Essays on the Impact of Information Asymmetry

Thesis by
Myungkoo Song

In Partial Fulfillment of the Requirements for the
degree of
Doctor of Philosophy

Caltech

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2017
Defended May 15, 2017

Myungkoo Song
ORCID: 0000-0001-8602-3825

ACKNOWLEDGEMENTS

I am especially grateful to Jakša Cvitanić, Jean-Laurent Rosenthal and Michael Ewens for their guidance and encouragement. I have also benefited from conversations with Federico Echenique, Leeat Yariv, Lawrence Jin and Omer Tamuz.

I would also like to thank my colleagues, Marcelo A. Fernández, Tatiana Mayskaya, Lucas Núñez, Welmar E. Rosado-Buenfil, Li Song, Pengfei Sui and Jun Zhang for their support and friendship.

Last but not least, I dedicate this thesis to my family and parents.

ABSTRACT

This dissertation consists of three essays focusing on how information asymmetry affects agents' behavior across different environments. The first essay characterizes the optimal contract when a firm can employ two incentive schemes, promotion and pay for performance, simultaneously (Chapter 2). In the second essay, I study how information asymmetry can lead a firm to choose a less profitable short-term over a more profitable long-term project (Chapter 3). The other essay analyzes a career choice problem when agents have private information about their ability (Chapter 4).

Chapter 2 presents the effect of information asymmetry on executive pay structure to examine the cause of the rise in CEO compensation and wage inequality between CEO and other executives. To analyze the effect of the interaction of two incentive schemes, promotion and pay for performance, on CEO compensation and within-firm wage inequality, I embed a pay for performance framework into a tournament structure. The model shows that when CEO and managers contribute to a firm's output independently, it is optimal for the firm to provide the CEO a compensation far beyond her reservation value in order to provide promotion incentives for managers. However, I find that the promotion incentive motive can disappear if there is interdependency between the CEO's and managers' outputs. In this case, the main purpose of a high CEO compensation is to induce the CEO to exert effort. The tension between incentives for CEO and managers makes it difficult to interpret the meaning of within-firm wage gap. As a possible solution, this paper suggests the use of CEO's base salary to identify which incentive factor is driving the pay gap.

In Chapter 3, I study the optimal contract problem when a firm faces a long-term project. I consider a long-term project as one that requires an indefinite amount of time to complete its objective. I assume that the long-term project generates profits once it is accomplished. Using a continuous-time moral hazard model, I characterize the incentive compatibility condition in a relatively general contracting space. Moreover, I find a unique optimal contract under a restricted contracting space which consists of the two components: the termination level and the completion payment. The firm might invest in a short-term project: one that generates an instantaneous profit to the firm without any effect on the future, as analyzed by DeMarzo and Sannikov (2006). Comparison of optimal contracts for long and short-term projects provides an interesting insight to managerial short-termism: the

firm, not the agent, could prefer a short-term project to a long-term project if there is a moral hazard problem.

Chapter 4 analyzes the role of asymmetry information on one's career choice. I examine how people choose their career when they do not know ability of the rest of the applicant pool. The goal is to understand labor supply in the markets where ability is widely distributed. In particular, I consider a situation where there are two exclusive labor markets and the upper and lower bounds of one market's payoffs are both higher than those of the other market. Under the market setting, agents decide which market to participate in. I find that the symmetric Bayesian Nash equilibrium of this problem is unique. In the equilibrium, agents are divided into two groups according to their ability. Members of the high ability group use a pure strategy and only apply to the more desirable market. Members of the low ability group apply to both markets with positive probability.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

Chapter 1

INTRODUCTION

This dissertation consists of three essays focusing on how information asymmetry affects agents' behavior across different environments. Asymmetric information is an important source of friction in many markets and it can also lead a firm to raise its executive compensation and engage in short-termism. The first essay characterizes the optimal contract when a firm can employ two incentive schemes, promotion and pay for performance, simultaneously (Chapter 2). This analysis helps us understand which incentive schemes contributed to the rise in CEO compensation. In the second essay, I study how information asymmetry can lead a firm to choose a less profitable short-term over a more profitable long-term project (Chapter 3). The other essay analyzes a career choice problem when agents have private information about their ability (Chapter 4). This information asymmetry yields a tension between pursuing a more desirable career and the probability of employment since agents cannot identify her or his ranking which determines their career path.

Chapter 2 presents the effect of information asymmetry on executive pay structure to examine the cause of the rise in CEO compensation and wage inequality between CEO and other executives. To analyze the effect of the interaction of two incentive schemes, promotion and pay for performance, on CEO compensation and within-firm wage inequality, I embed a pay for performance framework into a tournament structure. The model shows that when CEO and managers contribute to a firm's output independently, it is optimal for the firm to provide the CEO a compensation far beyond her reservation value if there are many managers who compete for the CEO position. In particular, promotion incentives for managers generate the high CEO pay and the wage gap between CEO and managers. However, I find that the promotion incentive motive can disappear if there is interdependency between the CEO's and managers' outputs. In this case, the main purpose of a high CEO compensation is to induce the CEO to exert effort, not to provide promotion incentives for managers. The tension between incentives for CEO and managers makes it difficult to interpret the meaning of within-firm wage gap. As a possible solution, this paper suggests the use of CEO's base salary to identify which incentive factor is driving the pay gap.

In Chapter 3, I study the optimal contract problem when a firm faces a long-term project. In many cases, a firm cannot pursue several projects simultaneously because of limited resources. Particularly, the firm might need to choose between long and short-term projects. I consider a short-term project as one that generates an instantaneous profit to the firm without any effect on the future, as analyzed by DeMarzo and Sannikov (2006). The firm can also invest in a long-term project: one that requires an indefinite amount of time to complete its objective. I assume that the long-term project generates profits once it is accomplished. Using a continuous-time moral hazard model, I characterize the incentive compatibility condition in a relatively general contracting space. Moreover, I find a unique optimal contract under a restricted contracting space which consists of the two components, the termination level and the completion payment. Comparison of optimal contracts for long and short-term projects provides an interesting insight to managerial short-termism: the firm, not the agent, could prefer a short-term project to a long-term project if there is a moral hazard problem.

Chapter 4 analyzes the role of asymmetry information on one's career choice. I examine how people choose their career when they do not know ability of the rest of the applicant pool. The goal is to understand labor supply in the markets where ability is widely distributed. Specifically, I consider a situation where there are two exclusive labor markets with different payoff distributions, but people share common preferences over possible career paths. More precisely, the upper and lower bounds of one labor market's payoffs are both higher than those of the other market. Under the market setting, agents decide which market to participate in. The presence of private information produces a tension between the probability of being employed and the expected payoff. I find that the symmetric Bayesian Nash equilibrium of this problem is unique. In the equilibrium, agents are divided into two groups according to their ability. Members of the high ability group use a pure strategy and only apply to the more desirable market. Members of the low ability group apply to both markets with positive probability.

Chapter 2

HIGH CEO COMPENSATION: INCENTIVES FOR CEO OR MANAGERS?

2.1 Introduction

The provision of incentives has been considered as an important factor that contributed the rise in executive compensation since 1970s. In particular, researchers have focused on two incentive schemes: promotion and pay for performance.[1] However, each incentive scheme has a drawback in explaining the trend in executive compensation. First, pay for performance schemes have been criticized because theoretical analysis predicts a reduction in base salary with the rise in incentive pay, which contradicts the empirical evidence. On the other hand, promotion incentives do not capture the weak correlation between executive compensation and firm size. This paper studies whether a hybrid incentive scheme that includes both pay for performance and promotion can account for these empirical findings.

This paper contributes to the theoretical literature on executive compensation based on a principal-agent framework. The moral hazard literature attributes high CEO compensation to the need to provide incentives. There are two incentive schemes relevant to this problem. The first incentive scheme is pay for performance as studied by Grossman and Hart (1983) and Mirrlees (1999). These works suggest that firms need to provide high compensation to a CEO in order to make him/her exert more effort. The second incentive scheme is internal promotion associated with a rank-order tournament model proposed by Lazear and Rosen (1981). They argue that firms set up internal labor markets and promote workers who do well. In this setting, high wages at the top of the firms can make lower level employees work harder to increase their promotion probability. Following the analysis of Lazear and Rosen (1981), the CEO may well have a wage that is much higher than his or her contribution to the firm because it acts as the prize for lower level managers. The main difference between these two incentive schemes is that the former is based on absolute performance while the latter depends on relative performance. In this paper, I analyze how the two incentive schemes interact within a firm to understand

[1] For example, Bognanno (2001) shows that CEO compensation increases with the number of competitors for the position. Also, Frydman and Jenter (2010) illustrate that the increase in incentive pay such as option compensation significantly contribute to the rise in CEO compensation.

when a firm might want to combine the two in a hybrid incentive scheme: a contract based on both absolute and relative performance.

To study how the hybrid incentive scheme works, it is important to understand which incentive scheme is relevant to each agent in a firm. In terms of incentives, the pay for performance scheme works for both the CEO and other top executives (from now on, I call these executives “managers") in the same way. However, the promotion based incentive scheme only matters for managers who have a possibility of being promoted. On the other hand, with respect to compensation level, promotion incentives affect two types of agents, CEO and managers, in opposite ways. When other conditions are fixed, raising promotion incentives increases CEO compensation, but decreases managers’ compensation. The pay for performance scheme yields a higher compensation to both agents if a firm wants them to exert a higher effort. Given that raising promotion incentives impacts managers and CEOs in different directions, it is natural to ask how the two types of incentive payments constitute the executives’ total compensation.

In order to answer this question, I consider a model with a hierarchical structure in which an infinitely lived risk-neutral firm hires N + 1 agents (one CEO and N managers) in each period. For tractability, I assume that the firm offers a contract based on agents’ positions. That is, there are only two types of contracts which do not depend on time: one for CEO and one for managers.[2] Each agent conducts her/his own tasks and produces an outcome that depends on her/his effort level.[3] In the benchmark model, I assume that each task affects the firm’s profit independently. That is, there is no complementarity between agents. All outcomes are observable and contractible. This structure closely follows Grossman and Hart (1983). In addition, I introduce a tournament structure to add promotion based incentives. If there are qualified managers for the CEO position according to a promotion rule, the firm fills the position by internal promotion when the previous CEO leaves.

The most important contribution of this paper is that I analyze how the two incentive schemes interact in a dynamic environment similar to Rogerson (1985a). In particular, the reward for promotion is not an immediate monetary remuneration, but a position that can provide high compensation in the future. Under this distinctive structure, the analysis of the model shows that firms can rationally employ the

[2]This is a stronger assumption than the commitment assumption in Lazear and Rosen (1981) However, this makes the firm’s problem simple when I consider a complex situation in Section 2.7

[3]Throughout the rest of this paper, I use she as a personal pronoun for a CEO and he for a manager.

two incentive schemes simultaneously. Therefore, it is important to understand the interaction between the two incentive tools within a firm as well as the effect of each scheme.

An extensive literature investigates the two incentive schemes separately.[4] Since the seminal work of Lazear and Rosen (1981), promotion incentives based on the tournament theory have been extensively studied in various settings.[5] Although the tournament theory has broad applications[6] , unlike athletic tournaments, it requires caution to apply the theory to a firm's internal organization problem. The distinction of the firm's internal labor market is that the winner of the tournament stays in the firm as a worker to whom the firm needs to provide incentives. Ke, Li, and Powell (2014) and Goel and Thakor (2008) reflect this idea in their papers. However, Ke, Li, and Powell (2014) only consider a risk-neutral agent and one of two incentive schemes, while Goel and Thakor (2008) do not explicitly analyze managers' compensation. Different from these works, I investigate the role of agents' risk aversion and the relationship between managers' and CEO's compensations in a hierarchical structure.

On the other hand, the pay for performance literature investigates how the optimal level of incentives changes according to numerous factors affecting the contract.[7] For example, Baker and Hall (2004) examine how the measure of CEO incentives could change according to the effect of a CEO on the firm's value, and Prendergast (2002) studies the effect of uncertainty on incentives and shows that a positive relationship between uncertainty and incentives can arise. In this paper, I examine the relationship between CEO incentives and the internal labor market.

A number of papers compare the two incentive schemes from a theoretical perspective. However, most papers do not consider a firm that uses both schemes together though most of firms implement an incentive system based on both schemes in practice. For example, Nalebuff and Stiglitz (1983) and Mookherjee (1984) investigate conditions under which one of the incentive schemes is optimal. That is, these papers consider the two incentive schemes as substitutes. A notable exception is Ekinci and Waldman (2015), where they combine pay for performance and promotion incentives. Different from my paper, Ekinci and Waldman (2015) use a

[4]See Prendergast (1999).

[5]See Lazear and Oyer (2012).

[6]Konrad (2009) and Dechenaux, Kovenock, and Sheremeta (2015) review a broad literature on contests.

[7]For a comprehensive review, see Bolton and Dewatripont (2005).

market-based tournament theory, where the reward of promotion is determined by expected wage offers of competing firms.[8] Therefore, their paper focuses on how the market wage is determined, while I illustrate that a CEO can earn more than her market wage because of the internal labor market.

The formal analysis yields the following results. First, the firm provides compensation beyond the CEO's reservation value when there are enough candidates for the promotion and if agents are risk-averse. That is, the CEO's participation constraint can optimally be slack to maximize the firm's profit. I interpret the gap between the CEO's expected utility under the optimal contract and the CEO's reservation value as the promotion incentive because this utility gap, rather than the monetary gap, is the reward from the agent's point of view.[9] The non-binding participation constraint implies that a large wage gap between a CEO and managers could be optimal for the firm in order to incentivize managers below the CEO even when the firm employs an absolute performance based incentive scheme. Intuitively, when there is a small number of managers, the profit generated by the CEO accounts for a substantial part of the firm's profit. Thus, the firm does not want to raise promotion incentives beyond the CEO's reservation value since this makes it more costly for the firm to provide incentives to her (particularly when CEO is risk-averse). However, when the firm is large, the CEO's contribution is marginal compared to the profit created by managers. Thus, the firm uses its CEO position as the prize to the winner of internal competition rather than an output producer, which leads to the rise in CEO compensation and larger wage gap.

Second, for a fixed number of managers, analyzing the comparative statics gives interesting results for a CEO's utility and managers' wages. I show that there is a negative relationship between the promotion incentive and the CEO's effort level. In other words, the optimal contract provides a higher expected utility to the CEO when the firm requires less effort from her. The contract could even offer her a higher expected compensation. This result captures how promotion incentives interact with CEO incentives. I will discuss this issue in more detail later. The other important comparative static is the relationship between promotion incentives and managers' reservation value. The positive correlation between them illustrates that the two extreme allocations between the two incentive schemes for managers can happen. That is, the firm can only adopt the promotion incentive without pay for

[8] See Waldman (2013) for a survey of the literature on the two tournament theories.

[9] Note that the monetary gap can widen if the firm requires a higher effort from the CEO since agents are risk-averse.

performance scheme if managers' reservation value is high enough. The opposite situation (no promotion incentives) can occur when the reservation value is very low. In particular, I expect that the former case can provide an explanation of why performance-tied compensation is not prevalent in the workplace as illustrated in Lemieux, MacLeod, and Parent (2009).[10] Firms may optimally mute the absolute performance channel even though they consider this incentive scheme. Namely, introducing relative performance-based incentives can make managers' compensation less dependent (even independent) on their performance.

As an extension of the benchmark model, I analyze a situation where there is complementarity between the CEO's task and managers' outputs. Specifically, I consider the CEO's task with a multiplication effect on the sum of other managers' outputs. This specification is closely related to Baker and Hall (2004), where the marginal product of the CEO's marginal effort is increasing in the size of the firm.[11] This extension indicates that the level of promotion incentive depends on the role of a CEO. Under the multiplication specification, the promotion incentive can disappear. That is, if the firm hires enough managers, the slackness of the CEO's individual rationality constraint can alter to the binding constraint. Because the CEO's marginal effort is much more valuable for the firm as the number of managers increases, the firm prefers to focus on the CEO's absolute performance based incentive rather than the managers' relative performance based incentive.

The extension provides a unique implication regarding CEO compensation. Although the benchmark model and the extension can produce a positive link between compensation and firm size, they make a different prediction about the optimal compensation structure. That is, the two specifications link the rise in CEO compensation to different channels. When the CEO's marginal productivity is independent of firm size, compensation grows with the size of the firm because of the raised promotion incentive. Under the alternative specification, however, the higher compensation stems from enhanced pay for performance incentives for the CEO because her demanded effort grows with the size. I connect this implication with a measure of promotion incentives in section 2.10.2.

In addition, this paper contributes to the literature on the trend in executive compensation. The hybrid incentive scheme model predicts that CEO compensation remains stable and the wage gap between lower-level executives and CEO does not

[10] Macera (2016) provides a behavioral explanation based on loss aversion on this problem.
[11] They use the market value of the company or firms' sales as the measure of the size of the firm.

expand until the size of the firm reaches the point that it begins to raise its CEO compensation to induce managers' effort. The main conditions for the result are executives' risk aversion and the independence of agents' outputs. Therefore, if these conditions are satisfied, CEO compensation has a non-monotonic relationship with firm size measured by the number of candidates for promotion. [12] In other words, the compensation remains at the same level although the number of competitors increases since CEO's participation constraint binds. However, beyond a certain point, the firm wants to raise its CEO compensation beyond her reservation value in order to provide incentives to its managers. Also, the wage gap widens when the firm starts providing compensation to the CEO beyond her reservation value but not before that point. These results suggest a possible explanation for two empirical facts shown in Frydman and Jenter (2010). First, they find a non-monotonic increase of executive pay: the rapid growth in executive pay only started in the mid-1970s. Also, their results show that the compensation gap between CEO and other top executives rapidly grown during the past 30 years but not before 1980. I discuss this issue in more detail in section 2.10.1.

The rest of the paper is organized as follows. Section 2.2 presents the basic model and the promotion rule. Before I introduce the formal model, I present a preview of the tension between promotion incentives and pay for performance in Section 2.3. In Section 2.4, I simplify the firm's problem and provide the primitive trade-off of the problem. In Section 2.5, I consider the problem without any agency problem to illustrate the effect of information asymmetry. In Section 2.6, I discuss the main properties of the firm's problem demonstrating the effect of promotion on CEO's compensation. In Section 2.7, I introduce several dynamics into the basic model and analyze the effect of them on CEO compensation. Section 2.8 examines how the role of a CEO impact on promotion incentives. I compare two promotion rules in Section 2.9 to demonstrate the benefit of external CEO recruitment. I discuss the important implications of this paper in Section 2.10. Section 2.11 contains concluding remarks.

2.2 The Benchmark Model

2.2.1 Firm structure and executives

I start with a simple discrete time model, where an infinitely lived risk-neutral firm with a discount factor $\delta \in (0,1)$ maintains its employment structure. In particular,

[12] For instance, Acs and Audretsch (1987) use the number of employees as a measure of firm size.

the profit maximizing firm employs one CEO and N managers over time. The firm hires the managers from a labor market every period, but it can promote one of the previous managers to be the next CEO according to a promotion rule. If there is no manager satisfying the promotion criteria, the firm hires a CEO from an external labor market. Throughout most of the paper, I consider the case where the CEO leaves the firm or retires after one period, and a manager remains in the firm if he is promoted to CEO. Otherwise, managers leave the firm after one period. I consider a moral hazard situation where agents' efforts (e) are not observable to the firm and incur a cost to agents, $g(\cdot)$. Agents are risk-neutral or risk-averse with an additively separable (von Neumann-Morgenstern) utility function $U(C, e)$, where C is consumption[13] , in each period satisfying the following assumption:

Assumption 1 *Agents' utility function is of the form*

$$U(C, e) = u(C) - g(e),$$

where $u'(\cdot) > 0, u''(\cdot) \leq 0$, $u(\cdot)$ *is defined over the real interval* $(\underline{C}, \infty)$*, and* $g(0) = 0$*,* $g'(\cdot) \geq 0$*,* $g''(\cdot) > 0$ *over* $e \in [0, 1]$*. Also, there exists* $\widehat{C} \in (\underline{C}, \infty)$ *such that* $u(\widehat{C}) > 0$ *and* $\lim_{C \downarrow \underline{C}} u(C) = -\infty$.

This assumption comes from Grossman and Hart (1983) and Rogerson (1985b), which guarantees the existence of solution and validates the first-order approach. Additionally, I assume that agents do not discount the future, and a CEO and managers have an outside option $\underline{U}_C \geq 0$, and $\underline{U}_M$, respectively, if they do not accept the offer from the firm. Also, in their second period, managers are assumed to obtain a reservation utility, normalized to zero, if they leave the firm. Hence, the positivity assumption on $\underline{U}_C$ assures that promotion is beneficial for managers. Agent i does an independent task X_i, $i = 0, \ldots, N$, which can end in a good or bad outcome.[14] The probability of good outcome depends on the agents' choice of effort level e_i by the function $s(e_i)$ in the following way:

$$X_i(e) = \begin{cases} \mathcal{G}_i & \text{with probability } s(e) \\ \mathcal{B}_i & \text{with probability } 1 - s(e) \end{cases}.$$

[13] I do not consider the possibility of private saving in this paper.

[14] Agent 0 represents CEO.

I assume that $s(e)$ is a linear function of the effort level e.[15] That is, $s(e)$ satisfies

$$s(e) = \alpha + \beta e,$$

where $e \in [0, 1]$ is the agent's effort level. For two parameters α and β, I assume that $\alpha \geq 0$, $\beta > 0$, and $\alpha + \beta \leq 1$, which guarantees that $s(e) \in [0, 1]$.

I assume that each manager's task has an identical effect on the firm's output. That is, $\mathcal{G}_i = \mathcal{G}_j$ and $\mathcal{B}_i = \mathcal{B}_j$ for $i, j \in \{1, 2, \ldots, N\}$. After observing managers' outputs, the firm makes the promotion decision according to the following rule:

(**Promotion Rule 1**) *Among managers whose outcome is good ($\mathcal{G}$), the firm chooses one manager randomly for promotion. If there are no such managers, the firm hires a CEO from an external labor market.*

This promotion rule can be understood as a situation where a firm sets up a certain requirement such that it considers managers who satisfy the requirement as candidates for internal promotion. Note that this promotion rule is different from the rule considered in Lazear and Rosen (1981), where they use the following one:[16]

(**Promotion Rule 2**) *The firm promotes the best manager, whose outcome could be good ($\mathcal{G}$) or bad ($\mathcal{B}$). If more than one agent makes the best outcome, the firm chooses one of them randomly.*

The difference between two promotion rules is the possibility of external hiring. When all managers' outcomes are bad, the firm appoints its CEO from an external labor market under promotion rule 1 while it still uses the internal labor market under promotion rule 2.[17] I show that promotion rule 1 could be preferred by firms to promotion rule 2 if they do not know a managers' ability and want to promote a more talented candidate in section 2.9.

While the CEO's job can also have a good or bad outcome, the effect on the output can be different from that of managers. For brevity, I denote CEO's good and bad outcomes by $\mathcal{G}_C$ and $\mathcal{B}_C$, respectively, with manager's outcomes denoted by $\mathcal{G}_M$ and $\mathcal{B}_M$. Also, the marginal productivity of the CEO's effort depends on the firm's operational structure. As the benchmark model, I study a firm where the CEO's

[15] This is not so much restrictive. With **Assumption 1**, this condition embraces any concave function $s(e)$ with $s'(e) > 0$.

[16] Strictly speaking, there is a difference. Lazear and Rosen (1981) consider a continuous outcome space. This yields no ties with probability 1.

[17] As Kale, Reis, and Venkateswaran (2009) show, the firm recruits its CEO from outside as well although internal promotion is more common.

task and managers' jobs are independent of each other. That is, CEO and managers are substitutable in terms of the firm's profit. I relax this assumption later on.

2.2.2 Contracts

In order to make the problem tractable, I assume that contracts only depend on agent's positions regardless of time as well as internal and external hiring.

Assumption 2 *The firm offers contracts based on agents' positions.*

This assumption implies that a firm offers the same contract to a future CEO as the current one. That is, in every period, the firm offers a contract (e_C, W_C^G, W_C^B) to the CEO choosing each component in order to maximize its profit. In the contract, e_C represents the firm's recommended effort level while W_C^O is the wage if the CEO's output turns out to be $O \in \{G,\ B\}$. On the other hand, the firm provides a contract $(e_{Mi}, W_{Mi}^G, W_{Mi}^B)$ to manager i, where the role of each component is the same as that of the CEO contract. Note that the firm does not need to specify the prize for the winner of promotion since **Assumption 2** allows current managers to know what they will get if promoted to CEO in the next period. This is in line with the tournament literature in the sense that firms can commit to the prize for the winner. For simplicity, I focus on a symmetric equilibrium in this paper. That is, the firm requires the same effort e_M from every manager. Therefore, the firm offers the same contract (e_M, W_M^G, W_M^B) to all managers. After I solve this simple benchmark model, I extend the model by including job security issues, complex operational structure, and heterogeneous managers.

2.3 The Effect of Promotion Incentive on Agents' Wages

In this section, before I move to a general problem, I study how promotion incentive affects agents' compensation. Since each output is binary, the two types of agents' ex-ante utility can be described by the following two terms for each agent:

$$\text{Manager} \begin{cases} u(W_M^G) + P(\mathbf{e}_{-M})\mathcal{V} & \text{when } X = \mathcal{G}_M \\ u(W_M^B) & \text{when } X = \mathcal{B}_M \end{cases} \quad \text{and,}$$

$$\text{CEO} \begin{cases} u(W_C^G) & \text{when } X = \mathcal{G}_C \\ u(W_C^B) & \text{when } X = \mathcal{B}_C \end{cases},$$

where $P(\mathbf{e}_{-M})$ represents the probability that a manager is promoted to CEO when he gets a good outcome and other managers' effort levels are $\mathbf{e}_{-M}$. $\mathcal{V}$ is the CEO's expected utility, which the manager will get if promoted in the next period. Note that the value of $\mathcal{V}$ determines the power of promotion incentive since this is the benefit the winner of the tournament enjoys.

How does this promotion incentive affect the incentive based on absolute performance? Note that for agents' incentives, the difference between two utility levels is all that matters. Hence, when the firm requires a certain effort from its executives, the firm chooses wages fixing the value of these two gaps at some positive values: 1) $\left[u(W_M^G) + P(\mathbf{e}_{-M})\mathcal{V}\right] - u(W_M^B)$ and 2) $u(W_C^G) - u(W_C^B)$.

From the manager's perspective, raising the promotion incentives makes the firm reduce incentives based on pay for performance, that is, the gap between W_M^G and W_M^B. This leads to a decrease in managers' compensation. However, for the CEO incentives, higher promotion incentives yield a bigger gap between W_C^G and W_C^B if the CEO is risk-averse. This makes the compensation for the CEO rise.

Therefore, there is a tension between incentives for managers and CEO as well as a trade-off between the two types of agents' wages: higher promotion incentives make it more difficult for the firm to incentivize its CEO but easier for managers. This feature is captured by the movement of the gap between wages associated with a good and bad outcome. It is worth mentioning that this incentive trade-off rises because the CEO is risk-averse. If the CEO is risk-neutral, the wage gap between W_C^G and W_C^B is always a constant unless the firm requires a different effort level.

Hence, adjusting promotion incentives affects two types of agents' absolute performance based incentive schemes in different ways. In the following sections, I analyze how the firm optimally sets up the level of promotion incentives and how this decision changes when the firm's contracting environment alters.

2.4 Formulation of the Firm's Problem

In this section, I explicitly state the firm's problem and simplify it. First, I consider the agents' problem. The CEO's utility maximization problem is straightforward. For a given compensation scheme (W_C^G, W_C^B), the CEO chooses an effort level e_C maximizing her expected utility

$$s(e_C)u(W_C^G) + (1 - s(e_C))u(W_C^B) - g(e_C).$$

The strict convexity of $g(\cdot)$ guarantees a unique solution to the CEO's problem. Notice that this problem does not depend on other agents' effort choices.

On the other hand, the managers' problem depends on other managers' effort choices $\mathbf{e}_{-M}$ in the same cohort. Since I consider a symmetric equilibrium, it is enough to focus on $\mathbf{e}_{-M}$ such that $\mathbf{e}_{-M} = (e_{-M}, e_{-M}, \ldots, e_{-M}) \in [0,1]^{N-1}$. Therefore, for a compensation scheme $(W_C^G, W_C^B, W_M^G, W_M^B)$, and an effort level of $\mathbf{e}_{-M}$, the managers' problem can be rewritten as

$$\max_{e_M} s(e_M)u(W_M^G) + (1 - s(e_M))u(W_M^B) - g(e_M) + s(e_M)P(\mathbf{e}_{-M})V_C,$$

where

$$P(\mathbf{e}_{-M}) = \frac{1 - (1 - s(e_{-M}))^N}{Ns(e_{-M})}$$

represents the conditional probability that a manager will be promoted to the next period's CEO when he achieves a good outcome and other managers' effort level $\mathbf{e}_{-M}$ is given.[18] Also,

$$V_C = s(e_C)u(W_C^G) + (1 - s(e_C))u(W_C^B) - g(e_C)$$

is the expected utility a manager will get if he is promoted to CEO. Recall that managers know the contract they will get if promoted under **Assumption 2**.

Under the fixed employment structure, **Assumption 2**, and the symmetric equilibrium condition, the firm's objective is to offer contracts (e_C, W_C^G, W_C^B) and (e_M, W_M^G, W_M^B) to CEO and managers that maximize its profit under the incentive compatibility and individual rationality constraints. Mathematically, the problem is:

$$\max_{\{(e_C, W_C^G, W_C^B), (e_M, W_M^G, W_M^B)\}} \sum_{t=1}^{\infty} \delta^{t-1} \Big\{ s(e_C)(\mathcal{G}_C - W_C^G) + (1 - s(e_C))(\mathcal{B}_C - W_C^B) + N \left[s(e_M)(\mathcal{G}_M - W_M^G) + (1 - s(e_M))(\mathcal{B}_M - W_M^B) \right] \Big\}$$

subject to

[18] The derivation of this equation is found in the Appendix.

$$s(e_C)u(W_C^G) + (1 - s(e_C))u(W_C^B) - g(e_C) \geq \underline{U}_C \ (IR_C) \tag{2.1}$$

$$s(e_M)u(W_M^G) + (1 - s(e_M))u(W_M^B) - g(e_M) + s(e_M)P(\mathbf{e}_{-M})V_C \geq \underline{U}_M \ (IR_M) \tag{2.2}$$

$$e_C \in \arg\max_{\hat{e}} s(\hat{e})u(W_C^G) + (1 - s(\hat{e}))u(W_C^B) - g(\hat{e}) \ (IC_C) \tag{2.3}$$

$$e_M \in \arg\max_{\hat{e}} s(\hat{e})u(W_M^G) + (1 - s(\hat{e}))u(W_M^B) - g(\hat{e}) + s(e_M)P(\mathbf{e}_{-M})V_C \ (IC_M). \tag{2.4}$$

Since contracts are not time-dependent, the above problem is equivalent to solve the following problem:

$$\begin{aligned} \max_{\{(e_C, W_C^G, W_C^B), (e_M, W_M^G, W_M^B)\}} \quad & s(e_C)(\mathcal{G}_C - W_C^G) + (1 - s(e_C))(\mathcal{B}_C - W_C^B) \\ & + N\left[s(e_M)(\mathcal{G}_M - W_M^G) + (1 - s(e_M))(\mathcal{B}_M - W_M^B)\right] \\ & \text{subject to the four constraints, (2.1), (2.2), (2.3), and (2.4).} \end{aligned}$$

From now on, I call this the firm's problem. It is worth mentioning the role of **Assumption 2** in the simplification of the firm's problem. This allows one to focus on a repeating part by restricting the firm to offer the same contract to the CEO. In section 2.7, I investigate a more complex model with dynamics in a simplified form with the same argument.

2.4.1 The basic characterization of the firm's problem

In this section, I analyze some preliminary features of the firm's problem before examining its main properties.

First of all, by a similar argument with Grossman and Hart (1983), I can show that the firm's problem has a solution.

Lemma 1 *There exists a solution to the firm's problem.*

Proof 1 *All proofs are presented in the Appendix.*

In order to analyze agents' incentive compatibility constraint, I can apply the first-order approach from Rogerson (1985b). The two first-order conditions yield the following results:

$$u(W_C^G) = u(W_C^B) + \frac{g'(e_C)}{h'(e_C)}, \text{ and}$$
$$u(W_M^G) = u(W_M^B) + \frac{g'(e_M)}{h'(e_M)} - P(\mathbf{e}_{-M})V_C.$$

Since the firm can control (W_M^G, W_M^B) without affecting the CEO's problem, the manager's individual rationality constraint must bind to maximize its profit. In particular, the firm can decrease W_M^B such that the managers' incentive compatibility constraint is satisfied at the given effort level e_M when W_M^G decreases for a given fixed value V_C. This means that the firm can increase its profit if the managers' individual rationality constraint does not bind.

Lemma 2 *Managers' individual rationality constraint binds.*

However, the same logic cannot be applied to the CEO's individual rationality constraint since adjusting W_C^G and W_C^B inevitably affects V_C, which directly enters into the managers' problem. Reducing a CEO's expected utility yields a higher W_M^G for given e_M and W_M^B, which could decrease the firm's profit. This observation makes it difficult to analyze the properties of the firm's problem. In the next section, I consider a modified method circumventing this obstacle.

2.4.2 CEO's individual rationality constraint

In this section, I examine a problem where the CEO's individual rationality constraint binds at a certain value in order to indirectly solve the firm's problem. In particular, I consider the following problem:

$$F(\mathcal{V}) \equiv \max_{\{(e_C, W_C^G, W_C^B),(e_M, W_M^G, W_M^B)\}} s(e_C)(\mathcal{G}_C - W_C^G) + (1 - s(e_C))(\mathcal{B}_C - W_C^B)$$
$$+ N\left[s(e_M)(\mathcal{G}_M - W_M^G) + (1 - s(e_M))(\mathcal{B}_M - W_M^B)\right]$$

subject to

$$s(e_C)u(W_C^G) + (1 - s(e_C))u(W_C^B) - g(e_C) = \mathcal{V} \ (IR_C) \tag{2.5}$$

$$s(e_M)u(W_M^G) + (1 - s(e_M))u(W_M^B) - g(e_M) + s(e_M)P(\mathbf{e}_{-M})\mathcal{V} = \underline{U}_M \ (IR_M) \tag{2.6}$$

$$e_C \in \arg\max_{\hat{e}} s(\hat{e})u(W_C^G) + (1 - s(\hat{e}))u(W_C^B) - g(\hat{e}) \ (IC_C) \tag{2.7}$$

$$e_M \in \arg\max_{\hat{e}} s(\hat{e})u(W_M^G) + (1 - s(\hat{e}))u(W_M^B) - g(\hat{e}) + s(e_M)P(\mathbf{e}_{-M})\mathcal{V} \ (IC_M) \tag{2.8}$$

for $\mathcal{V} \in [0, \infty)$. The difference from the original firm's problem is that the CEO's individual rationality constraint binds at a positive value $\mathcal{V}$. Since both individual rationality constraints are binding now, the compensation scheme $(W_C^G, W_C^B, W_M^G, W_M^B)$ can be expressed by functions of (e_C, e_M):

$$\begin{aligned}
u(W_C^G) &= \mathcal{V} + g(e_C) + (1 - s(e_C))\frac{g'(e_C)}{h'(e_C)} \\
u(W_C^B) &= \mathcal{V} + g(e_C) - s(e_C)\frac{g'(e_C)}{h'(e_C)} \\
u(W_M^G) &= \underline{U}_M + g(e_M) + (1 - s(e_M))\frac{g'(e_M)}{h'(e_M)} - P(\mathbf{e}_{-M})\mathcal{V} \\
u(W_M^B) &= \underline{U}_M + g(e_M) - s(e_M)\frac{g'(e_M)}{h'(e_M)}.
\end{aligned}$$

These equations characterize the compensation scheme for a given effort level (e_C, e_M) and a parameter $\mathcal{V}$. However, the solution to this modified problem is not necessarily the same as that of the firm's original problem. One special case when the two problems have the same solution is $\mathcal{V} = \underline{U}_C$ and the CEO's individual rationality constraint binds.

In order to find the optimal compensation schemes, I exploit the modified problem. Suppose that $F(\cdot)$ is a strictly quasi-concave function. Then, I can solve the original problem indirectly using $F(\cdot)$. In particular, consider a maximization problem:

$$\max_{\mathcal{V}} F(\mathcal{V}).$$

Let $\mathcal{V}^*$ denotes $\inf \{\arg\max_{\mathcal{V}} F(\mathcal{V})\}$.[19] Then, the strict quasi-concavity ensures that the CEO's individual rationality constraint binds and the firm's expected profit

[19] From now on, I use the superscript star to denote the optimal variable.

is equal to $F(\underline{U}_C)$ if $\mathcal{V}^* \leq \underline{U}_C$. On the other hand, the CEO's individual rationality constraint does not bind and the firm's expected profit is equal to $F(\mathcal{V}^*)$ if $\mathcal{V}^* > \underline{U}_C$. Before analyzing this modified problem, I impose some parametric assumptions on $\underline{U}_M$.

Assumption 3 *When $N = 1$ and agents are risk-averse, $\underline{U}_M$ is such that*

$$\mathcal{V}^* = \inf\left\{\arg\max_{\mathcal{V}} F(\mathcal{V})\right\} > 0.$$[20]

Under this assumption, I can exclude a corner solution implying that promotion is strictly beneficial for a manager if a firm hires only one manager. Note that the manager who is promoted to CEO enjoys the expected utility $\mathcal{V}^*$ during the second period. Hence, I can interpret this value as the promotion incentive from the manager's point of view.[21]

2.5 The Effect of Promotion Incentive on Agents' compensation when the Effort is Observable

Before analyzing the firm's problem with incomplete information, I first study the solution without any agency problem. In this section, I focus on risk-averse agents and use **Promotion rule 2** not **Promotion rule 1** for simplicity.[22] Then, the firm's modified problem is reduced to:

$$\max_{e_C, e_M, W_C^G, W_C^B, W_M^G, W_M^B} s(e_C)(\mathcal{G}_C - W_C^G) + (1 - s(e_C))(\mathcal{B}_C - W_C^B)$$
$$+ N(s(e_M)(\mathcal{G}_M - W_M^G) + (1 - s(e_M))(\mathcal{B}_M - W_M^B))$$

subject to

$$s(e_C)u(W_C^G) + (1 - s(e_C))u(W_C^B) - g(e_C) = \mathcal{V} \quad (IR_C)$$
$$s(e_M)u(W_M^G) + (1 - s(e_M))u(W_M^B) - g(e_M) + \frac{1}{N}\mathcal{V} = \underline{U}_M \quad (IR_M).$$

[20]Since $\frac{\partial \mathcal{V}^*}{\partial \underline{U}_M} > 0$ and there exists $\underline{U}_M^*$ such that $F'(\mathcal{V}|\underline{U}_M^*)|_{\mathcal{V}=0} > 0$, there is $\underline{U}_M$ satisfying the condition.

[21]More formally, the promotion incentive is $\max\{\mathcal{V}^* - \underline{U}_C, 0\}$ since the firm has to offer $\underline{U}_C$ in order to hire a CEO. However, I simply consider $\mathcal{V}^*$ as the promotion incentive since $\underline{U}_C$ is a constant.

[22]There are two reasons. First of all, **Promotion rule 2** simplifies the analysis if there is no agency problem. Also, if there is no agency problem, the argument in Section 2.9 does not work since the adverse selection issue is also related to incomplete information.

Since there is no agency problem, incentive compatibility constraints are dropped. Moreover, because the agents are risk-averse while the firm is risk-neutral, the two individual rationality constraints can be rewritten as:[23]

$$u(W_C) - g(e_C) = \mathcal{V} \quad (IR_C)$$
$$u(W_M) - g(e_M) + \frac{1}{N}\mathcal{V} = \underline{U}_M \quad (IR_M),$$

where the fixed payments (W_C, W_M) are paid to the CEO and managers when they exert the required effort levels (e_C, e_M), respectively.

Therefore, for given effort levels (e_C, e_M), the optimal wage levels are determined by

$$u(W_C) = \mathcal{V} + g(e_C), \text{ and}$$
$$u(W_M) = \underline{U}_M + g(e_M) - \frac{1}{N}\mathcal{V}.$$

Also, the following two first order conditions characterize the optimal effort choice levels by the firm:[24]

$$\beta(\mathcal{G}_C - \mathcal{B}_C) = \frac{g'(e_C)}{u'(W_C)}, \text{ and}$$
$$\beta(\mathcal{G}_M - \mathcal{B}_M) = \frac{g'(e_M)}{u'(W_M)}.$$

Recall that a strict quasi-concavity of $F(\cdot)$ allows the use of the modified method in order to solve the firm's original problem. The following lemma confirms that this is the case when there is no information friction between the firm and agents.

Lemma 3 *$F(\mathcal{V})$ is strictly concave.*

Now, I analyze the property of the firm's problem using the modified one. The main concern is the CEO's expected utility level $\mathcal{V}^*$.

Proposition 1 *If agents are risk-averse, $\mathcal{V}^*$ strictly increases as N increases.*

In other words, it is optimal for the firm to offer a contract providing a higher expected utility to the CEO if the internal pool of candidates for promotion is getting bigger. Combining this result with **Lemma 3** yields the following result.

[23]This can be easily checked using Jensen's inequality.
[24]The second order conditions can be checked easily.

Corollary 1 *CEO's expected utility is an increasing function in N.*

This implies that a CEO can obtain utility beyond her reservation value if a firm is big enough.[25] Although this does not mean that the firm provides a higher compensation to the CEO, this is true when there is no agency problem.

Corollary 2 *The optimal CEO compensation W_C^* is an increasing function in N.*

However, the following proposition illustrates that there are two restrictions on the CEO's compensation level when agents' effort levels are contractible.

Proposition 2 *If every agent is risk-averse, $W_C^* = W_M^*$. Moreover, if $\mathcal{G}_C - \mathcal{B}_C = \mathcal{G}_M - \mathcal{B}_M$, then $\mathcal{V}^*$ is equal to $\frac{N}{N+1}\underline{U}_M$.*

The first restriction is that CEO and managers receive exactly the same level of compensation when there is no information asymmetry. This is true since the firm wants to reduce the total pay by equally distributing remuneration. According to this result, if there is no informational friction, there is no wage gap between CEO and managers.

Moreover, if CEO and managers have the same marginal productivity[26] , then CEO's expected utility is strictly bounded by the manager's reservation value. Recall that the firm offers the utility $\mathcal{V}^*$ to its CEO only when $\mathcal{V}^* \geq \underline{U}_C$. In other words, if CEO's reservation value is greater than managers', the CEO's individual rationality constraint always binds regardless of the number of managers. This implies that the CEO's reservation value should be less than managers' to obtain both the wage gap and a positive relationship between CEO compensation and the firm size, In the next section, I consider the asymmetric information case.

2.6 The Effect of Promotion on CEO Compensation under Moral Hazard

2.6.1 Preliminary - no promotion possibility

In this section, I briefly illustrate several properties of the solution to the firm's problem when there is no promotion possibility. Comparing the optimal contract

[25] Here, the size of a firm is determined by the number of managers or candidates for internal promotion.

[26] In this paper, agents' marginal productivity is measured by $\beta(\mathcal{G}_i - \mathcal{B}_i)$, where $i \in \{C, M\}$.

with and without promotion possibility can clarify the effect of promotion incentives on CEO compensation.[27] If there is no promotion possibility, the firm's problem turns into a simple moral hazard problem and the CEO's individual rationality constraint must bind.[28] Binding individual rationality constraint yields a limitation in explaining the trend in CEO compensation. If changes in the firm's contracting environment make it require higher effort from the CEO, the firm provides higher expected pay to its CEO and the ratio of the fixed pay to the incentive pay falls. That is to say, the theory predicts the fixed pay and the incentive pay should move in the opposite directions. In practice, however, the base salary, i.e., fixed pay for CEO has not fallen enough while the incentive pay has risen. In the following section, I illustrate how this prediction changes if a firm uses the two incentive schemes, absolute performance based and relative performance based compensation, together.

2.6.2 Risk-neutral agents

First, I consider risk-neutral agents. The following proposition says that it is optimal for the firm not to offer a contract providing a higher utility than CEO's reservation value to her if all agents are risk-neutral.

Proposition 3 *If every agent is risk-neutral, the CEO's individual rationality constraint binds. Specifically, the firm's profit is a strictly decreasing function in* $\mathcal{V}$.

This implies that promotion incentive scheme is dominated by pay for performance if agents are risk-neutral. According to Lazear and Rosen (1981), a compensation scheme based on relative ranking achieves the first best allocation if agents are risk-neutral, which is also true for the incentive scheme based on absolute performance. Since the promotion rule considered in this paper has a penalty[29] compared to the tournament structure in Lazear and Rosen (1981), the incentive based on promotion is less efficient than absolute performance-based incentives.[30] I analyze how this observation changes if agents are risk-averse as follows.

[27] The problem in this section is exactly the same as Grossman and Hart (1983).

[28] This is true with any finite number of possible outcomes when CEO's utility function is additively separable.

[29] There is a possibility of external hiring.

[30] If I adopt **Promotion rule 2**, the two schemes are perfect substitutes. However, under **Promotion rule 2**, pay for performance scheme still dominates promotion incentive scheme if managers discount the future.

2.6.3 The first stage

In this section, I consider a simple situation where the firm's desired effort levels are given exogenously. That is, the firm requires fixed effort levels $(e_C, e_M) \in (0,1) \times (0,1)$ from the CEO and managers, respectively. First of all, the following lemma guarantees that I can use the modified method to solve the firm's problem. Moreover, the strict concavity ensures that the solution of the firm's problem is unique.

Lemma 4 *If agents are risk-averse, $F(\mathcal{V})$ is strictly concave.*

The next question is how the CEO's expected utility changes when there is a competition among managers for promotion. First, I consider the effect of the number of managers on the CEO's expected utility, which determines the power of promotion incentives. This is important because the number of candidates for promotion might be a good proxy of the firm size. For example, Zabojnik and Bernhardt (2001) use the number of competitors for a promotion as the measure of the firm size. From now on, I use the number of managers N to refer the size of a firm.

Proposition 4 *If agents are risk-averse, $\mathcal{V}^*$ strictly increases as N increases.*

Again, this proposition and the previous lemma lead to the following result.

Corollary 3 *CEO's expected utility rises as N increases.*

Hence, the firm raises the promotion incentive, although the power of it decreases as the competition gets severe. From this result, one can confirm that the firm wants to divide payments between absolute and relative performance based compensation depending on the size of internal labor market. In other words, firms effectively employ a hybrid incentive scheme not just one of the two schemes. Moreover, if the CEO's required effort level does not change as the number of candidates varies, the rise in CEO's expected utility yields a higher expected compensation to CEO.

Corollary 4 *The expected compensation to CEO increases as N grows.*

Therefore, it is optimal for a firm to provide a higher compensation to the CEO when there are more candidates for internal promotion. Therefore, the increase in promotion incentives might play a significant role in the rise of CEO compensation. When it comes to the wage inequality between CEO and managers, I can assert that the size of the firm has a positive correlation with the expected wage gap between CEO and managers as well as CEO compensation if agents have the log utility function.

Corollary 5 *Suppose that agents have the log utility function. Then, the expected compensation gap between CEO and managers widens as N increases.*

Then, can the CEO compensation still have the upper bound of **Proposition 2**?[31] This question is important since if the compensation has the same upper bound, the CEO compensation is always lower than managers' one when the firm requires the same effort level. However, the following proposition indicates that this is not the case if there is a moral hazard problem.

Proposition 5 *When $0 < e_C \leq e_M < 1$, there is $\widehat{N}$ such that $\mathcal{V}^* > \underline{U}_M$ if $N > \widehat{N}$.*

Figure 2.1 numerically illustrates previous results. In this example, the CEO's individual rationality constraint binds when $\underline{U}_C = 3$, which is the same as the manager's reservation value in the example, and $N = 1, 2$, or 3. This means that the CEO's compensation level remains at the same level even when the firm size increases. Namely, the size of the firm and CEO compensation do not have a strictly monotonic relationship. This feature makes this model different from Gabaix and Landier (2008), where CEO payment monotonically moves with changes in the size of the firm.[32]

By comparing this result with **Proposition 2**, I derive two channels to explain why the firm increases promotion incentives as N grows. The first channel comes from the difference between two reservation values $\underline{U}_C$ and $\underline{U}_M$. If $\underline{U}_C$ is less than $\underline{U}_M$, by offering a higher compensation to the CEO than to the managers, the firm can reduce overall payouts to agents. The other channel is to reduce the gap between W_M^G and W_M^B by increasing the CEO's expected utility. Note that the inefficiency of information asymmetry arises when agents are risk-averse if the firm does not

[31] If I consider **Promotion rule 1** under the first-best case, it can be shown that $\mathcal{V}^*$ is strictly less than $\frac{N}{N+1}\underline{U}_M$.

[32] Gabaix and Landier (2008) use earnings of firms as a proxy of the size of a firm. Analyzing the relationship between the number of candidates and other measures of the firm size such as earnings, market values, and sales will be a topic for future research.

use internal promotion schemes. This is because the wage gap incurs an additional cost to the firm. Hence, the firm can partly remove the inefficiency by increasing promotion incentives, which reduces the wage gap. This effect gets stronger as N grows because promotion incentives affects N managers' wage gaps.

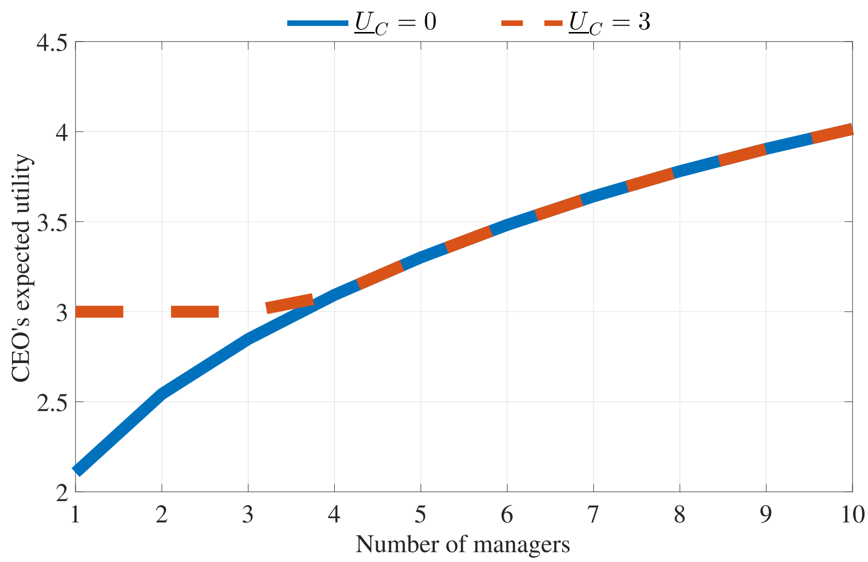

Figure 2.1: CEO's expected utility for $\underline{U}_C = 0$ (solid line) or $\underline{U}_C = 3$ (dashed line) when $\underline{U}_M = 3$. Other parameters are $\mathcal{G}_C = \mathcal{G}_M = 1{,}000$, $\mathcal{B}_C = \mathcal{B}_M = 0$, $\alpha = 0.05$, $\beta = 0.45$, $e_C = 0.35$, $e_M = 0.43$, $u(x) = \log(x)$ and $g(a) = -\frac{1}{a-1} - a - 1$.

As the next step, I examine how the level of promotion incentives behaves according to the adjustment of required effort levels. The following proposition shows how this changes.

Proposition 6 *Suppose that $\frac{u''(x)}{u'(x)^3}$ is a decreasing function in x. If agents are risk-averse, $\mathcal{V}^*$ strictly decreases as e_C increases while $\mathcal{V}^*$ strictly increases when e_M increases.*

One interesting result is that the firm provides a higher expected utility to the CEO when it requires less effort from her. This makes sense when CEO compensation is related to her incentives as well as promotion incentives for managers. This result gives an important implication about promotion incentives. Recall that the degree of promotion incentives is determined by the level of $\mathcal{V}^*$. Therefore, if the firm requires a high (low) effort from the CEO, it reduces (increases) the promotion incentives. This implies that if a change in contracting environments yields a growth in the CEO's marginal productivity, then the firm decreases the CEO's expected utility reducing promotion incentives.

The sufficient condition regarding agents' utility functions for **Proposition 6** holds for all CARA utility functions and CRRA utility functions with a coefficient of relative risk aversion higher than one half.

When the firm uses relative as well as absolute performance based incentive schemes, the gap between the managers' wage for a good and a bad performance could shrink to zero.[33] In other words, it can be optimal for a firm to focus on promotion incentives instead of performance tied compensation.

Proposition 7 *Assume that* $\lim_{x\to\infty} u'(x) = 0$. *Then, for a given* (N, e_C, e_M), *there is* $\widehat{\underline{U}}_M$ *such that*

$$(W_M^G)^* \leq (W_M^B)^*$$

if $\underline{U}_M \geq \widehat{\underline{U}}_M$.

Therefore, if people do not consider the possibility of promotion when they analyze executives' compensations they can misinterpret the incentives behind them. That is, although managers' compensation does not depend on their output, their wages have already reflected incentives through the promotion possibility. **Figure 2.2** illustrates this result. As the managers' reservation value increases from 8 to 9, the situation of $(W_M^G)^* \leq (W_M^B)^*$ happens when $N < 4$. It is worth mentioning that this result does not mean that managers have less utility when they make a good outcome rather than a bad outcome. If managers achieve a good outcome, they obtain utility through two channels: 1) utility from the first period's wage, W_M^G, in the first period and 2) utility from the promotion possibility, $P(\mathbf{e}_{-M})\mathcal{V}$, in the second period. On the other hand, a bad outcome only gives them the wage in the first period, W_M^B. For the sum of expected utilities, the firm has to offer a higher utility for a good outcome than a bad outcome in order to induce a positive effort from managers. That is, $u(W_M^G) + P(\mathbf{e}_{-M})\mathcal{V}$ must be greater than $u(W_M^B)$.

The previous result illustrates that promotion incentive can dominate the incentive associated with pay for performance. The following result shows that the opposite situation can also happen.[34]

[33] Similar result holds under the Promotion Rule 2. The difference is that $W_g^* > W_b^*$ always holds when $N = 1$.

[34] For this result, I drop **Assumption 3**. That is, I allow corner solutions.

Corollary 6 *For a given (N, e_C, e_M), there is $\widetilde{\underline{U}}_M$ such that*

$$\mathcal{V}^* = 0$$

if $\underline{U}_M \leq \widetilde{\underline{U}}_M$.

That is, the pay for performance incentive scheme might dominate the promotion based scheme. In this case, there is no incentive for the firm to introduce a hierarchical structure. One can interpret the previous two results in two ways. First, if a managers' reservation value is really high (low) and the CEO's individual rationality constraint does not bind, a firm should focus on the promotion (pay for performance) incentive scheme. Also, if the CEO's reservation value is really high, the firm would prefer the promotion based incentive scheme to the pay for performance incentive scheme.

2.6.3.1 The effect of internal competition on firm's profit per agent

Up to this point, I focus on how the two incentive schemes interact in a firm. Here, I turn to another question: why does the firm use a hierarchical structure based on the competition among managers by paying higher compensation to the CEO? In order to answer this question, I focus on the firm's profit per agent measured by

$$\Pi(\mathcal{V}^*|N) \equiv \frac{1}{N+1} F(\mathcal{V}^*|N).$$

In order to concentrate on the effect of competition between managers, I assume that $\mathcal{G} \equiv \mathcal{G}_C = \mathcal{G}_M$ and $\mathcal{B} \equiv \mathcal{B}_C = \mathcal{B}_M$. In particular, I try to answer the question: can firms increase the profit per agent by adopting a hierarchical structure? The following result shows what the firm can do by introducing internal competition between managers.

Proposition 8 *If $1 > e_M > e_C > 0$, there is $(\mathcal{G}^*, \mathcal{B}^*)$ such that for $\mathcal{G} - \mathcal{B} \geq \mathcal{G}^* - \mathcal{B}^*$, $\Pi(\mathcal{V}^*|N) > \Pi(\mathcal{V}^*|1)$ for a given N. Moreover, for a given $\mathcal{G} - \mathcal{B} \leq \overline{O}$, there is $\widehat{N}$ such that $\Pi(\mathcal{V}^*|N) < \Pi(\mathcal{V}^*|1)$ if $N > \widehat{N}$, where $\overline{O}$ is derived in the Appendix.*

Moreover, for a given $\mathcal{G} - \mathcal{B} \leq \overline{O}$, there is $\widehat{N}$ such that $\Pi(\mathcal{V}^*|N) < \Pi(\mathcal{V}^*|1)$ if $N > \widehat{N}$ and $e_M \in (e_C, \bar{e}_M]$, where $\overline{O}$ and $\bar{e}_M$ are derived in the Appendix.

Figure 2.3 illustrates this result graphically. Under the given parameters, the profit per agent is maximized when $N = 2$ if $\underline{U}_C = 0$ and when $N = 3$ if $\underline{U}_C = 3$. The latter case means that the firm can increase its profit by hiring three managers and one

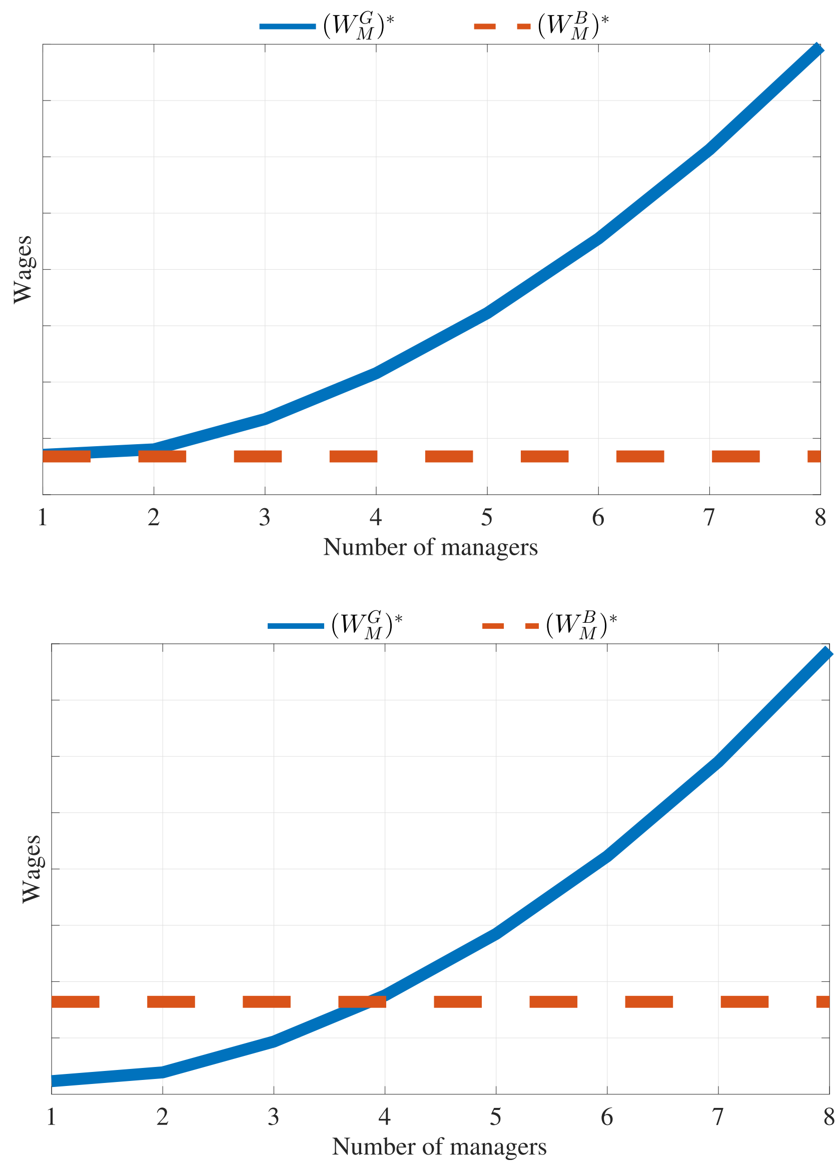

Figure 2.2: Managers' wage levels for $\underline{U}_M = 8$ (up) and $\underline{U}_M = 9$ (down) when $\underline{U}_C = 0$. Other parameters are the same as those in **Figure 2.1**.

CEO with internal competition compared to the situation where there are two lower-level managers and two upper-level managers without any internal competition. In the second scenario, there is no competition for promotion among managers, but the decision is only determined by their own performances. These two scenarios are represented by $N = 3$ case and two $N = 1$ structures.

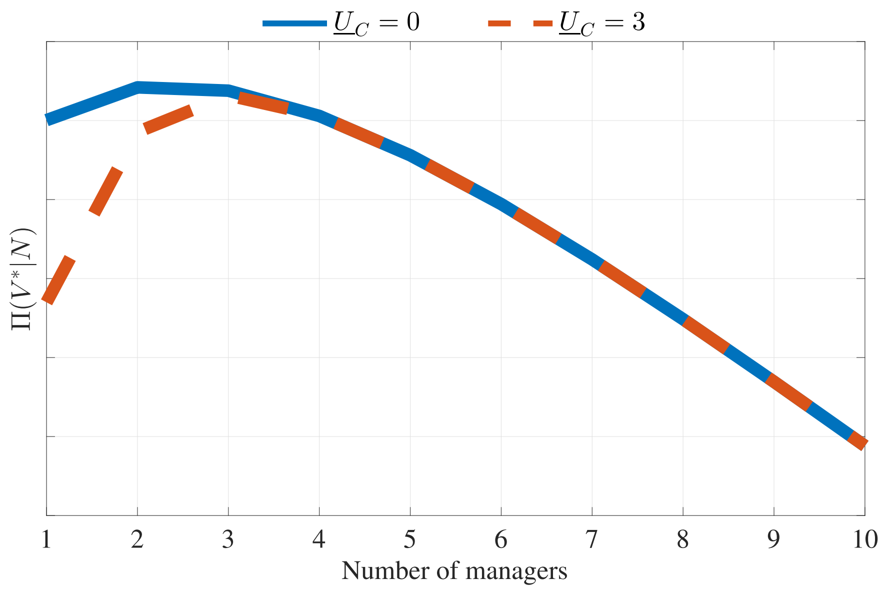

Figure 2.3: Firm's profit per agent for $\underline{U}_C = 0$ (solid line) or $\underline{U}_C = 3$ (dashed line) when $\underline{U}_M = 3$. Other parameters are the same as those in **Figure 2.1**.

Moreover, the following result extends **Proposition 8** with the log utility function and $e_C = e_M = e$. The difference is that the previous result comes from the managers' higher contribution to the firm's output than the CEO's. However, the following proposition is related to the reduction in managers' wages. That is, as the number of managers increases, the firm reduces W_M^G by raising promotion incentives.

Proposition 9 *Suppose that agents have the log utility function,* $u(x) = \log(x)$*. Also, assume that* $\lim_{e\to 1} g'(e) = \infty$*. Then, for a given* $N \geq 2$*, there is* $(\underline{U}_M^*, \underline{U}_M^{**}, e^*)$ *such that* $\Pi(\mathcal{V}^*|N) > \Pi(\mathcal{V}^*|1)$ *if* $e > e^*$ *and* $\underline{U}_M \in [\underline{U}_M^*, \underline{U}_M^{**})$*, where* $\underline{U}_M^* < \underline{U}_M^{**}$*. Moreover, for a fixed* $(\underline{U}_M, e)$*, there is* $N^* \geq 1$ *such that* $\Pi(\mathcal{V}^*|N) < \Pi(\mathcal{V}^*|1)$ *if* $N \geq N^*$*.*

Both **Proposition 8** and **9** illustrate that the profit per agent decreases beyond some point of N. In other words, the relationship between the profit per agent and the number of managers is not monotonic.

2.6.4 The second stage

In this section, I relax the condition that the effort levels are determined exogenously. In order to make the analysis more tractable, I impose the following assumption.

Assumption 4 *The agents' cost function $g(\cdot)$ satisfies $g'''(\cdot) \geq 0$ with $\lim_{e \to 1} g'(\cdot) = \infty$. Also, the two parameters in the probability of good outcome $s(e) = \alpha + \beta e$ satisfies $\alpha + \beta \leq \frac{1}{2}$. $u''(x)/u'(x)^3$ is a decreasing function in x.*

This assumption provides a sufficient condition for concavity of the firm's objective function with respect to agents' effort level for a given $\mathcal{V}$. This allows one to focus on the first order condition with respect to agent's effort when solving the problem. Intuitively, these conditions together make the firm's payout more rapidly increase as the firm requires a higher effort from an agent. It is worth mentioning the condition that $\alpha + \beta \leq \frac{1}{2}$. Different from Lazear and Rosen (1981), the agent's output has only two possible outcomes, good or bad. Under this binary output process, change of its mean by controlling the agent's effort inevitably leads to change of the variance. Moreover this variance is equal to zero if $s(e)$ is zero or one, which eliminates the moral hazard problem since the realized output perfectly reveals the agent's exerted effort. In particular, a probability around $s(e) = 1$ could be problematic since the variance of the output decreases as $s(e)$ increases. This yields a tension between a higher wage and a lower agency problem around $s(e)$ equal to one. By restricting the two parameters in the given way, I can constrain the variance and the mean of the output to have a positive relationship. However, this assumption can be relaxed if the cost function $g(e)$ is sufficiently convex. This condition makes the growth of the wage component dominate the decrease of the agency problem when the firm requires a higher effort even if this reduces the variance of the output.

2.6.4.1 Optimal CEO effort choice

I relax the condition of exogenous effort choice in this section: I allow the firm to optimally choose the CEO's effort level in order to maximize its profit, still fixing the managers' effort level. I assume that the function $F(\mathcal{V})$ satisfies the strict quasi-concavity condition in order to solve the firm's original problem using the modified one.[35]

Assumption 5 *If agents are risk-averse, $F(\mathcal{V})$ is strictly quasi-concave.*

The following result confirms that the CEO's expected utility is still an increasing function in the size of the firm. Moreover, the CEO's optimal effort level decreases as the firm size increases.

Proposition 10 *When agents are risk-averse, $\mathcal{V}^*$ is a strictly increasing function in N while e_C^* is a strictly decreasing function in N.*

Intuitively, under the operational structure characterized by agents' independent outputs, the relative importance of the CEO's output, compared to the managers' total output, decreases as the number of managers increases. Therefore, the firm uses the CEO's position as a bonus rather than an output source. For the purpose of reducing the cost of increasing promotion incentives, the firm decreases the CEO's effort level.

It is important to note that a higher promotion incentive does not directly imply a higher expected compensation in this case. The movements of the optimal effort level and the promotion incentive predict the change of expected compensation in the opposite directions when the firm size grows. **Figure 2.4** and **Figure 2.5** illustrate an example for given parameters to see how CEO compensation changes as the firm size grows. **Figure 2.4** shows that the relationship between CEO effort level and the ratio of CEO's fixed pay to incentive pay remains the same as the prediction without promotion possibility: the ratio increases as the firm requires a lower effort level from the CEO. Also, a moral hazard model without promotion possibility predicts lower compensation and incentive pay when a firm requires a lesser effort level. However, the two variables in **Figure 2.5** present the opposite results. This result demonstrates that the trade-off between the required effort level and compensation could be reversed if people take into account the presence of internal labor markets. As a special case, if agents have the log utility function, the following corollary shows that CEO's incentive pay rises as the firm size increases although it requires lower effort from the CEO.

[35]Numerically, this condition is investigated, and I confirm that the condition holds generally.

Corollary 7 *Suppose that agents have the log utility function,* $u(x) = log(x)$*. Then, CEO's incentive pay measured by* $(W_C^G)^* - (W_C^B)^*$ *rises as N increases.*

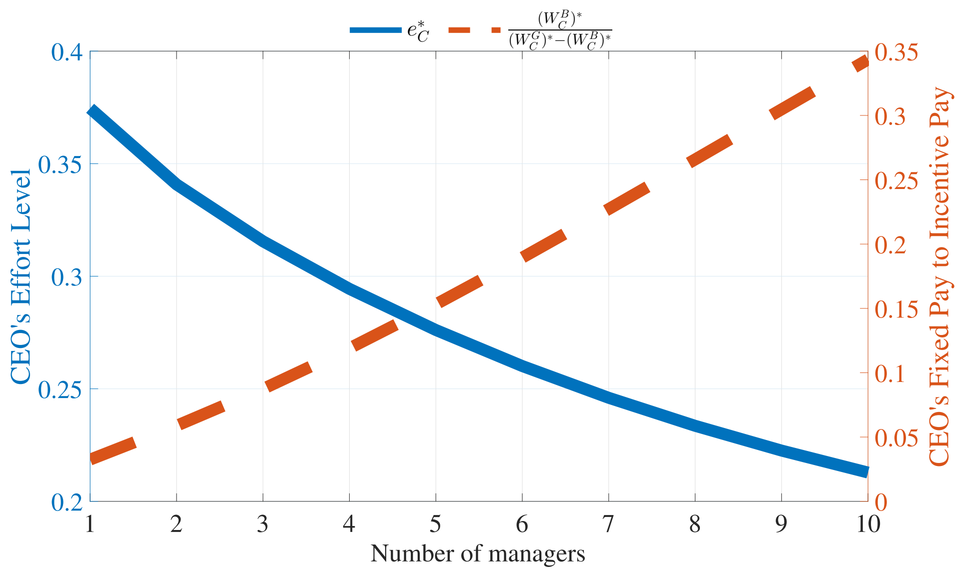

Figure 2.4: CEO effort level (solid line) and the ratio of CEO's fixed pay relative to incentive pay (dashed line) for $\underline{U}_C = 0$ and $\underline{U}_M = 3$. Other parameters are the same as those in **Figure 2.1** except e_C.

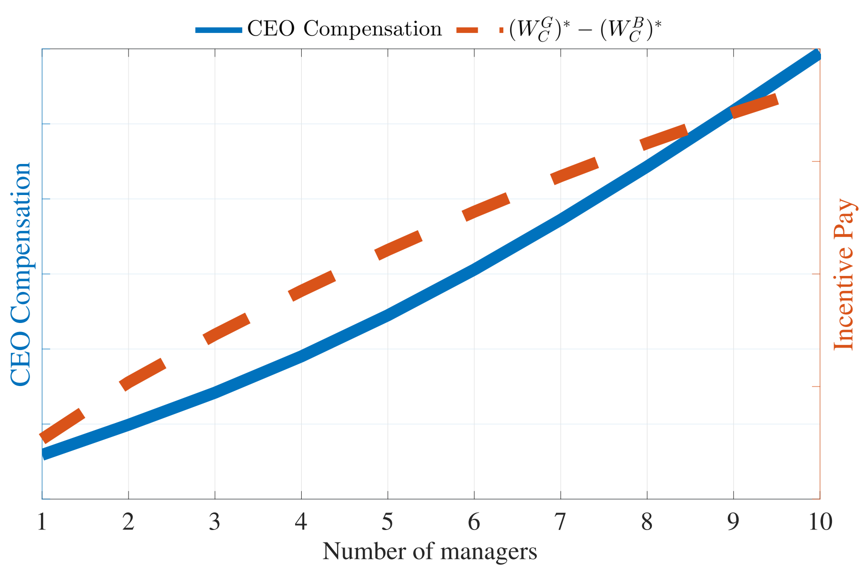

Figure 2.5: CEO compensation (solid line) and the level of incentive pay (dashed line) for $\underline{U}_C = 0$ and $\underline{U}_M = 3$. Other parameters are the same as those in **Figure 2.4**.

2.6.4.2 General case

In this section, I mainly deal with the numerical examples illustrating that the previous results still hold when the firm chooses both effort levels e_C and e_M optimally. In addition, I analytically show that $e_C^* < e_M^*$ could be a consequence of the firm's optimal decision.[36] Again, I assume that the function $F(\mathcal{V})$ satisfies the desired condition.[37]

Assumption 6 *If agents are risk-averse, $F(\mathcal{V})$ is strictly quasi-concave.*

First, **Proposition 11** shows that the firm can optimally require a higher effort from managers than that of the CEO. **Figure 2.6** illustrates this numerically.

Proposition 11 *Suppose that $\underline{U}_M = 0$, then there exists $\widehat{N}$ such that $e_M^* > e_C^*$ if $N > \widehat{N}$.*

This implies that it can be optimal for a firm to provide an expected utility to the CEO beyond her and managers' reservation values since **Proposition 5** only requires $e_C \leq e_M$. Also, the firm can raise its profit per agent by adopting the internal labor market based on competition between managers as the same reason. **Figure 2.7** and **2.8** confirm these observations. In particular, in this example, CEO's expected utility is higher than managers' reservation value if N is greater than or equal to three. Also, **Figure 2.8** shows that the firm's profit per agent is maximized when the number of managers is equal to three.

Furthermore, **Figure 2.9** demonstrates that $(W_M^B)^*$ could be greater than $(W_M^G)^*$ as $\underline{U}_M$ increases from 6 to 7. In particular, $(W_M^B)^*$ is greater than $(W_M^G)^*$ if $N \in [5, 54]$. That is, the firm can optimally reduce the dependency of managers' compensation on their performance.

2.7 The Effect of Job Security on Promotion Incentives

As the first extension of the benchmark model, I analyze the effect of employment structure on promotion incentives in this section. In particular, I focus on how the level of promotion incentives changes according to the firm's policy on the executives' job security. This question is important since issues related to executives' job

[36] Recall that **Proposition 5** and **8** depend on this feature.
[37] The condition is checked numerically.

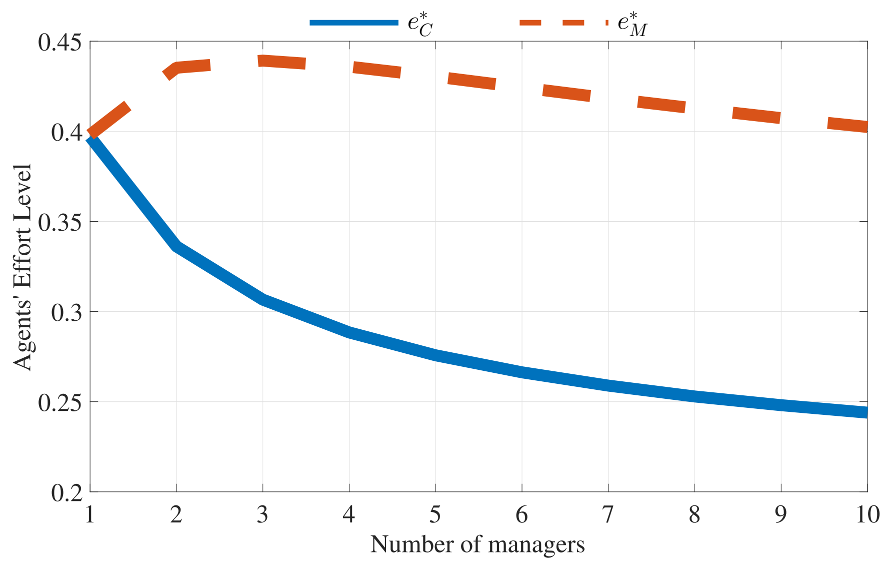

Figure 2.6: CEO effort level (solid line) and Manager's effort level (dashed line) for $\underline{U}_C = 0$ and $\underline{U}_M = 3$. Other parameters are the same as those in **Figure 2.1** except e_C and e_M.

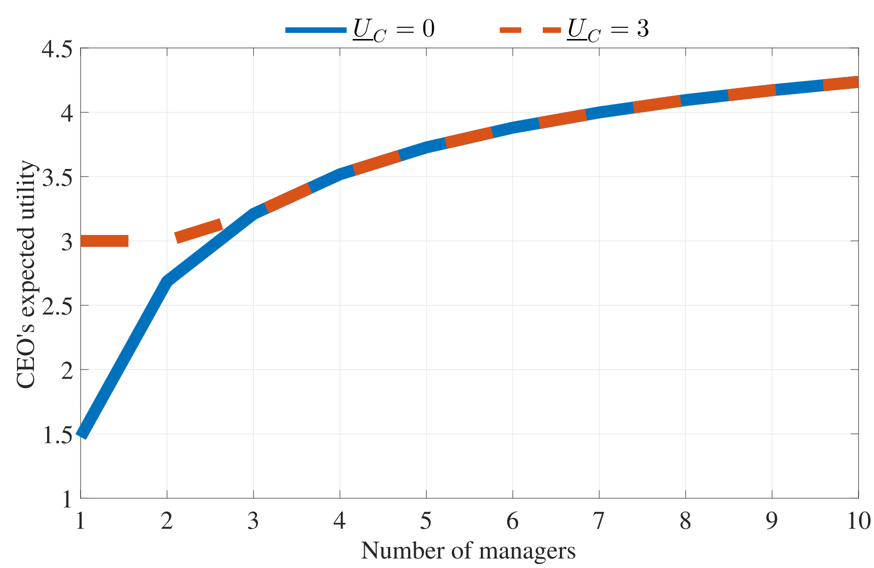

Figure 2.7: CEO's expected utility for $\underline{U}_C = 0$ (solid line) or $\underline{U}_C = 3$ (dashed line) when $\underline{U}_M = 3$. Other parameters are the same as those in **Figure 2.7**.

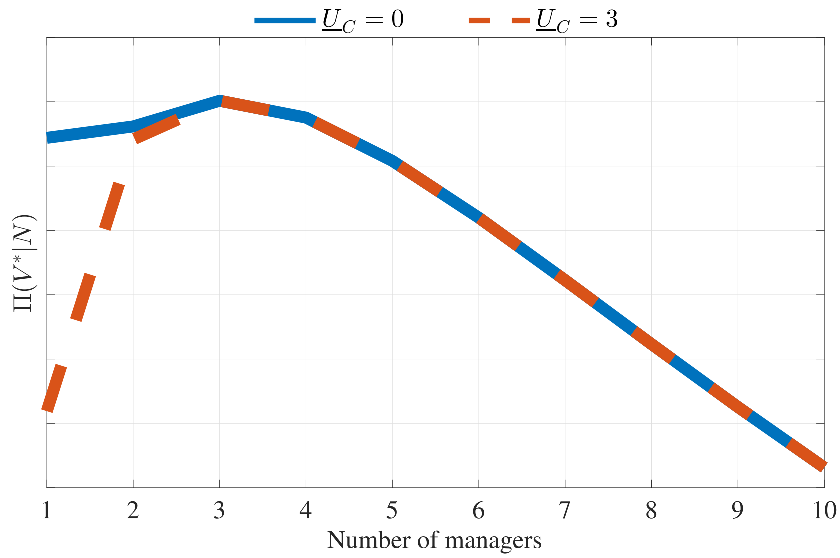

Figure 2.8: Firm's profit per agent when $\underline{U}_C = 0$ (solid line) or 3 (dashed line) when $\underline{U}_M = 3$. Other parameters are the same as those in **Figure 2.7**.

security have received significant attention from both researchers and practitioners. For example, Jenter and Kanaan (2015) illustrate that bad performance related to negative shocks causes a forced CEO turnover. Hence, in order to understand executives' compensation more clearly, it is valuable to examine the effect of executives' job security on their compensation. Throughout the section, I treat agents' effort levels as exogenously given.

2.7.1 The effect of managers job security on promotion incentives

So far I assume that managers leave the firm if they are not promoted to CEO although they perform well in their position. In this section, I consider a slightly different employment structure. Namely, even if managers fail to be promoted to CEO, they can stay in the firm one more period as senior managers if their performance is good.[38] However, the senior managers do not have any chance for promotion. That is, senior managers retire regardless of their outcomes. Except this strengthened job security for managers, every employment structure is the same as the benchmark model. That is, the firm still hires N managers from an external labor market every period, and uses the same promotion rule. Then, does the firm have an incentive to provide more utility to the CEO than other successful managers? That is,

[38] Since managers do not have any chance of promotion if they face a bad outcome in the first period under promotion rule 1, it does not affect the result in terms of promotion incentives whether managers can stay in the firm after a bad outcome.

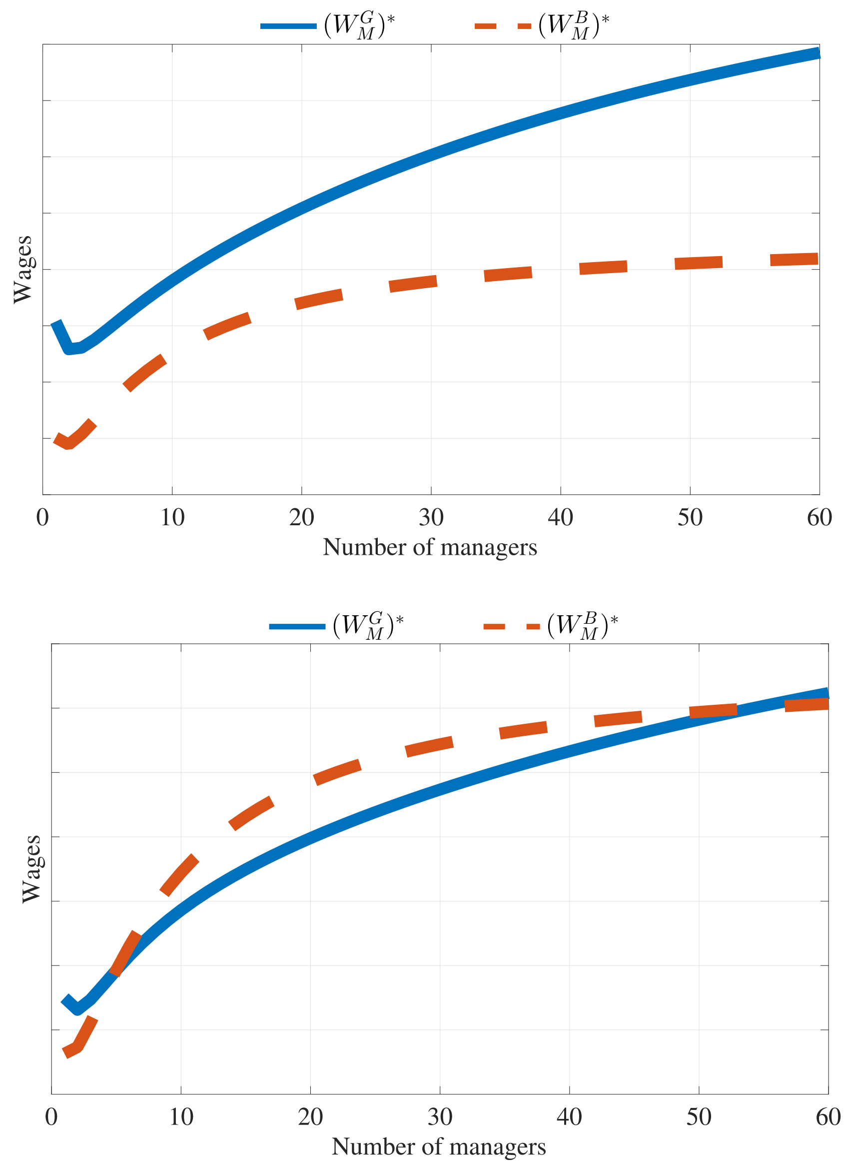

Figure 2.9: Managers' wage levels for $\underline{U}_M = 6$ (up) and $\underline{U}_M = 7$ (down) when $\underline{U}_C = 0$. Other parameters are the same as those in **Figure 2.7**.

can promotion incentives still contribute to the increase in CEO compensation?

Denote the senior manager's compensation by (W_M^{GG}, W_M^{GB}), which are tied to a good and bad outcome, respectively. Note that only managers with a good outcome remain in the firm as senior managers. Also, I use the term e_{M2} to distinguish the senior manager's effort from the (young) manager's effort, e_{M1}. The difference from the benchmark model is that successful managers obtain the expected utility,

$$s(e_{M2})u(W_M^{GG}) + (1 - s(e_{M2}))u(W_M^{GB}) - g(e_{M2}),$$

though they are not picked as the CEO.[39] I denote this utility by U_{M2} for brevity. Then, I show that managers' compensation scheme satisfies

$$\begin{aligned}
u(W_M^G) &= \underline{U}_M + g(e_{M1}) + (1 - s(e_{M1}))\frac{g'(e_{M1})}{\beta} - P(\mathbf{e}_{-M})\mathcal{V} - (1 - P(\mathbf{e}_{-M}))U_{M2} \\
u(W_M^B) &= \underline{U}_M + g(e_{M1}) - s(e_{M1})\frac{g'(e_{M1})}{\beta} \\
u(W_M^{GG}) &= U_{M2} + g(e_{M2}) + (1 - s(e_{M2}))\frac{g'(e_{M2})}{\beta} \\
u(W_M^{GB}) &= U_{M2} + g(e_{M2}) - s(e_{M2})\frac{g'(e_{M2})}{\beta}.
\end{aligned}$$

The first equation illustrates that the manager's wage tied to good outcome depends on the senior manager's expected utility, U_{M2} as well as the promotion incentive, $\mathcal{V}$. Moreover, this wage and the senior manager's compensation satisfy the following relationship according to Rogerson (1985a)

$$\frac{1}{u'(W_M^G)} = \delta \left[\frac{s(e_{M2})}{u'(W_M^{GG})} + \frac{1 - s(e_{M2})}{u'(W_M^{GB})} \right]. \tag{2.9}$$

For this problem, the main concern is the difference between the CEO's expected utility ($\mathcal{V}^*$) and senior managers' expected utility (U_{M2}^*). If senior managers have higher expected utility than the CEO, promotion incentives do not play any effective role as an incentive tool. In this case, managers do not want to be promoted to the CEO. Hence, from now on, I add one more constraint in the firm's problem, call this promotion constraint.

$$\mathcal{V} \geq U_{M2} \qquad \text{(Promotion Constraint)}$$

The following result indicates when the promotion constraint does not bind.

[39] The detail of the firm's problem can be found in Appendix A.2.1.

Proposition 12 *Assume that* $\frac{u''(x)}{u'(x)^3}$ *is a decreasing function in* x*. Then, if* δ *is sufficiently large*[40] *, there is* $(\hat{e}_C, \widehat{N})$ *such that* $\mathcal{V}^* > U^*_{M2}$ *when* $e_C \leq \hat{e}_C$ *and* $N \geq \widehat{N}$*, where* $\hat{e}_C < e_{M2}$*. Moreover,* $\mathcal{V}^*$ *is an increasing function in* N *when* $\mathcal{V}^* \geq U^*_{M2}$*.*

Therefore, the firm can still optimally provide its CEO a compensation beyond her reservation value even if managers have strong job security. That is, **Proposition 4** still holds under this extension. This means that promotion incentives contribute the rise in CEO compensation when the size of the firm is large and the firm requires less effort from CEO than managers, even if managers have job security.

2.7.2 The effect of CEO job security on promotion incentives

In the benchmark model, a CEO leaves the firm after one period regardless of her performance. In this section, I assume differently that a CEO will stay and work one more period if she performs well in the first period. In other words, the CEO will be sacked as punishment for bad performance. Regarding managers, I maintain the same structure as the benchmark model. That is, only the manager promoted to the CEO remains in the firm and works one more period. In this section, I slightly modify **Assumption 2**.

Assumption 7 *The firm offers contracts based on agents' positions and the seniority of its CEO. That is, it offers three (possibly) different types of contracts: 1) CEO, 2) managers with new CEO, and 3) managers with CEO close to retirement.*

Under this assumption, the firm can offer different contracts to managers according to the length of the CEO's remaining term. Therefore, the extension allows one to examine managers' cohort effects as well as the effect of the CEO's job security. [41] Similar with section 2.4, I reduce the firm's problem to a more tractable form under **Assumption 7**[42] .

The following proposition tells how the level of promotion incentives changes according to the change of CEO job security. More formally, call the case where a CEO is sacked if her performance is bad in the first period “unguaranteed job security". On the contrary, name it “guaranteed job security" if a CEO is not fired regardless of the first period's performance. By comparing these situations, one can

[40] The condition for δ can be found in the proof.

[41] By the cohort effects, I mean that a cohort who earn more on entry maintains its advantage through time according to Baker, Gibbs, and Holmstrom (1994).

[42] The detail of the firm's problem can be found in Appendix A.2.2.

see when the firm puts more weight on promotion incentives. In particular, I focus on the level of promotion incentives by fixing the required effort at the same level for both cases.

Proposition 13 *Assume that* $\lim_{e\to 1} g'(e) = \infty$ *and* $\lim_{x\to\infty} u'(x) = 0$*. Also, suppose that the firm requires an effort level* e_C *anf* e_M *from the CEO and managers, respectively, in every period. Then, there is* $e_C^* \in (0,1)$ *such that* $\mathcal{V}^*(e_C) \geq \widehat{\mathcal{V}}^*(e_C)$ *if* $e_C \in [e_C^*, 1)$ *and* δ *is sufficiently high*[43] *, where* $\mathcal{V}^*(e_C)$ *and* $\widehat{\mathcal{V}}^*(e_C)$ *are the optimal level of promotion incentives for unguaranteed and guaranteed job security cases, respectively, for a given* e_C*. Moreover, for the sufficiently high* δ*, if* $\widehat{\mathcal{V}}^*(0) > 0$*, there is an interval* $[e_C^*, \bar{e}_C]$*, where* $\bar{e}_C \in (e_C^*, 1)$*, such that* $\mathcal{V}^*(e_C) > \widehat{\mathcal{V}}^*(e_C)$ *for* e_C *in the interval.*

It is important to know that the CEO's individual rationality constraint binds in both cases if one does not consider promotion possibilities. However, taking into account promotion possibilities gives a different answer. That is, it shows that a CEO can receive a favorable contract when her job security is less guaranteed. This result is also related to the tension between CEO incentives and managers' incentives. If the firm does not guarantee CEO job security, it gives more promotion chances to its managers. On the other hand, raising promotion incentive is more costly to the firm under the guaranteed situation than the other case since this make the firm need to pay more in the second period. Hence, when the firm requires a higher effort from the CEO, it wants to emphasize more the incentive for the CEO under the guaranteed situation. This emphasis yields fewer promotion incentives than the unguaranteed case.

Moreover, it can be shown that $W_{M1}^G > W_{M2}^G$ if $e_{M1} = e_{M2}$. Hence, managers' wage can exhibit a gap between cohorts although their abilities and required effort are the same. The reason is straightforward. The promotion possibility of managers relies on the timing they enter the firm: managers working with a new CEO expect less to be promoted than managers below a CEO who is about to retire, other things being equal. This affects managers' wages in an unambiguous way.

Now, I extend the above model allowing managers stay in the firm one more period if their performance is good although they are not promoted to the CEO. Then, do the managers who start a career with a new CEO still earn more money than

[43]The condition for δ can be found in the proof.

managers starting with CEO close to retirement the second period? The following result shows that this is the case if the promotion is desirable for managers.

Proposition 14 *Suppose that the firm requires the same effort* $(e_{M11}, e_{M12}) = (e_{M21}, e_{M22})$*, where* e_{Mij} *is the required effort level in the period* j *from managers with a new CEO (*$i = 1$*), or managers with CEO close to retirement (*$i = 2$*). Then,* $(U^2_{M1})^* > (U^2_{M2})^*$ *if* $\mathcal{V}^* > (U^2_{M1})^*$*. This also implies that* $(W^G_{M1})^* > (W^G_{M2})^*$*.*

Hence, this model can also predict a cohort effect. That is, the expected wage level of managers who have a lesser chance of promotion can be larger than that of managers having better chance in their whole careers. The condition that being the CEO is strictly beneficial than staying in the manager position is closely related to **Proposition 12**. Although I do not formally analyze the condition, if the firm does not require a high effort from the CEO in her second period, **Proposition 12** might hold since the firm optimally distributes the cost of raising promotion incentives.

2.8 Complementary Tasks: Multiplication Specification

I study how the firm's operational structure affects the firm's optimal contract in this section. Until now, I assume that a CEO's contribution to the firm's profit is independent of managers' outputs. As the second extension of the benchmark model, I investigate how the previous results change if there is a dependency between them. In particular, I consider the following firm's output structure:

$$[s(e_C)\mathcal{G}_C + (1 - s(e_C))\mathcal{B}_C]\, E\left[f\left(\sum_{i=1}^{N} X_i\right)\right],$$

where the CEO's task has a multiplication effect on the managers' aggregated output through the function $f(\cdot)$. For the function $f(\cdot)$, I impose some assumptions.

Assumption 8 *The function* $f(\cdot)$ *satisfies* $f(\cdot) > 0$*,* $f'(\cdot) > 0$*, and* $\lim_{x\to\infty} f'(x) > 0$*.*

The first two conditions are quite general. The last condition rules out the case when the size effect disappears since the firm size N only affects the firm's output through the function $f(\cdot)$. For example, a linear function taking positive numbers on its domain satisfies all requirements. Also, I adopt **Assumption 4** for the same reason as mentioned above. Note that the argument of the function $f(\cdot)$ is $[i\mathcal{G}_M + (N - i)\mathcal{B}_M]$,

$i = 0, \ldots, N$. In order to assure that the argument is a strictly increasing function in the number of managers, I assume that $\mathcal{B}_M > 0$. Under these assumptions, I analyze the following problem with an exogenously given managers' effort level e_M like in Section 2.6.4.1.

$$\max_{\mathcal{V} \in [0,\infty)} F(\mathcal{V}), \text{ where}$$

$$\begin{aligned} F(\mathcal{V}) = \max_{\{(e_C, W_C^G, W_C^B), (W_M^G, W_M^B)\}} & [s(e_C)\mathcal{G}_C + (1 - s(e_C))\mathcal{B}_C]\, E\left[f\left(\sum_{i=1}^{N} X_i\right)\right] \\ & - s(e_C)W_C^G - (1 - s(e_C))W_C^B \\ & - N[s(e_M)W_M^G + (1 - s(e_M))W_M^B] \\ & \text{subject to four constraints, (2.5), (2.6), (2.7), (2.8).} \end{aligned}$$

The distinction between the independent and the multiplication specification is whether the marginal productivity of the CEO changes according to the size of the firm. Since the multiplication specification assumes an increasing productivity, the CEO's task is more crucial to the output when the firm is larger. This change of the CEO's role yields a different result from **Proposition 10**.

Proposition 15 *Assume that* $\lim_{x\to\infty} u'(x) = 0$. *Then, there is* $\widehat{N}$ *such that* $\mathcal{V}^*$ *is a decreasing function in* N *while* e_C^* *is a strictly increasing function in* N. *Also, if* $\mathcal{V}^* > 0$, $\mathcal{V}^*$ *strictly decreases as* N *grows.*

Recall that **Proposition 10** says that the promotion incentive increases as the firm size grows, and the firm requires less effort from the CEO. However, **Proposition 15** reveals that the exactly opposite situation happens if the size of the firm is sufficiently large, and the marginal contribution of the CEO's effort increases with the firm size. **Figure 2.10** and **2.11** show numerical examples under the multiplication specification. **Figure 2.10** illustrates a non-monotonic relationship between the promotion incentive and the number of managers. Especially, if CEO's reservation value is equal to one, which is exactly the same as managers' reservation value in the example, the CEO's individual rationality constraint binds when N is less than 7 or greater than 22. Also, **Figure 2.11** shows that the CEO's effort level increases as N grows. Intuitively, the firm wants the CEO to exert more effort since her marginal productivity grows as the number of managers increases. However, the concavity of the CEO's utility function makes it very costly for the firm to induce her to exert more effort especially when e_C is high. Hence, the firm puts more emphasis on the CEO's absolute performance-based compensation not the promotion incentive for managers. Namely, by reducing the promotion incentive, the firm can increase the CEO's effort level at a lower cost.

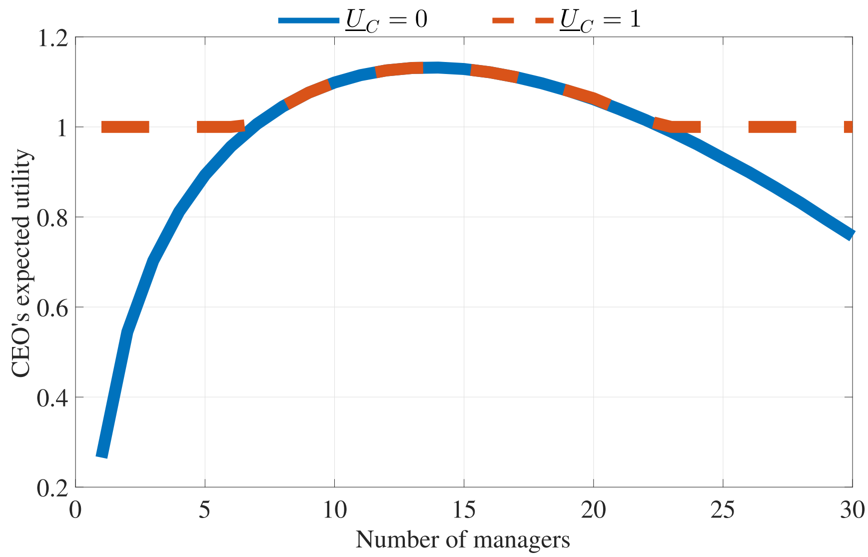

Figure 2.10: CEO's expected utility under a multiplication specfication for $\underline{U}_C = 0$ (solid line) or $\underline{U}_C = 1$ (dashed line) when $\underline{U}_M = 1$. Other parameters are $\mathcal{G}_C = 1.7$, $\mathcal{B}_C = 0.7$, $\mathcal{G}_M = 35$, $\mathcal{B}_M = 5$, $f(x) = x$, $\alpha = 0.1$, $\beta = 0.4$, $e_M = 0.22$, $u(x) = \log(x)$ and $g(a) = -\frac{1}{a-1} - a - 1$.

2.9 Comparing Two Promotion Rules

In this section, I explain when promotion rule 1 can be preferred to promotion rule 2 by the firm. Consider the situation where managers' abilities are private information to them. Recall that a manager's effort affects the probability of good outcome by the function $s(e) = \alpha + \beta e$. In the following model, managers' ability is determined by β. Specifically, assume that β can be one of two values, $\bar{\beta}$ and $\underline{\beta}$, with the probability $q \in (0, 1)$ and $1 - q$, respectively. Also, assume that the two values satisfy $\bar{\beta} > \underline{\beta} > 0$ and $\bar{\beta} \leq \frac{1}{2} - \alpha$. Call the agent with $\beta = \bar{\beta}$ by high type (H) and the agent with $\beta = \underline{\beta}$ by low type (L). For brevity, I denote the probability of good performance for a given e by $s_H(e)$ and $s_L(e)$ according to manager's type. If the firm hires the CEO from the external labor market, they also have the same prior probability about types of agents. That is, the firm's prior probability that a CEO from the external labor market is high type is q. In order to focus on the effect

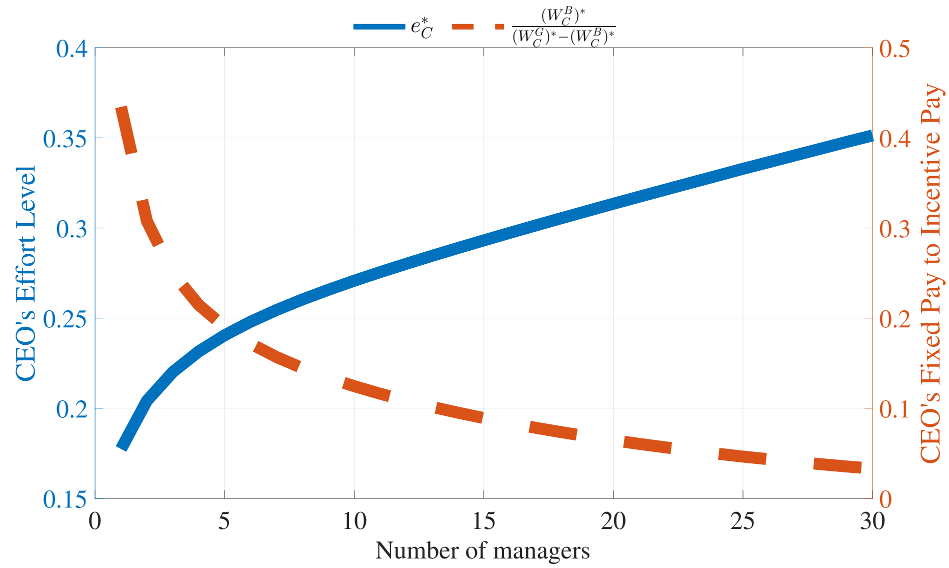

Figure 2.11: Under a multiplication specification, CEO effort level (solid line) and the ratio of CEO's fixed pay relative to incentive pay (dashed line) for $\underline{U}_C = 0$ and $\underline{U}_M = 1$. Other parameters are the same as those in **Figure 2.10**.

of promotion rules, I additionally assume that N is equal to 2 in this section. Then, the firm's problem is the following:

$$\max_{(W_C),(e_L,e_H,W_M^G,W_M^B)} \gamma E[\beta|e_H,e_L] - W_C + E[\Pi_M|(e_H,e_L,W_M^G,W_M^B)]$$

subject to

$$u(W_C) \geq \underline{U}_C \ (IR_C)$$

$$s_H(e_H)[u(W_M^G) + P(\mathbf{e}_{-M})V_C] + (1 - s_H(e_H))[u(W_M^B) + R(\mathbf{e}_{-M})V_C] - g(e_H) \geq \underline{U}_M \ (IR_H)$$

$$s_L(e_L)[u(W_M^G) + P(\mathbf{e}_{-M})V_C] + (1 - s_L(e_L))[u(W_M^B) + R(\mathbf{e}_{-M})V_C] - g(e_L) \geq \underline{U}_M \ (IR_L)$$

$$e_H \in \arg\max_{\hat{e}} s_H(\hat{e})[u(W_M^G) + P(\mathbf{e}_{-M})V_C] + (1 - s_H(\hat{e}))[u(W_M^B) + R(\mathbf{e}_{-M})V_C] - g(\hat{e}) \ (IC_H)$$

$$e_L \in \arg\max_{\hat{e}} s_L(\hat{e})[u(W_M^G) + P(\mathbf{e}_{-M})V_C] + (1 - s_L(\hat{e}))[u(W_M^B) + R(\mathbf{e}_{-M})V_C] - g(e_L) \ (IC_L),$$

where $E[\Pi_M|(e_H,e_L,W_M^G,W_M^B)]$ represents the expected profit from two managers[44], and $E[\beta|e_H,e_L]$ is the expected β when the firm requires e_H and e_L from high-type and low-type managers, respectively. γ determines the firm's benefit of hiring high-type CEO. More formally, γ is the firm's marginal benefit from higher expected CEO's β. Also, the managers' promotion probability when their outcome is

[44]The exact expression can be found in Appendix A.2.3.

good ($P(\mathbf{e}_{-M})$) and bad ($R(\mathbf{e}_{-M})$) are

$$P(\mathbf{e}_{-M}) = \frac{2 - s_H(e_H)q - s_L(e_L)(1-q)}{2}, \text{ and}$$

$$R(\mathbf{e}_{-M}) = \begin{cases} 0 & \text{under promotion rule 1} \\ \frac{1-s_H(e_H)q-s_L(e_L)(1-q)}{2} & \text{under promotion rule 2} \end{cases}.$$

Here, I only consider a pooling equilibrium so that the firm provides the same contract regardless of managers' types.[45] In this problem, moral hazard aspects related to the CEO are ignored. However, a hired CEO's ability, which determines her productivity, depends on the promotion rule and managers' compensation. Specifically, the choice of promotion rule determines the functional form of $E[\beta|e_H, e_L]$ while the compensation affects managers' effort choice.[46] Moreover, I simply assume that agents' cost function is a quadratic function, that is, $g(e) = \frac{\kappa e^2}{2}$ with $\kappa > 0$.

The following result illustrates that the firm will prefer the promotion rule 1 to the promotion rule 2 if the benefit from hiring a high type CEO is big enough.

Proposition 16 *There is $\hat{\gamma} > 0$ such that the profit under promotion rule 1 is strictly greater than that under promotion rule 2 if $\gamma \geq \hat{\gamma}$.*

That is, if the firm emphasize the role of internal promotion as a screening device, promotion rule 1 will be optimally chosen rather than promotion 2.

2.10 Discussion

In this section, I connect the previous results with two lines of literature: 1) CEO compensation, and 2) Empirical studies in tournament literature.

2.10.1 Implications for the trend in executive compensation

CEO compensation and the wage gap between CEO and other managers have received extensive attention from both researchers and media. Since the influential work by Gabaix and Landier (2008),[47] researchers have focused on the size of the firm in order to explain the rise of CEO compensation. However, the relationship is not monotonic from a long-term perspective as Frydman and Saks (2010) and Frydman and Jenter (2010) illustrate. Here, I connect my model with the non-monotonic relationship between the two factors.

[45] The argument in this section still holds even when one considers a separating equilibrium if the firm determines its promotion decision based on managers' outputs not revealed types.

[46] The functional form can be found in Appendix A.2.3.

[47] See also Tervio (2008).

First, I illustrate that the promotion incentive channel predicts that CEO compensation increases non-monotonically as the size of the firm, measured by the number of internal candidates for the CEO position, grows. **Corollary 4** and **Figure 2.1** demonstrate this relationship. In **Figure 2.1**, when the CEO' reservation value is equal to the managers', the CEO's expected utility and compensation remains a constant up to the point that N is equal to 3 although the firm size increases since the CEO's individual rationality constraint binds. However, beyond this point, the CEO's expected utility and the size of the firm show a monotonic relationship. This rise of CEO's expected utility leads to the increase of CEO's compensation according to **Corollary 4** if her effort level remains a constant. Moreover, **Figure 2.5** shows that this can also be the case when the firm optimally choose the CEO's effort level.

Also, my model suggests that the rise of promotion incentives might play a key role in the growth of the wage gap between CEO and managers during the past 30 years but not before 1980. Therefore, similar with the previous argument, the relationship between the wage gap and the firm size is not monotonic since the CEO's compensation does not change if her individual rationality constraint binds. However, if the number of managers is high enough such that the CEO's individual rationality constraint does not bind, the wage gap might start to widen as **Figure 2.12** illustrates.[48]

It is worth mentioning that the rise of promotion incentives can also affect the CEO's pay for performance incentives. Basically, the CEO's incentive payment, such as stock options and restricted stock, has a positive correlation with the required effort level. However, if researchers take into account promotion incentives, this relationship can be reversed. **Figures 2.4** and **2.5** show this possibility. In this example, the CEO's incentive payment expressed by $(W_C^G)^* - (W_C^B)^*$ increases although the firm requires less effort from her as the size of the firm grows.

2.10.2 Implications for empirical research on tournament theory

Since the seminal work of Lazear and Rosen (1981), an extensive empirical literature studies the implications of tournament theory. Executive compensation is one of the most studied applications in order to test the implications. In these applications, researchers have used the wage gap between CEO and the next level executives as the measure of promotion incentives based on Lazear and Rosen (1981).[49]

[48]When $\underline{U}_C = \underline{U}_M = 3$, the wage gap reduces up to N is equal to 3, which is the smallest N that the CEO's individual rationality constraint does not bind, in this example.

However, to the best of my knowledge, the interaction between CEO incentives, which depend on the role of CEO in a firm, and promotion incentives has not been considered. The interaction, however, should be useful in understanding the cause of the wage gap in a firm as well as CEO compensation. For example, the Securities and Exchange Commission announced that it adopted a rule that required a public company to disclose the ratio of its CEO to the median compensation of its employees. Hence, if researchers misinterpret the meaning of the ratio, it could lead to a distortion in the wage structure or incentives.

In this section, I compare **Proposition 10** and **15**, and derive an important implication regarding a measure of promotion incentives. The important message is that it can be misleading to use the wage gap between CEO and managers as the measure of promotion incentives. The reason for this is that two factors, high-powered incentives for CEO and strong promotion incentives, both should result in a high CEO compensation and a large wage gap when other conditions remain the same. Moreover, when the CEO is risk-averse, the two factors are negatively correlated according to **Proposition 6**. Therefore, researchers need to disentangle these effects in order to measure promotion incentives properly.

One possible way of identifying promotion incentives is to see CEO's base salary (W_C^B in my model). To increase CEO's effort level, the base salary should decrease because of moral hazard problem. Thus, the fall of base salary is related to the demand for higher effort from CEO. On the other hand, when a firm wants to increase promotion incentives for the lower-rung executives or employees, the CEO's base salary should increase. In particular, **Proposition 10** predicts that enhancing promotion incentives leads to the increase in CEO's base salary.

Unfortunately, this identification strategy based on base salary may not work when a firm is small. Under the multiplication specification, for instance, a CEO's effort level and promotional incentives can move in such a way that they affect the base salary in the opposite directions. Despite this limited applicability, I expect that the CEO's base salary can still be useful for identifying promotion incentives for two reasons. First, this undesired situation happens when the size of the firm is small, which makes the problem less of a concern. For small firms, the CEO's individual rationality constraint binds or promotion incentives may dominate the effect of high-powered incentives for CEO. In the former case, the growth in the firm size does not yield the increase in promotion incentives and base salary. Therefore,

---

[49]Lazear and Oyer (2012) and Waldman (2012) provide reviews on the literature.

the base salary can weakly identify promotion incentives. In the latter case, as the firm size grows, I expect that the promotion incentive will be the major driving force of the base salary movement. Second and more importantly, the dataset that economists mostly use consists of large firms. For example, *ExecuComp* dataset only contains public firms. For firms large enough, I expect that promotion incentives move in the opposite direction of CEO effort level based on **Proposition 15**.

As an example of this prediction, **Figure 2.12** and **2.13** illustrate two numerical examples based on given parameters. As these two figures show, the expected wage gap measured by:

$$E[W_C] - E[W_M] = [s(e_C)W_C^G + (1 - s(e_C))W_C^B] - [s(e_M)W_M^G + (1 - s(e_M))W_M^B]$$

grows as the number of managers increases regardless of the direction of promotion incentives. On the other hand, CEO's base salary, overall, captures the change of promotion incentives. This is especially true when the number of managers is large enough as the theory predicts. The future research is to study this relationship under a more complex operational structure. However, unless promotion incentives and the CEO's required effort level move in opposite directions, researchers can exploit the CEO's base salary in order to identify promotion incentives. Moreover, this property is generally applicable to other problems if the organizer of a contest needs to provide incentives to the winner of the contest.

2.11 Conclusion

I examine the effect of a managers' possibility of promotion on executive compensation. Using a theoretical model which incorporates both relative and absolute performance based compensation, I find that it is optimal for a firm to provide a compensation to its CEO higher than her reservation value in order to incentivize lower-rung executives. In particular, this is true if agents are risk-averse, the CEO's marginal productivity is independent of firm size, and there are enough number of competitors for promotion. Therefore, the promotion possibilities yield the growth of CEO compensation as well as create a wage gap between the CEO and other executives. Moreover, promotion incentives reduce the dependence of managers' compensation on their performance. As a justification for the hybrid incentive scheme, I examine the effect of it on a firm's profit per agent. The result illustrates that firms can increase their profit per agent by introducing a promotion structure based on competition among managers. In addition, the relation between the CEO's task and the managers' outputs significantly affects the CEO's compen-

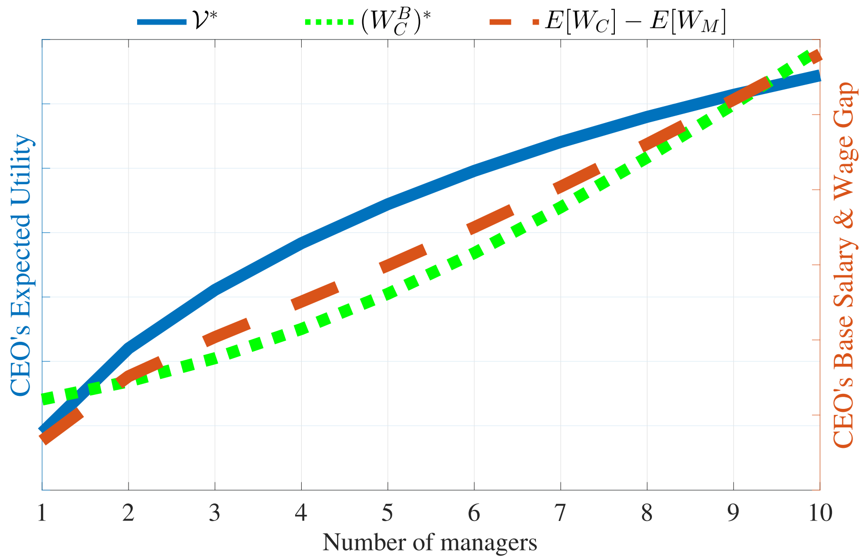

Figure 2.12: CEO's expected utility (solid line), CEO's base salary (dotted line), and Wage gap between CEO and manager (dashed line) under the independent specification for $\underline{U}_C = 0$. Other parameters are the same as those in **Figure 2.4**.

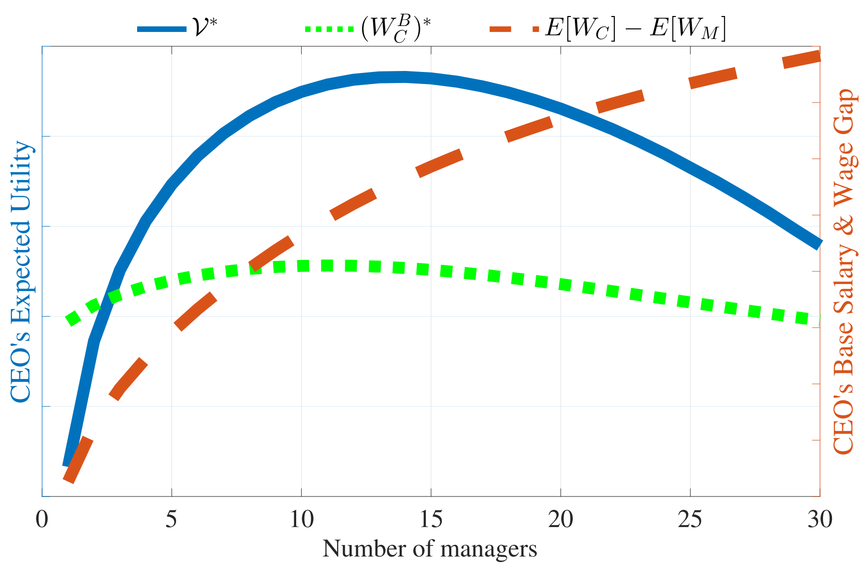

Figure 2.13: CEO's expected utility (solid line), CEO's base salary (dotted line), and Wage gap between CEO and manager (dashed line) under a multiplication specification for $\underline{U}_C = 0$. Other parameters are the same as those in **Figure 2.10**.

sation. In particular, if there is a complementarity between them, the contribution of promotion incentives to CEO compensation can vanish. These results give valuable implications for CEO compensation. Researchers need to take into account the promotion possibility besides performance based compensation when they analyze CEO compensation. Also, interpreting the role of the CEO in a firm is of importance when one tries to answer why CEO compensation has risen.

Chapter 3

INCENTIVE CONTRACT FOR A LONG-TERM PROJECT WITH MORAL HAZARD

3.1 Introduction

In many cases, a firm's project requires long-term investment or R&D before it starts generating a positive cash flow to the firm. For example, an automobile company invests its resources in order to develop a new model, and the new car provides a profit after the development is completed. A hired manger is in charge of such a project. However, it is difficult for investors to observe the manager's effort level. The manager may enjoy some private benefits instead of exerting effort. In this paper, I analyze the optimal contract problem under a moral hazard setting where an agent controls the investment process for a project.

I study a continuous-time moral hazard model in order to examine the optimal contract. This modeling is desirable since it is tractable and easily comparable to the existing literature. Specifically, I model a long-term investment by an arithmetic Brownian motion. An agent's effort choice is reflected in the drift term in the investment process. That is, the agent's effort helps the investment process complete more quickly on average. This investment does not generate any cash flow before it reaches a fixed threshold. However, once the project is completed, it delivers a stream of positive cash flow to the investors without agent's further effort.

First, I characterize the incentive compatibility condition in a general contracting space. Also, I study two essential elements of the incentive contract, the completion payment and the termination level. However, the presence of two state variables, the agent's continuation value and the investment level, makes the model intractable. Hence, I focus on a restricted contracting space for the baseline model. In the baseline model, contracting space is restricted to include only the termination level and the final completion payment fixed at the beginning of the contract. Under these restrictions, I find a unique incentive-compatible contract maximizing the investors' profit. Despite the inherent limitation of the contract, it can serve as a benchmark for a more complex contracting space. In addition, I extend the baseline model to allow for one-time adjustment of the termination level with an intermediate compensation. In my extended model, the payoff for the investors slightly increases compared to the baseline model.

Moreover, I compare my model with DeMarzo and Sannikov (2006). In their model, the agent directly controls the drift of the cash flow process. I will call such projects as short-term projects hereafter since such projects do not require any time interval between the agent's effort and cash flow. Under a comparison rule, I compare the expected profit the investors obtain from the two different types of projects. More specifically, I fix a set of parameters that make the two projects yield the same profit when there is no information asymmetry. However, the comparison of two projects with moral hazard problem shows that the short-term project can be preferred to the long-term project by the investors. This implies that the short-termism (often criticized as a moral hazard problem of agent) can actually arise for the sake of the principal. I expect that the results can provide a new insight to the short-termism issue in the literature.

The remainder of the paper is as follows. Section 2 reviews the related literature. Section 3 analyzes the agent's incentive compatibility and essential components of the incentive-compatible contract under a general contracting space. In Section 4, I examine the optimal contract under a restricted contracting space and provide some comparative statics. Section 5 compares the optimal contract for long and short-term projects. Section 6 concludes.

3.2 Related Literature

This paper is closely related to several streams of the literature. First of all, this paper builds on the literature of a hidden-action principal-agent problem, introduced by Hölmstrom (1979). Among many works in this literature, Spear and Srivastava (1987) and Rogerson (1985a) are related to my paper. They analyze a dynamic principal-agent model in discrete-time setting. Since the seminal work of Sannikov (2008), many researchers follow the novel technique to analyze an agency problem in continuous-time setting.[1] For example, DeMarzo and Sannikov (2006), Zhu (2013b), Biais et al. (2007), and He (2009) have used the methodology in order to analyze a dynamic principal-agent model where the agent controls a cash flow process. While the existing literature has full tractability since they consider only one state variable, my model requires two state variables, which makes the problem intractable. One notable exception is Cvitanic, Wan, and Yang (2013), who analyze

[1] Cvitanić, Wan, and Zhang (2009) use a stochastic maximum principle approach in order to characterize optimal contract in a similar setting.

moral hazard and adverse selection in continuous-time setting. In this paper, they consider two states variables: the agent's continuation value and temptation value.

Also, there is literature on optimal contracting problem for a long-term project. For example, Zhu (2013a) studies a myopic agency problem where there exists a tension between a short-term benefit and long-term cost. On the other hand, Sannikov (2013) studies the situation where the agent's effort has a persistent effect on the future output. My modeling of long term project is closely related to Georgiadis (2014). However, Georgiadis (2014) focuses more on agents' behavior in a group although he also considers a simple contracting problem. In this paper, I consider an extended contracting space compared to the one in Georgiadis (2014).

Another closely related research area is experimentation. For instance, Manso (2011), Hörner and Samuelson (2013), and Guo (2014) study the contracting problem when players do not know the profitability of a risky project. The main difference between this literature and my model is that players know the quality or profitability of the project in my model. However, the profitability is unknown in the literature of experimentation. While the agent's past behavior is reflected in the posterior belief on the quality of the project in the experimentation literature, it is directly reflected in the current investment level in my model. Hence, the two models have different implications. If players' main concern is the unknown quality of project, the experimentation model would be more appropriate. However, if the main concern is the accumulated effort or development to complete a project, my model specification would be more suitable.

3.3 The Model

I consider a continuous-time principal-agent model, where a principal or investors need to hire an agent in order to operate an investment process or a R&D process. If the principal decides not to hire the agent, both players receive their reservation values. The firm's cash flow process evolves according to

$$dY_t = \kappa \mathbb{1}_{\{\tau=\tau_u, \tau \leq t\}} dt,$$

where κ is a constant, and τ is a stopping time depending on the investment process. Specifically, the stopping time $\tau = \min[\tau_u, \tau_d]$ is decided by the investment process $\{I_t\}_{0\leq t\leq\tau}$ such that

$$\tau_u = \inf\{s|I_s \geq \bar{I} \text{ for } s \in [0,\infty)\} \text{ and } \tau_d = \inf\{s|I_s \leq \underline{I}(H_s) \text{ for } s \in [0,\infty)\},$$

where $\bar{I}$ is exogenously given but $\underline{I}(t)$ is determined by investors. Also, H_s denotes a history of investment process until time s. Therefore, τ is a I- measurable stopping time. $\bar{I}$ and $\underline{I}(H_t)$ represent the completion level and the termination level of the investment, respectively. Therefore, the term $\kappa \mathbb{1}_{\tau=\tau_u}$ means that if the investment process reaches the completion level before the principal terminates the project, the successful project starts generating cash flow at rate κ from the moment without any agency problem. Note that $\underline{I}(H_t)$ could be a negative infinity for every H_t. In this case, the principal never terminates the investment. In this paper, I model the publicly observable investment process by the following arithmetic Brownian motion:

$$dI_t = a_t dt + \sigma dZ_t,$$

where σ is a constant and $Z = \{Z_t, \mathcal{F}_t; 0 \leq t < \infty\}$ is a standard Brownian motion. The drift term $a_t \in \{0, \mu\}$, where $\mu > 0$, is decided by the agent's binary effort choice. Each choice gives a different cost to the agent. If the agent chooses "shirking" ($a_t = 0$), then she enjoys private benefit ϕdt for each time t. On the other hand, if she chooses "working" ($a_t = \mu$), then there is no private benefit.

Under this environment, a contract $\Gamma = (C, \underline{I}, \mathcal{B}, \mathcal{A})$ specifies a cumulative intermediate compensation $C = \{C_t\}_{t\geq 0}$ to the agent, a lower bound $\underline{I}$, the bonus payment $\mathcal{B} = (\mathcal{B}_u, \mathcal{B}_d)$ at time τ, where $\mathcal{B}_u$ is compensated to the agent if $\tau = \tau_u$ and $\mathcal{B}_d$ is provided if $\tau = \tau_d$, and a recommended effort process $\mathcal{A} = \{a_t\}_{t\geq 0}$. All four components are adapted to I.

Two players, the principal and the agent, are both risk-neutral. The principal or investors discount the future at rate $r > 0$, and the agent discounts at $\rho > r$. The agent is protected by limited liability. This implies $dC_t \geq 0$ for all t and $\mathcal{B} \geq 0$. For simplicity, assume that both players' reservation values are 0. Also, I assume that investors possess full bargaining power.

In this paper, I say that a contract Γ is *incentive-compatible* if it induces the agent to work until completion or termination. That is, a contract Γ is incentive-compatible if $\mathcal{A} = \{a_t = \mu\}_{0\leq t<\tau}$ is a solution to the following agent's problem:

$$\max_{a=\{a_t\in\{0,\mu\} \mid 0\leq t<\tau\}} E^a \left[\int_0^\tau e^{-\rho t} \left(dC_t + \phi \left(1 - \frac{a_t}{\mu} \right) dt \right) + e^{-\rho\tau} (\mathcal{B}_u \mathbb{1}_{\{\tau=\tau_u\}} + \mathcal{B}_d \mathbb{1}_{\{\tau=\tau_d\}}) \right].$$

Note that the expectation depends on the effort process $a = \{a_t \in \{0, \mu\} \mid 0 \leq t < \tau\}$. From now on, I suppress a in the expectation operator if the effort process is $\mathcal{A} = \{a_t = \mu\}_{0\leq t<\tau}$ for brevity. Moreover, I assume that parameters κ and ϕ satisfy

$\kappa > \phi$. This is a necessary condition for the incentive-compatible contract to be socially optimal.

The principal's problem is to find an incentive-compatible contract Γ maximizing his discounted expected profit

$$E\left[-\int_0^\tau e^{-rt}dC_t + e^{-r\tau}\left(\frac{\kappa}{r}\mathbb{1}_{\{\tau=\tau_u\}} - \mathcal{B}_u\mathbb{1}_{\{\tau=\tau_u\}} - \mathcal{B}_d\mathbb{1}_{\{\tau=\tau_d\}}\right)\right] - C_0,$$

where a constant C_0 is the setup cost for the project. Note that if $\phi = 0$, the principal can achieve the first best profit by choosing $\underline{I} = -\infty$, $\{C_t = 0\}_{0\le t<\tau}$, and $\mathcal{B}_u = \mathcal{B}_d = 0$, and the agent always exerts effort until the completion.[2] This policy gives the profit

$$\exp\left(-\frac{-\mu+\sqrt{\mu^2+2r\sigma^2}}{\sigma^2}(\bar{I}-I_0)\right)\frac{\kappa}{r} - C_0$$

to the principal and the reservation value to the agent. From now on, I call this profit the first best profit.

3.4 Incentive Compatibility

In this section, I characterize the agent's incentive compatibility condition and two essential components of the optimal incentive-compatible contract, $\underline{I}$ and $\mathcal{B}_u$. Among two components, $\mathcal{B}_u$ does not appear in DeMarzo and Sannikov (2006) since the contract is only terminated when the agent's continuation value reaches zero in their problem. On the other hand, the possibility of a finite $\underline{I}$ is not considered in Georgiadis (2014).

Before I analyze the incentive compatibility condition, I put some restriction on the choice of C and $\underline{I}$ in order to obtain tractability. Specifically, I only allow a finite number of intermediate compensation and a finite number of termination level updating based on the investment level. Denote K_c and K_d as the number of intermediate compensation and the number of termination level updating. Then, i-th intermediate compensation C^i is provided according to the threshold $I_{c,i}$ such that

$$dC_t = C^i \text{ if } t = \inf\{s|I_s \ge I_{c,i} \text{ for } s \in [\tau_{c,i-1},\infty)\} \text{ and } t < \inf\{s|I_s \le \underline{I}(s) \text{ for } s \in [0,\infty)\},$$

where

$$\tau_{c,i} \equiv \inf\{s|I_s \ge I_{c,i} \text{ for } s \in [\tau_{c,i-1},\infty)\} \text{ for } i = 1,2,3,\dots,K_c \text{ and } \tau_{c,0} = 0.$$

[2] I implicitly assume that the agent works if both actions give the same utility to the agent.

On the other hand, the j-th termination level I_j is adjusted by the thresholds $I_{d,j}$ such that

$$\underline{I}(H_t) = \underline{I}_j \text{ if } t = \inf\{s|I_s \geq I_{d,j} \text{ for } s \in [\tau_{d,j-1}, \infty)\} \text{ and } t < \inf\{s|I_s \leq \underline{I}(s) \text{ for } s \in [0,\infty)\},$$

where

$$\tau_{d,j} \equiv \inf\{s|I_s \geq I_{d,j} \text{ for } s \in [\tau_{d,j-1}, \infty)\} \text{ for } i = 1,2,3,\ldots,K_d \text{ and } \tau_{d,0} = 0.$$

Later, I demonstrate that the second restriction is closely related to the first restriction. For brevity, I define a threshold $I_{u,k} \in \{I_{c,1}, \ldots, I_{c,K_c}, I_{d,1}, \ldots, I_{d,K_d}\}$ such that

$$dC_t = C^k, \text{ and } \underline{I}(H_t) = \underline{I}_k$$

when

$$t = \inf\{s|I_s \geq I_{u,k} \text{ for } s \in [\tau_{u,k-1}, \infty)\} \text{ and } t < \inf\{s|I_s \leq \underline{I}(H_s) \text{ for } s \in [0,\infty)\},$$

where

$$\tau_{u,k} \equiv \inf\{s|I_s \geq I_{u,k} \text{ for } s \in [\tau_{u,k-1}, \infty)\} \text{ for } i = 1,2,3,\ldots,K_u \text{ and } \tau_{u,0} = 0,$$

allowing $C^k = 0$ and $\underline{I}_k = \underline{I}_{k+1}$, where K_u is the number of any intermediate compensation or termination level updating. Therefore, $K_u \leq K_c + K_d$. Therefore, $I_{u,k}$ represents the updating point of the intermediate compensation or the termination level.

First, I analyze the agent's problem and find the incentive-compatible condition. I denote the agent's continuation value at time t by $W(I_t|\underline{I}, C, \mathcal{B}, a)$. That is,

$$W(I_t|\underline{I}, C, \mathcal{B}, a) = E_t^a\left[\int_t^{\tau} e^{-\rho(s-t)}\left(dC_s + \phi\left(1 - \frac{a_s}{\mu}\right)ds\right) + e^{-\rho\tau}(\mathcal{B}_u \mathbb{1}_{\{\tau=\tau_u\}} + \mathcal{B}_d \mathbb{1}_{\{\tau=\tau_d\}})\right].$$

I can rewrite this continuation value using $\tau_{u,k}$. Denote $W^k(I_t|\underline{I}, C, \mathcal{B}, a)$ as the agent's continuation value for $I_t \in [I_{u,k-1}, I_{u,k})$. That is,

$$W^k(I_t) = E_t^a\left[\int_t^{\hat{\tau}} e^{-\rho(s-t)}\phi\left(1 - \frac{a_s}{\mu}\right)ds + e^{-\rho\hat{\tau}}[(C^{k+1} + W^{k+1}(I_{u,k}))\mathbb{1}_{\{\hat{\tau}=\tau_{u,k}\}} + \mathcal{B}_d \mathbb{1}_{\{\hat{\tau}=\tau_d\}}]\right],$$

where $\hat{\tau} = \min[\tau_{u,k}, \tau_d]$ is a stopping time.

The agent chooses her effort level maximizing her continuation value each time t. This maximization problem satisfies the Hamilton-Jacobi-Bellman equation

$$\rho W^k dt = \max_{\alpha} \left[\phi \left(1 - \frac{\alpha}{\mu} \right) dt + \alpha \frac{\partial W^k}{\partial I} dt + \frac{1}{2} \sigma^2 \frac{\partial^2 W^k}{\partial I^2} dt \right]$$

subject to the boundary conditions

$$W^k(\underline{I}_{k-1}) = \mathcal{B}_d \text{ and } W^k(I_{u,k}) = C^{k+1} + W^{k+1}(I_{u,k})$$

for $k = 1, 2, \ldots, K_u$.

I can characterize the incentive-compatibility condition using the equation.

Proposition 17 *The contract* Γ *is incentive-compatible if*

$$\frac{\partial W^k}{\partial I} \geq \frac{\phi}{\mu} \text{ for every } \underline{I}_{k-1} < I < I_{u,k} \text{ and } k = 1, 2, \ldots, K_u.$$

Intuitively, if the agent shirks at time t, she obtains a private benefit ϕdt. However, she loses $\mu W_I dt$ since the drift term is 0. Hence, by setting $\mu \partial W^k / \partial I \geq \phi$, the principal can incentivize the agent. Combining the incentive-compatibility condition and the limited liability condition yields the following proposition.

Proposition 18 *There is no incentive-compatible contract satisfying* $\mathcal{B}_u = 0$.

This says that the principal has to compensate for the completion of the investment in order to incentivize the agent. Unless the principal provides the payment for completion, the agent has an incentive to delay the completion since she can only enjoy the private benefit before the completion. Note that the limited liability excludes a negative payment which can make the agent incentivized with $\mathcal{B}_u = 0$. In addition to this bonus payment, **Proposition 19** shows that a finite lower bound $\underline{I}$ is the other essential component for the incentive compatibility and a positive profit to investors.

Proposition 19 *The optimal incentive-compatible contract* Γ *satisfies* $\underline{I}_k > -\infty$ *for every* k.

Although the probability of completion is equal to one if $\underline{I}(H_t)$ equal to negative infinity for every H_t, it is not optimal for the principal to set $\underline{I}(H_t)$ as such. If the current investment level is really low at time t, it takes long time to complete the investment. Therefore, for the agent, it would be better to enjoy the private benefit than to exert effort to complete the project unless she will be compensated big enough in the future. However, the compensation that makes the agent work yields a negative profit to the principal at the very low level I. Also, the investment level can get to the low level with a positive probability. Hence, the principal cannot achieve both objectives, incentive-compatibility and positive profit, simultaneously if the principal does not set a finite termination level.

Proposition 18 and **19** indicate that the optimal incentive-compatible contract has to include a positive $\mathcal{B}_u$ and a finite $\underline{I}(H_t)$ for every t. Generally, those $\mathcal{B}_u$ and $\underline{I}(H_t)$ can change as the investment level and the agent's continuation value change. However, this general case is difficult to analyze since this problem requires to solve a partial differential equation instead of an ordinary differential equation. Therefore, I characterize the incentive compatibility condition by imposing some restrictions on the dependence of C and $\underline{I}$ upon the investment level.

3.5 Lower Bound

It is difficult to find a general optimal contract since the principal's value function depends on two state variables I and W. In this section, I restrict the contracting space and find the optimal contract under that restriction. Formally, I restrict the contracting space to $\Gamma = (\underline{I}, \mathcal{B}, \mathcal{A})$ such that $\underline{I}$ and $\mathcal{B}$ are constant. That is, I do not allow any intermediate compensation, adjustable lower bound, and adjustable bonus payment. This contract can be interpreted as a lower bound for the general optimal contract since the restricted contracting space includes two essential components in the simplest way. Although this is very restrictive, the contract could be close to the optimal one in some cases. For instance, companies may not have enough cash or budget to provide any intermediate compensation. Also, they may not be able to update the initial agreement for some reasons. In these cases, investors may focus on the final payment or fix the termination level at the beginning of employment.

3.5.1 Optimal Contract under the Restricted Contract Space

Under this restricted contracting space, the agent's continuation value is

$$W_t(\underline{I}, \mathcal{B}, a) = E_t^a \left[\int_t^\tau e^{-\rho s} \left(\phi \left(1 - \frac{a_s}{\mu} \right) ds \right) + e^{-\rho \tau} (\mathcal{B}_u \mathbb{1}_{\{\tau = \tau_u\}} + \mathcal{B}_d \mathbb{1}_{\{\tau = \tau_d\}}) \right].$$

If the effort process is $A = \{a_t = \mu\}_{0\le t<\tau}$, $W_t(I, \mathcal{B}, a)$ can be written as[3]

$$W_t(\underline{I}, \mathcal{B}, A) = \frac{\exp(\eta^-\underline{I} + \eta^+ I_t) - \exp(\eta^- I_t + \eta^+\underline{I})}{\exp(\eta^-\underline{I} + \eta^+\bar{I}) - \exp(\eta^-\bar{I} + \eta^+\underline{I})}\mathcal{B}_u + \frac{\exp(\eta^- I + \eta^+\bar{I}) - \exp(\eta^-\bar{I} + \eta^+ I)}{\exp(\eta^-\underline{I} + \eta^+\bar{I}) - \exp(\eta^-\bar{I} + \eta^+\underline{I})}\mathcal{B}_d,$$

where

$$\eta^- = \frac{-\mu - \sqrt{\mu^2 + 2\rho\sigma^2}}{\sigma^2} \text{ and } \eta^+ = \frac{-\mu + \sqrt{\mu^2 + 2\rho\sigma^2}}{\sigma^2}.$$

This equation enables one to find the final payment in a closed form. I characterize it later. Under this condition, the principal's problem is expressed by

$$\begin{aligned}&\max_{(\underline{I}, \mathcal{B}_u, \mathcal{B}_d)\in\mathbf{IC}} E\left[e^{-r\tau}\left(\frac{\kappa}{r}\mathbb{1}_{\{\tau=\tau_u\}} - \mathcal{B}_u\mathbb{1}_{\{\tau=\tau_u\}} - \mathcal{B}_d\mathbb{1}_{\{\tau=\tau_d\}}\right)\right]\\ &= \max_{(\underline{I}, \mathcal{B}_u, \mathcal{B}_d)\in\mathbf{IC}}\left[\frac{\exp(\nu^-\underline{I} + \nu^+ I_0) - \exp(\nu^- I_0 + \nu^+\underline{I})}{\exp(\nu^-\underline{I} + \nu^+\bar{I}) - \exp(\nu^-\bar{I} + \nu^+\underline{I})}\left(\frac{\kappa}{r} - \mathcal{B}_u\right) - \frac{\exp(\nu^- I_0 + \nu^+\bar{I}) - \exp(\nu^-\bar{I} + \nu^+ I_0)}{\exp(\nu^-\underline{I} + \nu^+\bar{I}) - \exp(\nu^-\bar{I} + \nu^+\underline{I})}\mathcal{B}_d\right],\end{aligned}$$

where

$$\nu^- = \frac{-\mu - \sqrt{\mu^2 + 2r\sigma^2}}{\sigma^2}, \quad \nu^+ = \frac{-\mu + \sqrt{\mu^2 + 2r\sigma^2}}{\sigma^2},$$

and **IC** means the set of $(\underline{I}, \mathcal{B}_u, \mathcal{B}_d)$ satisfying the incentive-compatibility condition. Since the setup cost does not affect the principal's choice of the optimal contract if he hires the agent, I suppress C_0 in this section.[4] Now, I characterize each component of the optimal contract.

Lemma 5 *The optimal $\mathcal{B}_d$ is equal to zero.*

When the principal can't adjust the lower bound and bonus payments, $\mathcal{B}_d$ only makes the incentivization more difficult since $\mathcal{B}_d$ gives an incentive to shirk. Hence, I define the principal's choice set as $(\underline{I}, \mathcal{B}_u)$ without loss of generality.

Lemma 6 *For given $\underline{I}$, the optimal bonus payment is*

$$\mathcal{B}_u(\underline{I}) = \frac{\exp(\eta^-\underline{I} + \eta^+\bar{I}) - \exp(\eta^-\bar{I} + \eta^+\underline{I})}{\eta^+\exp(\eta^-\underline{I} + \eta^+ I^*) - \eta^-\exp(\eta^- I^* + \eta^+\underline{I})}\frac{\phi}{\mu},$$

[3] The related mathematical result is stated in **Appendix B.2**.

[4] Unless the principal's discounted expected profit is greater than C_0, the principal does not hire the agent. However, if he decides to hire the agent, C_0 does not affect the choice of $\underline{I}$ and $\mathcal{B}$ since C_0 is a sunk cost. Hence, without loss of generality, I ignore C_0 or assume $C_0 = 0$ in the analysis.

where

$$I^* = \min\left[\frac{1}{\eta^+ - \eta^-}\ln\left(\frac{(\eta^-)^2}{(\eta^+)^2}\right) + \underline{I}, \bar{I}\right].$$

In order to incentivize the agent with any I on $[\underline{I}, \bar{I}]$, the slope of the continuation value with respect to I must be greater than $\frac{\phi}{\mu}$ for all $I \in [\underline{I}, \bar{I}]$ by **Proposition 17**. For given $\underline{I}$, the unique completion payment comes from the strict convexity of $\frac{\partial W}{\partial I}$ and the compactness of $[\underline{I}, \bar{I}]$. Here, one can see why the lower bound is essential part for the optimal contract. For a given $\underline{I}$, $\frac{\partial W}{\partial I}$ is minimized at the point I^*. If $\underline{I}$ approaches to negative infinity, I^* also goes to negative infinity. That is, there is no finite completion payment $\mathcal{B}_u$ providing incentives since $\mathcal{B}_u$ goes to positive infinity as $\underline{I}$ approaches to negative infinity. It is worth mentioning that I^* is strictly greater than $\underline{I}$. This means that $\frac{\partial W}{\partial I}$ is strictly greater than $\frac{\phi}{\mu}$ for $I \in [\underline{I}, I^*)$. That is, for I in that region, the completion payment is provided more than needed to incentivize the agent at each point. This means that when I decreases the principal can incentivize the agent without increasing the completion payment in this region. On the other hand, for $I \in (I^*, \bar{I}]$, the principal has to increase the completion payment as I decreases. Mathematically, this is reflected in the convexity of W for $I \in (I^*, \bar{I}]$, while W is concave in the region, $[\underline{I}, I^*)$. Therefore, when the investment level falls close to $\underline{I}$, there is a "self-incentive" effect. This effect does not arise if $\underline{I} = -\infty$. In this case, W is convex on the whole region. Now, the principal's problem is reduced to find the optimal $\underline{I}$ maximizing his discounted expected profit at time 0.

Proposition 20 *For given parameters* $(r > 0, \rho > r, \mu > 0, \sigma > 0, \kappa > 0, \phi, \bar{I}, I_0)$ *such that*

$$\frac{\kappa}{r} - \mathcal{B}_u(I_0) > 0, \tag{3.1}$$

there exists a unique optimal contract $(\underline{I}, \mathcal{B}_u)$ *providing a positive expected discounted profit to the principal. In this contract,* $\mathcal{B}_u$ *satisfies*

$$\mathcal{B}_u(\underline{I}) = \frac{\exp(\eta^-\underline{I} + \eta^+\bar{I}) - \exp(\eta^-\bar{I} + \eta^+\underline{I})}{\eta^+\exp(\eta^-\underline{I} + \eta^+ I^*) - \eta^-\exp(\eta^- I^* + \eta^+\underline{I})}\frac{\phi}{\mu},$$

where

$$I^* = \min\left[\frac{1}{\eta^+ - \eta^-}\ln\left(\frac{(\eta^-)^2}{(\eta^+)^2}\right) + \underline{I}, \bar{I}\right],$$

and $\underline{I}$ *is the solution to the equation*

$$\mathcal{P}'(\underline{I})\left(\frac{\kappa}{r} - \mathcal{B}_u(\underline{I})\right) - \mathcal{P}(\underline{I})\frac{\partial \mathcal{B}_u(\underline{I})}{\partial \underline{I}} = 0,$$

where

$$\mathcal{P}(\underline{I}) = \frac{\exp(\nu^-\underline{I} + \nu^+ I_0) - \exp(\nu^- I_0 + \nu^+\underline{I})}{\exp(\nu^-\underline{I} + \nu^+\bar{I}) - \exp(\nu^-\bar{I} + \nu^+\underline{I})}, \text{ and}$$

$$\mathcal{P}'(\underline{I}) = -\frac{2\sqrt{\mu^2 + 2r\sigma^2}}{\sigma^2} \exp\left(-\frac{2\mu}{\sigma^2}\underline{I}\right) \frac{\exp(\nu^- I_0 + \nu^+\bar{I}) - \exp(\nu^-\bar{I} + \nu^+ I_0)}{[\exp(\nu^-\underline{I} + \nu^+\bar{I}) - \exp(\nu^-\bar{I} + \nu^+\underline{I})]^2}.$$

If the condition (3.1) does not hold, there is no incentive-compatible contract providing a positive profit to the principal.

In the remaining paper, I call this contract the baseline contract. **Figure 3.1** shows the optimal choice of $\underline{I}$ when ($r = 0.1, \rho = 0.15, \mu = 10, \sigma = 3, \kappa = 15, \phi = 3, I_0 = 0, \bar{I} = 50$).

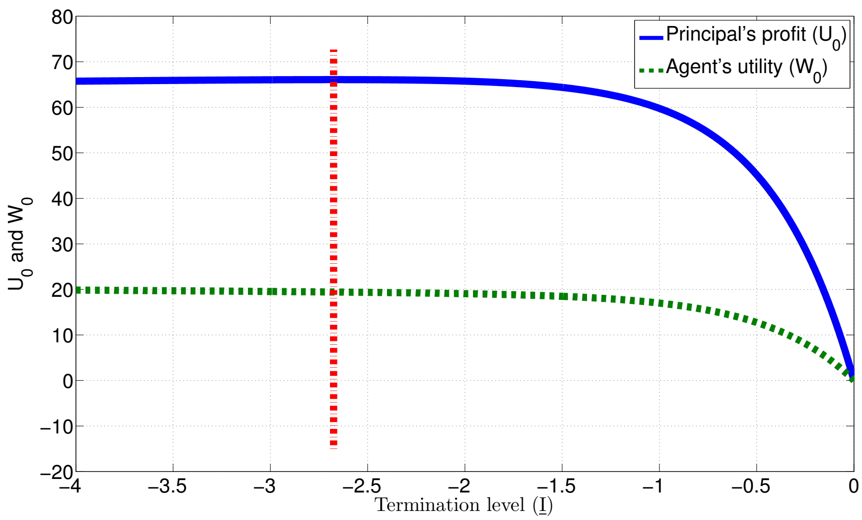

Figure 3.1: The optimal $\underline{I} = -2.6729$ and $\mathcal{B}_u = 41.0295$.

Although I drop the possibility of an intermediate payment to the agent in the baseline contract, this restriction is not severe if the principal can't adjust the termination level in the middle of investment for some reason. For example, if the verification of investment performance incurs some cost, the principal may want to avoid a frequent update of the termination level. The following corollary formally states that the intermediate compensation is closely related to the adjustment of the termination level.

Corollary 8 *For a fixed $\underline{I}$, a finite intermediate payment is not optimal in the set of incentive-compatible contract.*

Corollary 8 has two implications. First, it says that if the principal can't adjust the termination level after the initiation of the contract, he has to focus on the completion payment. Intuitively, the intermediate compensation has no role for incentivization after it is paid to the agent. Hence, the principal has to provide $\mathcal{B}_u(\underline{I})$ eventually in order to fully incentivize the agent. This means that the intermediate compensation only increases the principal's cost if the termination level is fixed. Second, this implies that if the principal provides an intermediate compensation, he has to modify the termination level. In subsection 3.5.3, I examine how the contract changes in this case.

3.5.2 Comparative Statics

In this section, I analyze some comparative statics of the optimal contract under the restricted contracting space as in the previous section. The main interests are the optimal termination level, the completion probability, and the players' discounted expected utilities. First, I define the completion probability as a function of I_0 and $\underline{I}$.[5] Formally,

$$P(I_0, \underline{I}) = Pr(\tau = \tau_u | I_0) = \frac{\exp(-\delta \underline{I}) - \exp(-\delta I_0)}{\exp(-\delta \underline{I}) - \exp(-\delta \bar{I})},$$

where $\delta = \frac{2\mu}{\sigma^2}$. **Table 3.1** summarizes the analytical findings of comparative statics.

	$\partial \underline{I}^*$	∂U_0	∂W_0	∂P
$\partial \phi$	$+$	$-$	$\pm$	$-$
∂I_0	$+$	$+$	$\pm$	$\pm$

Table 3.1: Comparative Statics for the Baseline Contract

It is particularly interesting to understand the effect of the private benefit and the initial investment level because they play an important role in the following section where I compare my model with another model based on a different specification. First, I analyze the effect of the private benefit the agent can enjoy. As the benefit increases, the optimal termination level increases and principal's expected profit decreases. The intuition is clear. The completion probability also decreases as the termination level increases. On the other hand, the agent's expected utility can go either way. Note that as ϕ increases above a certain level, the expected profit is negative and the agent will not be employed in the first place. Therefore, if ϕ is big enough, the agent's expected utility will be zero. For $\phi = 0$, it gives zero utility to

[5] See **Appendix B.2** for the related mathematical result.

the agent while the principal obtains the first-best profit. Hence, the effect depends on the value of ϕ.

Second, I also examine the effect of the initial investment level I_0. As the initial investment level gets close to $\bar{I}$, the project is getting closer to the completion. Therefore, the optimal termination level and the principal's expected profit increase. On the other hand, the direction of the agent's discounted expected utility and completion probability can be positive or negative. The completion probability decreases if the principal increases the optimal termination level significantly compared to the increment of the initial investment level. The decrease of the agent's utility can be demonstrated indirectly through the principal's first-best profit. Denote the first-best value of the profit by $FB(\mu, I_0)$. That is,

$$FB(\mu, I_0) = \exp\left(-\nu^+(\bar{I} - I_0)\right).$$

Then, doing simple calculations gives

$$\begin{aligned}
\frac{\partial FB(\mu, I_0)}{\partial \mu} &= \frac{1}{\sqrt{\mu^2 + 2r\sigma^2}} \nu^+(\bar{I} - I_0) FB(\mu, I_0) > 0, \text{ and} \\
\frac{\partial^2 FB(\mu, I_0)}{\partial \mu \partial I_0} &= \frac{1}{\sqrt{\mu^2 + 2r\sigma^2}} \nu^+ FB(\mu, I_0)(\nu^+(\bar{I} - I_0) - 1) < 0 \text{ if } \nu^+(\bar{I} - I_0) < 1.
\end{aligned}$$

As one can see, the effect of μ decreases if the initial level is really close to $\bar{I}$. This implies that the role of the agent's effort decreases when the initial level is close to $\bar{I}$. Hence, the agent's utility can decrease. Note that, in the extreme case, $I_0 = \bar{I}$, there is no need to hire the agent for the completion. Hence, in this case, the agent obtains zero utility. **Figure 3.2** and **Figure 3.3** illustrate the comparative static results for the case $(r = 0.1, \rho = 0.15, \mu = 10, \sigma = 3, \kappa = 15, \phi = 3, I_0 = 0, \bar{I} = 50)$ varying parameters I_0 and ϕ. One thing to note is that the agent's discounted expected utility sharply decreases when I_0 is really close to $\bar{I}$. From this result, I can anticipate that the agent prefers the project which is moderately far from the completion level. That is, she may prefer the project taking more time to the one that is closer to the completion. However, she does not prefer a project which is far from the completion. This result can be connected to the short-termism problem in the literature; agents prefer a moderately short-term project compared to a long-term project, but they prefer a moderately short-term project to a very short-term project.

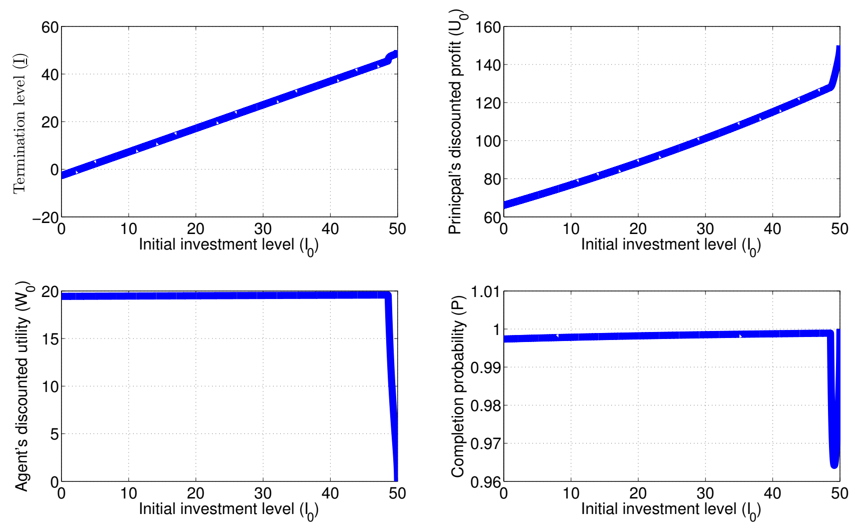

Figure 3.2: Comparative Statics with respect to I_0

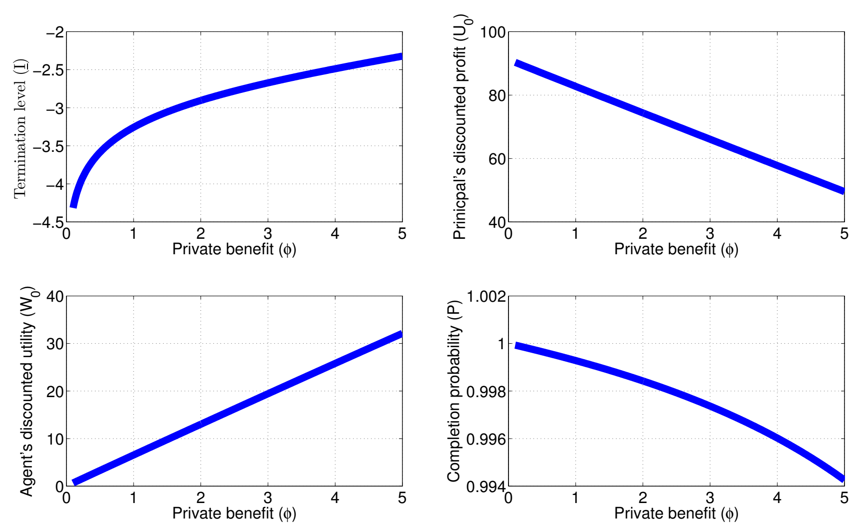

Figure 3.3: Comparative Statics with respect to ϕ

3.5.3 One Step Further

In this section, I extend the contracting space. In contrast to the baseline contract, I allow one time adjustment of the termination level with an intermediate compensation denoted by C, which can be zero. In this case, the principal's problem is much more complex since the final payment and the intermediate compensation depend on three thresholds $(\widehat{I}, \underline{I}_1, \underline{I}_2)$, where $\widehat{I}$, $\underline{I}_1$, and $\underline{I}_2$ denote the updating point, the first termination level, and the updated termination level, respectively. More specifically, the first termination level is set at $\underline{I}_1$. However, if the investment level reaches $\widehat{I}$ before it drops to the first termination level, the termination level is adjusted to $\underline{I}_2$ paying the intermediate compensation C to the agent. The following proposition shows the existence of the optimal contract and illustrates the optimal payments as the function of three threshold levels.

Proposition 21 *Suppose that parameters* $(r > 0, \rho > r, \mu > 0, \sigma > 0, \kappa > 0, \phi, \bar{I}, I_0)$ *satisfy the condition*

$$\frac{\kappa}{r} - \mathcal{B}_u(I_0) > 0.$$

Then, there exists a solution $(\widehat{I}, \underline{I}_1, \underline{I}_2, C, \mathcal{B})$ *to the principal's problem, and the optimal* $\mathcal{B}$ *and* C *are given as follows :*

$$\mathcal{B}(\underline{I}_2) \equiv \frac{\exp(\eta^- \underline{I}_2 + \eta^+ \bar{I}) - \exp(\eta^- \bar{I} + \eta^+ \underline{I}_2)}{\eta^+ \exp(\eta^- \underline{I}_2 + \eta^+ I^{**}) - \eta^- \exp(\eta^- I^{**} + \eta^+ \underline{I}_2)} \frac{\phi}{\mu},$$

where

$$I^{**} = \min\left[\frac{1}{\eta^+ - \eta^-} \ln\left(\frac{(\eta^-)^2}{(\eta^+)^2}\right) + \underline{I}_2, \bar{I}\right],$$

and

$$C(\widehat{I}, \underline{I}_1, \underline{I}_2) \equiv \max\left[\frac{\phi}{\mu} \frac{\exp(\eta^- \underline{I}_1 + \eta^+ \widehat{I}) - \exp(\eta^- \widehat{I} + \eta^+ \underline{I}_1)}{\eta^+ \exp(\eta^- \underline{I}_1 + \eta^+ I^*) - \eta^- \exp(\eta^- I^* + \eta^+ \underline{I}_1)} - \widetilde{\mathcal{P}}_A(\widehat{I}, \bar{I}, \underline{I}_2)\mathcal{B}(\underline{I}_2), 0\right],$$

where

$$I^* = \min\left[\frac{1}{\eta^+ - \eta^-} \ln\left(\frac{(\eta^-)^2}{(\eta^+)^2}\right) + \underline{I}_1, \widehat{I}\right].$$

From now on, I call this contract the extended contract. Note that this contract includes the baseline contract since setting $\widehat{I} = \bar{I}$, $\underline{I}_1 = \underline{I}^*$, and $\underline{I}_2 = \bar{I}$ gives the same profit to the principal as $(\underline{I}^*, \mathcal{B}_u(\underline{I}^*))$. Hence, this extended contract must provide at least the same profit as the baseline contract. **Figure 3.4** illustrates the principal's discounted expected profit when $(r = 0.1, \rho = 0.15, \mu = 10, \sigma = 3, \kappa = 15, \phi = 3, \bar{I} = 50)$ varying I_0 from 0 to 50.

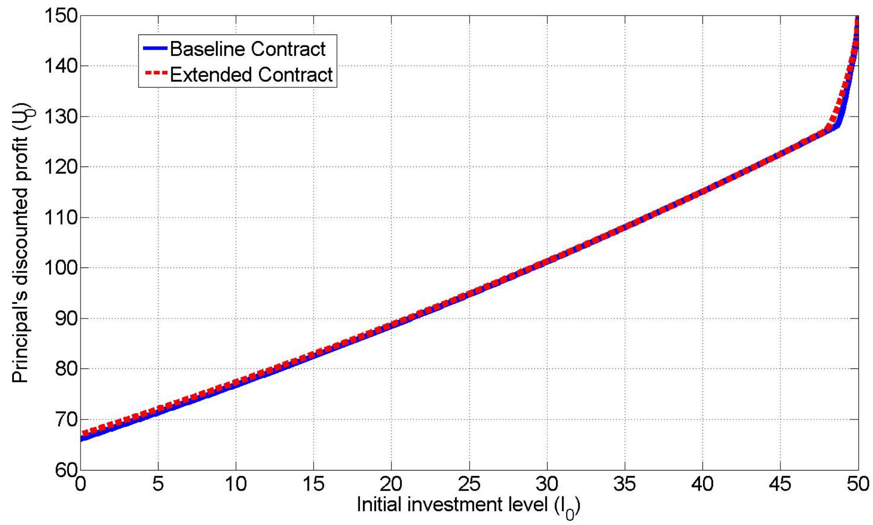

Figure 3.4: The baseline and extended contracts - Principal

This example shows that the extended contract provides slightly higher profit to the principal. According to the numerical example, the extended contract gives 1.29 percent higher profit to the principal when I_0 is equal to zero. At this point, the optimal $(\widehat{I}, \underline{I}_1, \underline{I}_2)$ is $(24.2084, -2.7348, 20.3077)$, and $(C, \mathcal{B})$ is $(8.1294, 29.1332)$. On the other hand, the baseline contract is $(\underline{I}, \mathcal{B}_u) = (-2.6729, 41.0295)$ with the same parameters. Also, the extended contract provides 2.85 percent higher profit to the principal when $I_0 = 48.78$. with $(\widehat{I}, \underline{I}_1, \underline{I}_2) = (49.3877, 47.1187, 47.7064)$ and $(C, \mathcal{B}) = (0, 10.7773)$. At this point, the baseline contract is $(\underline{I}, \mathcal{B}_u) = (47.1143, 16.3975)$. Surprisingly, the principal can achieve higher profit by adjusting the termination level without any intermediate payment. However, in this case, two termination levels should be very close in order to fully incentivize the agent. For the future research, it would be interesting to analyze a more extended case where frequent updating is allowed.

3.6 Comparison with DeMarzo and Sannikov (2006)

DeMarzo and Sannikov (2006) find the optimal contract when the agent controls the drift of a cash flow process. That is, the cash flow process evolves as

$$dY_t = \hat{a}_t dt + \hat{\sigma} d\widehat{Z}_t,$$

where $\hat{a}_t$ could be $\hat{\mu}$ or zero according to the agent's effort choice and $\widehat{Z} = \{\widehat{Z}_t, \widehat{F}_t; 0 \leq t < \infty\}$ is a standard Brownian motion. If the agent shirks ($\hat{a}_t = 0$), she enjoys the private benefit ϕdt.

The following Proposition rephrases the result in DeMarzo and Sannikov (2006), which characterizes the optimal contract implementing high effort until the termination.

Proposition 22 (Proposition 1 and 7 in DeMarzo and Sannikov (2006)) *The contract that maximizes the principal's profit and delivers the value* $W_0 \in [0, \widehat{W}]$ *to the agent takes the following form :* W_t *evolves according to*

$$dW_t = \rho W_t dt - dC_t - \phi\left(1 - \frac{a_t}{\hat{\mu}}\right) + \phi(dY_t - \hat{\mu}dt).$$

When $W_t \in [0, \widehat{W})$, $dC_t = 0$. *When* $W_t = \widehat{W}$, *payments* dC_t *cause* W_t *to reflect at* $\widehat{W}$. *If* $W_0 > \widehat{W}$, *an immediate payment* $W_0 - \widehat{W}$ *is made. The contract is terminated at time* τ *when* W_t *reaches* 0. *The principal's expected payoff at any point is given by a concave function* $b(W_t)$, *which satisfies*

$$rb(W) = \hat{\mu} + \rho W b'(W) + \frac{1}{2}\phi^2\hat{\sigma}^2 b''(W)$$

on the interval $[0, \widehat{W}]$, $b'(W) = -1$ *for* $W \geq \widehat{W}$, *and boundary conditions* $b(0) = 0$ *and* $rb(\widehat{W}) = \hat{\mu} - \rho\widehat{W}$.

In contrast to my model, DeMarzo and Sannikov (2006) consider a project in which the agent controls the cash flow directly. For instance, in an automobile company, a manager may control sales or production of existing models. Such tasks can be understood as a short-term project. On the other hand, development of a new model is considered as a long-term investment project. If the company faces limited resources, it should decide between short-term and long-term projects. Hence, it is worth comparing two different types of projects to see which project the firm will pursue.

The difference between DeMarzo and Sannikov (2006) and my model has an important implication. There are two main differences. First of all, the current investment level I_t plays a key role in my model. It measures the distance to the completion level. The lower the current level is, the longer the project remains incomplete on average. Secondly, the long-term project can be ended by both completion and termination. If the long-term project is completed, then it generates the cash flow without an agent. Hence, the firm does not suffer from an agency problem after it finishes the project. On the other hand, the firm pursuing a short-term project constantly encounters the agency problem unless they fire the agent giving up the additional cash flow. Hence, one can see that there is a tension between the length of agency problem and its intensity when a principal chooses its project. Specifically, if I_0 is closer to $\bar{I}$, the expected time to completion will be shorter. On the other hand, if I_0 is far from $\bar{I}$, it is expected to take a long time to finish the project. Therefore, the length of agency problem is determined by the distance between I_0 and $\bar{I}$. However, in the long term project, the agency problem can be more severe. Since the agent can't enjoy the private benefit once the project is completed, the agent has higher incentive to shirk delaying the completion of the project. When it comes to the short-term project, this type of incentive does not exist since the contract runs out only when the firm fires the agent. This implies that the long-term project can be worse than the short-term project with respect to the agency problem. Based on these implications, I numerically compare two projects from the principal's perspective in the following.

Before I compare the two contracts, I need to decide how to specify parameters because the comparison depends on the way I set the parameters. First, note that parameters $(r, \rho, \mu(\hat{\mu}), \phi, \sigma(\hat{\sigma}))$ appear in both specifications. In order to reduce complexity in comparison, I use the same values of those parameters for both models. Note that the long-term project also depends on four additional variables; κ, C_0, I_0, and $\bar{I}$. Since only the distance between I_0 and $\bar{I}$ matters, I fix I_0 as zero without loss of generality. After that, I specify κ and C_0 according to the following equations

$$\frac{\hat{\mu}}{r} = \exp\left(-\nu^{+}(\bar{I} - I_0)\right)\frac{\kappa}{r} - C_0, \text{ and} \tag{3.2}$$

$$0 = \exp\left(-\frac{\sqrt{2r}}{\sigma}(\bar{I} - I_0)\right)\frac{\kappa}{r} - C_0. \tag{3.3}$$

The equation (3.2) says that both specifications provide the same profit to the principal when there is no moral hazard problem. On the other hand, the equation (3.3) means that both projects provide the same profit if $a_t = \hat{a}_t = 0$ all the time. That is, if the principal can operate each project without a manager or agent, both projects give the same profit. These two conditions make the two project comparable in the sense that they have the same net present value ignoring the cost and the benefit related to the agent.

Figure 3.5, **Figure 3.6**, and **Figure 3.7** show the numerical results based on parameters ($r = 0.1, \rho = 0.15, \sigma = \hat{\sigma} = 12, \mu = \hat{\mu} = 10, I_0 = 0$) varying the degree of agency problem (ϕ) for three different $\bar{I}$, 10, 50, and 100. The results show that the long-term project could provide higher profit to the principal when the completion level is close enough to the initial level and the agent's private benefit is not high enough. However, if $\bar{I}$ is very far from the initial investment level, the long-term project provides lower profit than the short-term for all ϕ values. These results provide one important implication regarding "short-termism" : if one take into account the agency problem, the short-term project could be the best choice for the principal or investors.

This comparison has two critical limitations. First of all, the parameter σ ($\hat{\sigma}$) does not have the same effect on both specifications. In DeMarzo and Sannikov (2006), $\hat{\sigma}$ only affects the principal negatively since it reflects the unobservability of the agent's action. On the other hand, in my specification, higher σ can provide higher profit to the principal because a higher volatility can help the investment process to reach the completion level. This property is reflected in the equation (3.3). That is, the long-term project provides a positive profit to the principal although the drift term is equal to zero if there is no setup cost. This is different from the short-term project which gives zero profit if the drift term of cash flow process is equal to zero. The other limitation arises from the setup cost C_0. If C_0 is fixed, the principal's profit is a strictly increasing function in I_0 by the comparative static result. However, if C_0 satisfied the condition (3.3), the principal's profit can decrease as I_0 increases since the setup cost is also an increasing function in I_0. **Figure 3.8** shows that this could happen under the parameters ($r = 0.1, \rho = 0.15, \sigma = 12, \mu = 10, \bar{I} = 100$) varying I_0 from 0 to $\bar{I}$.

In summary, two observations (a higher σ can increase the principal's profit and I_0 closer to $\bar{I}$ can decrease the principal's profit) make the interpretation not clear. Nonetheless, this result can provide a research direction regarding the short-termism issues.

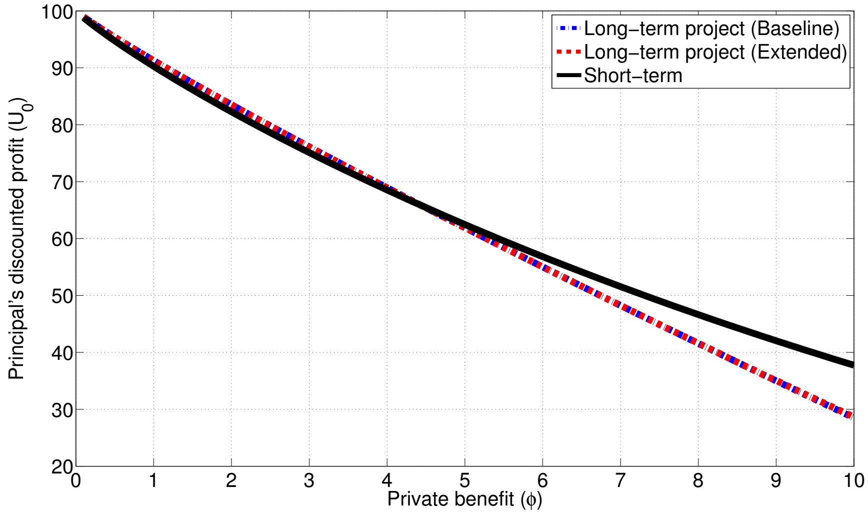

Figure 3.5: Comparison between the short-term and long-term contract when $\bar{I}$ = 10.

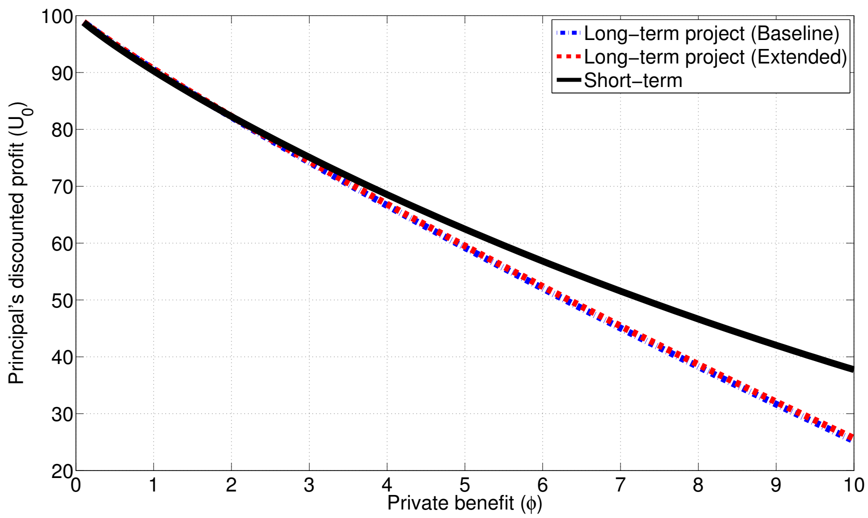

Figure 3.6: Comparison between the short-term and long-term contract when $\bar{I}$ = 50.

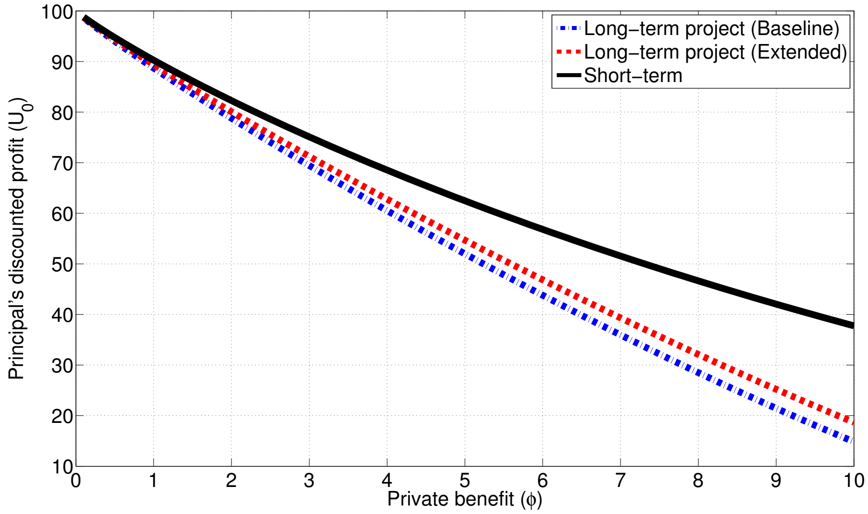

Figure 3.7: Comparison between the short-term and long-term contract when $\bar{I}$ = 100.

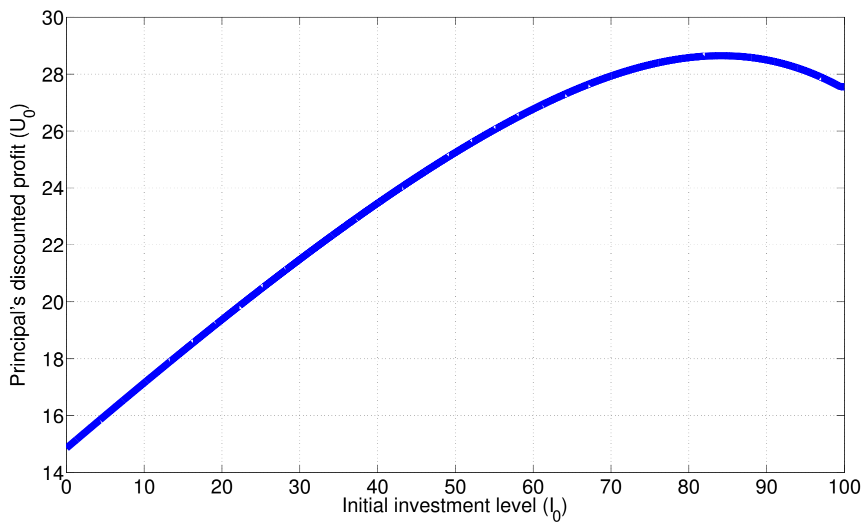

Figure 3.8: Principal's discounted profit as I_0 changes.

3.7 Conclusion

In this paper, I examine the optimal contract problem when the agent controls a long-term investment process. I characterize the incentive compatibility condition in the general contracting space. The characterization shows that there are two essential components for the optimal incentive-compatible contract, a termination level and a completion payment. Based on these results, I find the optimal contract under a restricted contracting space which includes the two components in a tractable way. This result shows that the principal can obtain a positive profit while fully incentivizing the agent. Moreover, the comparative static results demonstrate that the agent prefers a project which is moderately far from the completion level while the principal always prefers the project closer to the completion. Also, I extend the contracting space by adding a one-time intermediate compensation and termination level updating. The numerical results show that this extension makes the principal slightly better off. Finally, I compare my result with DeMarzo and Sannikov (2006). This comparison gives an interesting insight regarding the short-termism problem. That is, the principal herself could prefer a short-term project to a long-term project if there is an agency problem.

In the future research, a more complex or general contracting space can be considered. This analysis can provide a tighter comparison between a long-term and short-term project. Also, one can incorporate a repeated relation between the principal and the agent. That is, after the completion of the long-term project, the principal may re-hire the manager by assigning another long-term investment or short-term project to her. I expect that a repeated relation would reduce the agency cost. In the aspect of model specification of a long-term project, one can introduce a time-varying cost or κ. These settings will make the problem more difficult since the value of the project changes over time. Another interesting direction is to combine my model with the short-term project model such that an agent can assign her time or effort between two projects in order to maximize her utility according to the contract. To characterize optimal contract with multi-tasks is an interesting topic for future research.

C h a p t e r 4

A CAREER CHOICE PROBLEM WITH INFORMATION ASYMMETRY IN ABILITY

4.1 Introduction

Many scholars study two-sided matching problems between agents and organizations. They mostly focus on finding matching mechanisms satisfying some properties or the properties of a given mechanism. That is, researchers have examined how the mechanism works after agents decide a set of organizations to be matched with. When it comes to a career choice problem, however, people first think about which career path to pursue and then decide what to do in a specific labor market. In this context, a labor market should be interpreted as the aggregation of possible career paths in a particular field. For example, each labor market has a certain requirement such as MBA degree or a medical degree. In this paper, I examine how people behave when there is more than one labor market on the market side with an emphasis on agents' career choice.

I consider agents with different ability level such as academic ability or communication skill. I assume that the agents' abilities are independently and identically distributed and the distribution is common knowledge to every agent. However, each agent's ability is private information. On the market side, throughout the paper, I focus on the case with two labor markets. For example, I can consider two markets for undergraduate students: 1) career path with a postgraduate degree and 2) career path without a postgraduate degree. A college student can enter an economics Ph.D. program to be an economist or pursue an accounting career with a bachelor's degree. In order to focus on the information asymmetry between agents, I assume that agents and career paths are matched according to their ranking (positive assortative matching) after agents make their career decisions. In other words, an agent with higher ability experiences a better career path in each market.

Two labor markets are distinguished by its payoff distribution. I assume that one market's upper bound and lower bound are both higher than those of the other market. In terms of the above example, the best career path with a postgraduate degree is better than the best path with a bachelor's degree and also the worst situation with a graduate education is better than the worst case without a postgraduate degree.

Agents choose one labor market to enter and this career decision is irreversible. The irreversibility and incomplete information produce the tension between the probability of being matched and the expected payoff.

The tension results from two key elements of the model. First, the two labor markets may differ in competitiveness. In this paper, competitiveness of a market is defined as the ratio of the number of applicants who apply to the market to the number of career paths in it. Notice that the more competitive the market is, the lower the acceptance rate is in that market. Going back to the example of undergraduate students, if the number of agent pursuing a graduate education are much larger than the number of jobs requiring postgraduate degrees, then the employment rate of graduate schools would be very low. In this case, even though career paths with postgraduate degree provide much better payoffs, students may participate in the less desirable market. The other component to consider is the applicants' relative ranking in each market. The ranking is important even when markets have enough career paths for applicants since a low-ranked agent in one market could get a higher payoff in the other market. That is, the outcome of the career choice is not straightforward since the information about ability is private information when agents make their choice over two markets.

In order to analyze the agents' behavior in this setting, I consider a symmetric Bayesian Nash equilibrium. Importantly, I find that there is no pure strategy equilibrium and there is a unique symmetric equilibrium in which agents are divided into two groups according to their ability. Members of the high ability group uses a pure strategy and only apply to the more desirable market. On the other hand, members of the low ability group apply to both markets with positive probability. Also, in order to investigate the effect of asymmetric information on the market side, I consider the competitiveness of a market. The analysis illustrates which market is to be more competitive with the presence of information asymmetry. Despite some limitations, I believe that the results can serve as benchmark for more complex environment and research.

The remainder of the paper is as follows. Section 2 reviews the related literature. Section 3 describes the model of career choice. Section 4 introduces a simple example and section 5 defines the equilibrium and shows its existence and uniqueness. In section 7, I introduce the concept of competitiveness and compare the competitiveness of each market. Section 7 extends the model and section 8 concludes.

4.2 Related Literatures

This paper is closely related to several streams of literatures. First of all, the model is based on two-sided matching problem. Since Gale and Shapley (1962) propose the deferred acceptance algorithm, two-sided matching markets have been extensively studied[1] . Especially, Roth (1984) and Abdulkadiroglu and Sonmez (2003) introduce two-sided matching concept in order to analyze medical residencies and school choice problem, respectively. While they deal with only one market, hospitals or schools, on the side of organizations, I consider two markets case.

Second, this model contains a signaling problem. By choosing a market, agents can reveal or signal their private information to organizations. Since the seminal work by Spence (1973), signaling problems have been much studied in economics. Particularly, Spence (1973) analyzes a situation with costly signal when there is a private information. This paper shows that education may serve as signal conveying information of agents' type. Recently, Bilancini and Boncinelli (2013) develop a model where disclosure of information is costly only for one side although both sides have private information. On the other hand, Satterthwaite and Shneyerov (2007) consider dynamic matching model under two-sided incomplete information and participation costs. Unlike other papers, my model does not include any explicit cost of revealing the private information, but implicit cost incurred by exclusiveness of the two markets.

Finally, this paper can be interpreted as an extension of career or major choice problem in labor economics. There is a huge literature in labor economics focusing on major or career choice problem[2] . Most of the literature considers dynamic choice models where agents' ability and preferences are unknown and have some uncertainty. Therefore, these studies focuses on agents' optimal choice when they have imperfect information of their own characteristics. Noticeably, McCann et al. (2015) study dynamic model where agents can choose between becoming teacher and worker in a firm when individuals are characterized with two characteristics: communication and cognitive skills. However, they do not consider any informational friction. In contrast with these literatures, in my paper, every agent knows their own ability exactly and preferences and agents' types are fixed. Therefore, the key factor affecting agents' choice problem is the uncertainty on relative ranking among agents.

---

[1] See Roth (2008).

[2] See Altonji, Blom, and Meghir (2012)

4.3 Model

4.3.1 Preliminaries

Consider the following career choice problem. One side is composed of agents having private information on their ability. On the other side, there are two exclusive labor markets; each market consists of a set of career paths. The distributions of the two markets' payoffs are common knowledge and the possible outcomes' upper bound and lower bound of one market are both higher than those of the other market. From now on, I call the market with higher bounds a high-paying career, and the other market a low-paying career.

Agents decide which career to pursue between these two labor markets. For simplicity, I assume that the cost of market participation is zero, which implies that every agent enters one of two markets. Moreover, in this paper, I exclude a possibility of pursuing both careers simultaneously.[3] Once agents make their career choices over two markets, participants and possible career paths are matched according to their ranking (positive assortative matching). That is, the best agent among those pursuing the high-paying career experiences the best path of the high-paying career. If the number of participants in the high-paying career or the low-paying career is greater than the number of possible career paths in a market, low-ranked agents are unmatched and gain the outside option for the market. In the benchmark model, I assume that the unmatched agents cannot apply to the other market even when there are remaining positions in that market.

4.3.2 Agents

There are I agents and their (one dimensional) ability θ_i, $i = 1, \ldots, I$, is a private information. However, the fact that θ_i is independently distributed according to a distribution function $F(\cdot)$ is common knowledge. I assume that the function $F(\theta)$ is continuously differentiable on a finite support $[\underline{\theta}, \bar{\theta}]$ with $F'(\theta) > 0$ for all $\theta \in (\underline{\theta}, \bar{\theta})$. There are n possible career paths in the high-paying career which are represented by $H_1, H_2, \ldots, H_n$, and m paths in the low-paying career, represented by $L_1, L_2, \ldots, L_m$. I consider the case that $I \geq n + m$ and $n, m \geq 2$. Every agent has the following identical preferences over possible career paths:

$$H_1 \succ H_2 \succ \cdots \succ H_n, \; L_1 \succ L_2 \succ \cdots \succ L_m, \; H_1 \succ L_1, \text{ and } \; H_n \succ L_m.$$

[3] Although this is a limitation, it may not be severe if career paths include career-change in the future.

Moreover, if an agent pursuing the high-paying career is unmatched, she will get the outside option, O_H. Also, unmatched low-paying career participants obtain the outside option, O_L. The two outside options satisfy $H_n \succ O_H$, $L_m \succ O_L$ and $L_1 \succ O_H$, where the last condition makes the career choice problem not trivial.[4] I do not impose any other restriction on the relationship between O_H and O_L. That is, $O_H \succ O_L$, $O_L \succ O_H$, and $O_H \sim O_L$ are all possible. For completeness, I assume that every career path is acceptable to all agent and the two outside options are preferred to the outcome from non-participation. Therefore, agent i's action space consists of

$$A_i = \{\mathcal{H}, \mathcal{L}\},\ i = 1, 2, \ldots, N,$$

where $\mathcal{H}$ and $\mathcal{L}$ denote the participation in the high-paying career and the low-paying career, respectively.

I denote the utility of an agent by $u(M)$ when she is matched with a career path M. Here, I assume that agents' utility does not depend on their ability but only the matched career path. Moreover, I consider the case where agents and career paths in each market are matched based on the ranking in ability and payoff. That is, a higher ability agent gets a better payoff in each market. I also assume that every agent is appropriate to all career paths regardless of their ability.

4.3.3 Sequence of the Game

The game is played according to following steps:

1. Agents observe their ability θ_i, $i = 1, \ldots, I$.
2. Agents decide whether to pursue the high-paying career or the low-paying career.
3. Agents are matched with career paths according to their ranking.
4. If the number of agents who participate in one of two labor markets is greater than the number of possible paths, then low-ranked agents remain unmatched and obtain the outside option associated with the market.

[4] If $O_H \gtrsim L_1$, every agent pursues the high-paying career for sure.

4.3.4 Equilibrium Concept

I focus on a symmetric Bayesian Nash equilibrium to characterize the agents' behavior in the career choice problem. Recall that an agent i's action space is composed of two elements:

$$A_i = \{\mathcal{H},\ \mathcal{L}\}.$$

Therefore, for a given other agents' strategy s_{-i}, she chooses her strategy to maximize her expected utility:

$$s_i^*(s_{-i}) = \arg \max_{s_i \in \Delta(A_i)} E[u(\cdot)|(s_i, s_{-i})].$$

From now on, I suppress the subscript i for brevity since I consider a symmetric equilibrium.

4.4 Simple Example

In this section, I consider a simple example in order to illustrate the tension between pursuing a more desirable career and the probability of employment. In the economy, there are four agents and their abilities are uniformly distributed on $[0, 1]$. On the other hand, there are two high-paying career paths and two low-paying career paths. Every agent has the following identical utility function:

$$u(H_1) = 4, u(H_2) = 2, u(L_1) = 3, u(L_2) = 1, \text{ and } u(O_H) = u(O_L) = 0,$$

where H_i and L_i, $i = 1, 2$, represent high-paying and low-paying career paths, respectively. Also, O_H and O_L denote the outside options for each labor market. If there is no information asymmetry, agents and career paths are matched in the following way:

$$\mu_{Complete} = \begin{pmatrix} H_1 & H_2 & L_1 & L_2 \\ S_1 & S_3 & S_2 & S_4 \end{pmatrix},$$

where S_i represents ith ranked agent.[5] Therefore, under the complete information, agents choose their career based on their relative ranking compared to other agents since this ranking determines their payoffs. However, when there is information asymmetry, agents have to make decision without the information about the ranking.

[5]Without loss of generality, I can ignore the case where some agents have the same ability since the distribution function for the ability is continuous.

In equilibrium, there is a cut-off ($\widehat{\theta} \approx 0.8318$) such that agents above the cut-off ability pursue the high-paying career and agents below the cut-off attend both labor markets with a positive probability. The solid line in **Figure 4.1** indicates the equilibrium strategy for each type under incomplete information[6] . That is, $g(\theta)$ is the probability that an agent with ability θ pursues the high-paying career. In order to compare this with the complete information case, **Figure 4.1** also includes the ex-ante probability (dotted line) that each agent with ability θ concentrates on the high-paying career. In other words, the probability of being first or third ranked among four agents.

It is worth mentioning that high-ability agents ($\theta \geq \widehat{\theta}$) pursue the high-paying career with probability one under incomplete information although there is a chance that they are ranked second or fourth. Based on the equilibrium strategy and the ex-ante probability, the solid line and dotted line in **Figure 4.2** illustrate the expected utility when agents follow the strategy with information asymmetry and the ex-ante expected utility of each type with complete information, respectively. The result shows that the cut-off type agent experiences the biggest utility loss when agents' ability is private information compared to the complete information case. I generalize this example and characterize the equilibrium behavior in the following sections.

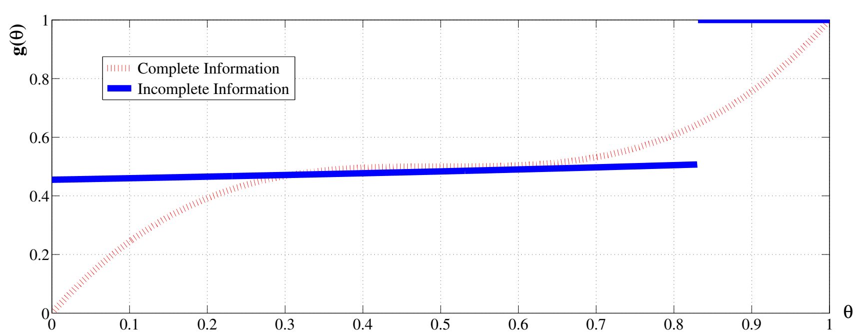

Figure 4.1: Equilibrium strategy $g(\theta)$

4.5 Equilibrium

4.5.1 Pure Strategy Equilibrium

First, I prove that there is no symmetric pure strategy equilibrium in this problem.

Proposition 23 *There is no symmetric pure strategy equilibrium.*

[6] See the appendix for the derivation.

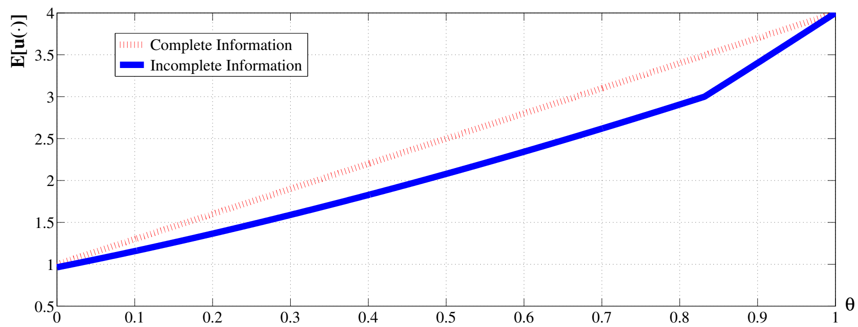

Figure 4.2: Expected Utilities

Intuitively, there must exist a cut-off such that agents with higher ability than the cut-off pursue the high-paying career with probability one since they are matched with H_1 with very high probability. Notice that an agent with the cut-off ability must obtain the same expected utility whatever she does in the equilibrium. Now, assume that there is a non-degenerate interval below the cut-off such that agents with ability in the interval use pure strategy. Denote this interval by $[a, b]$, where $a < b$. If an agent with ability a uses the same (pure) strategy as one with ability b, her expected utility should be strictly less than that of the agent with ability b in equilibrium. However, if she uses the other (pure) strategy[7] , she can obtain the same expected utility as the agent with ability b since there is no competitors with ability $\theta \in (a, b)$. This is a contradiction. Hence, there is no symmetric pure strategy equilibrium.

4.5.2 Mixed Strategy Equilibrium

In this section, I examine the existence and uniqueness of the symmetric mixed strategy equilibrium.

Proposition 24 *Among integrable functions $g(\theta)$, there exists a unique symmetric mixed strategy equilibrium (s^*). This has the following form:*

Agents with ability $\theta \geq \widehat{\theta}$ pursue the high-paying career with probability one. On the other hand, agents with ability $\theta < \widehat{\theta}$ pursue the high-paying career with probability $g(\theta)$ and the low-paying career with probability $1 - g(\theta)$, where $g(\theta)$ is given by:

$$g(\theta) = -\frac{\mathcal{N}(\theta)}{\mathcal{M}(\theta) - \mathcal{N}(\theta)}, \ \underline{\theta} \leq \theta < \widehat{\theta},$$

[7] Note that there are only two pure strategies: $\mathcal{H}$ and $\mathcal{L}$.

where

$$\mathcal{M}(\theta_i) = \frac{\partial E[u(\cdot)|(s_i(\theta_i) = \mathcal{H}, s^*_{-i}(\theta_{-i}))]}{\partial \theta_i} \ \ \textit{and} \ \ \mathcal{N}(\theta_i) = \frac{\partial E[u(\cdot)|(s_i(\theta_i) = \mathcal{L}, s^*_{-i}(\theta_{-i}))]}{\partial \theta_i}.$$

I provide the sketch of the proof of this proposition. Details are in the appendix.

Proof 2 ***Claim 1*** *The cut-off ability $\widehat{\theta}$ is strictly greater than $\underline{\theta}$, the lower bound of the type space.*

If every agent pursues the high-paying career, low ability agents can increase their expected utility by deviating from the strategy since they will be matched with L_1 for sure if they participate in the other labor market. Hence, the cut-off must be strictly greater than $\underline{\theta}$.

For the next claim and the rest of this paper, I define a function $D(\cdot)$ as the difference in the expected utility between pursuing the high-paying career and the low-paying career:

$$D(\theta) = E[u(\cdot)|(s_i(\theta) = \mathcal{H}, s^*_{-i}(\theta_{-i}))] - E[u(\cdot)|(s_i(\theta_i) = \mathcal{L}, s^*_{-i}(\theta_{-i}))].$$

Also, I introduce two functions:

$$\mathcal{P}(a,b) := \left(\int_a^b f(\theta)g(\theta)d\theta\right), \quad \textit{and} \quad Q(a,b) = \left(\int_a^b f(\theta)(1-g(\theta))d\theta\right).$$

$\mathcal{P}(a,b)$ represents the ex-ante probability that an agent pursues the high-paying career with ability $\theta \in (a,b)$ Similarly, $Q(a,b)$ is the ex-ante probability that an agent attend the labor market of the low-paying career paths with ability $\theta \in (a,b)$.

Claim 2 *There exists a unique value of $Q(\underline{\theta},\widehat{\theta}) \in (0,1)$ such that $D(\underline{\theta}) = 0$.*

For the lowest ability agent, the only factor to take into consideration is the probability that other agents participate in the labor market of the low-paying career paths ($Q(\underline{\theta},\widehat{\theta})$) or the probability that other agents pursue the high-paying career ($1-Q(\underline{\theta},\widehat{\theta})$) since it is not possible (probability is zero) that there is an agent whose ability is less than or equal to $\underline{\theta}$. Since $D(\underline{\theta})$ should be zero in equilibrium, $Q(\underline{\theta},\widehat{\theta})$ must have the unique value to guarantee the condition, $D(\underline{\theta}) = 0$.

Claim 3 *There exists a unique $\widehat{\theta} < \bar{\theta}$ such that $D(\widehat{\theta}) = 0$ and $D(\underline{\theta}) = 0$.*

First, note that $D(\bar{\theta}) = u(H_1) - u(L_1) > 0$ *since she will be matched with the best career path in the labor market she attends. Therefore, the cut-off* $\widehat{\theta}$ *should be strictly lower than the upper bound,* $\bar{\theta}$*. The condition* $D(\bar{\theta}) > 0$ *also implies that it is better for a high-ability to pursue the high-paying career since the probability of being matched with the career path* H_1 *is very high. However, this chance decreases as the ability gets lower. At the end, for an agent with the cut-off ability* $\widehat{\theta}$*, both strategies should provide the same expected utility in equilibrium. That is,* $D(\widehat{\theta})$ *should be zero, and there is a unique* $\widehat{\theta}$ *satisfying this condition and* $D(\underline{\theta}) = 0$*. In words, given an environment,* ***Claim 2*** *tells us that the ex-ante probability that an agent participates in the labor market of the low-paying career paths should be fixed. This fixed probability leads us to* ***Claim 3****, which determines the cut-off ability. That is, there is a certain cut-off such that an agent with a higher ability than the cut-off use the pure-strategy (pursue the high-paying career).*

Claim 4 *For given* $\widehat{\theta}^*$ *and* $Q(\underline{\theta}, \widehat{\theta}^*)$*, there is a unique value of* $Q(\underline{\theta}, \theta)$ *satisfying* $D(\theta) = 0$ *for each* $\theta in (\underline{\theta}, \widehat{\theta})$*.*

Since $Q(\underline{\theta}, \widehat{\theta})$ *and* $\widehat{\theta}$ *are fixed,* $Q(\underline{\theta}, \theta)$ *determines the value of* $D(\theta)$*. For a given* θ*, this claim means that there is a unique probability* $g(\theta)$ *since* $Q(\underline{\theta}, \theta)$ *only depends on* $g(\theta)$*. Therefore, in equilibrium, an agent with ability* $\theta \in [\widehat{\theta}, \bar{\theta})$ *pursues the high-paying career with probability one, and an agent with ability* $\theta \in [\underline{\theta}, \widehat{\theta})$ *pursues the low-paying career with probability* $1 - g(\theta) \in (0, 1)$*.*

4.5.3 Characterization

I have proved the existence of the equilibrium strategy $g(\theta)$ for each agent with ability θ. In this section, I characterize the equilibrium strategy. In particular, I show that there is a discontinuity of $g(\theta)$ at $\widehat{\theta}$.

Proposition 25 $g(\theta)$ *has a discontinuity at* $\widehat{\theta}$*. That is,* $\lim_{\theta \uparrow \widehat{\theta}} g(\theta) < 1 = g(\widehat{\theta})$*.*

Proof 3 *Proof is in the Appendix.*

This implies that the two groups, the high-ability group and the low-ability group, are strictly divided. In other words, around the cut-off ability, a small difference in ability causes a significant change in agent's behavior.

4.6 Competitiveness

In this section, I assume that $I = n + m$. Define the level of competitiveness of a labor market including $b \in \{n, m\}$ career paths by $\frac{I\int_{\underline{\theta}}^{\bar{\theta}} f(\theta)r(\theta)d\theta}{b}$, where $r(\theta)$ is the probability that an agent with ability θ participates in the market. Therefore, the competitiveness of a market measures the expected number of applicants to the market compared to the available career paths of the market.

Therefore, the competitiveness of the high-paying career is

$$\frac{I\int_{\underline{\theta}}^{\bar{\theta}} f(\theta)g(\theta)d\theta}{n} = \frac{I(1-Q^*)}{n},$$

and that of the low-paying career is

$$\frac{I\int_{\underline{\theta}}^{\bar{\theta}} f(\theta)(1-g(\theta))d\theta}{m} = \frac{IQ^*}{m}.$$

Note that if the information is complete, both markets have the same competitiveness equal to one. In order to compare competitiveness of the two markets in equilibrium, I define two variables:

$$E[\mathcal{H}] := \sum_{j=0}^{m-1}\binom{I-1}{j}u(O_H)\left(\frac{m}{I}\right)^j\left(\frac{n}{I}\right)^{I-1-j} + \sum_{j=m}^{I-1}\binom{I-1}{j}u(H_{I-j})\left(\frac{m}{I}\right)^j\left(\frac{n}{I}\right)^{I-1-j}$$

and

$$E[\mathcal{L}] := \sum_{j=0}^{n-1}\binom{I-1}{j}u(O_L)\left(\frac{m}{I}\right)^j\left(\frac{n}{I}\right)^{I-1-j} + \sum_{j=n}^{I-1}\binom{I-1}{j}u(L_{I-j})\left(\frac{m}{I}\right)^j\left(\frac{n}{I}\right)^{I-1-j}.$$

Here, $E[\mathcal{H}]$ ($E[\mathcal{L}]$) represents the expected utility of an agent with the lowest ability among all agents if she pursue the high-paying career (the low-paying career) and all other agents randomly choose their career path according to the ratio of available career paths of a marker to the total number of career paths. That is, other agents pursue the high-paying career path with probability $\frac{n}{I}$ and the low-paying career path with probability $\frac{m}{I}$.

Proposition 26 *If $E[\mathcal{H}]$ is greater(smaller) than $E[\mathcal{L}]$, the high-paying career path is more (less) competitive than the low-paying career path.*

Proof 4 *Proof is in the Appendix.*

Competitiveness of a labor market has important meaning since unbalanced competitiveness implies inefficient allocation of human capital. In particular, if the competitiveness of a market is greater than one, some agents in this market may end up unemployed, while other marker could experience a shortage of labor supply.

4.7 Extension

4.7.1 Endogenous Outside Option

Up to this point, I have considered a situation where agents choose their action simultaneously and two markets are exclusive. This situation is proper if two markets are divided physically or each career path has a specific requirement. For example, a migrant worker has to decide where he settles in order to find a job. Once he has decided where to move, then it is difficult to move again due to moving costs. Therefore, he has to choose where to go based on possible jobs and competition in that area. However, in some cases, agents can still try to pursue the other career after they experience a career failure. For instance, a student who fails to pass the CPA exam might pursue a finance career instead. Although his comparative advantage is low compared to students who have prepared for a long time to become a financial analyst, he might still prefer having a job to being unemployed. In order to incorporate this situation, in this section, I consider a model where the outside option O_H is determined endogenously. In particular, agents who pursue the high-paying career can be matched with the remaining slots in the other market if they do not succeed in the high-paying career. In addition, I add one more assumption on agents' preferences over career paths, $H_n \succ L_1$. That is, the worst outcome in the high-paying career is preferred to the best low-paying career path. This implies that the high-paying career dominates the low-paying career in all possible payoffs. The sequence of the game is the following:

1. Agents observe their type θ_i.
2. Agents decide whether to pursue the high-paying career or the low-paying career.
3. Agents are matched with career paths according to their ranking.

4. If the number of agents who purse the low-paying career is greater than the number of low-paying career paths, then low-ranked agents remain unmatched and obtain the outside option.

   If the number of agents who pursue the high-paying career is greater than the number of high-paying career paths, the unmatched agents will be matched with the remaining low-paying career paths according to their ranking after the first round matching.

   Unmatched agents after the second round matching get the outside option.

The following proposition shows that agents behave in the similar way as **Proposition 24** in equilibrium.

Proposition 27 *There does not exist any symmetric pure strategy equilibrium. Also, among integrable functions $g(\theta)$, there exists a unique symmetric mixed strategy equilibrium. This has the following form: Agents with ability $\theta \geq \widehat{\theta}$ pursue the high-paying career with probability one. On the other hand, agents with ability $\theta < \widehat{\theta}$ pursue the high-paying career path with probability $g(\theta)$, where $g(\theta)$ is given by:*

$$g(\theta) = -\frac{\mathcal{N}_e(\theta)}{\mathcal{M}_e(\theta) - \mathcal{N}_e(\theta)}, \ \underline{\theta} \leq \theta < \widehat{\theta},$$

where

$$\mathcal{M}_e(\theta_i) = \frac{\partial E[u(\cdot)|(s_i(\theta_i) = \mathcal{H}, s^*_{-i}(\theta_{-i}))]}{\partial \theta_i} \ \text{ and } \ \mathcal{N}_e(\theta) = \frac{\partial E[u(\cdot)|(s_i(\theta_i) = \mathcal{L}, s^*_{-i}(\theta_{-i}))]}{\partial \theta_i}.$$

Proof 5 *Proof is in the Appendix.*

Note that this situation can be interpreted as a dynamic situation with a discount factor equal to one. That is, agents who pursue the high-paying career at $t = 1$ can attend the other labor market in the second round if they are not successful in their first career.

The next question is how this opportunity for changing career affects agent's equilibrium behavior. In particular, when does the opportunity make the high-paying career more desirable? In order to answer this question, I compare the equilibrium cut-off ability for the exogenous outside option case with that for the endogenously determined outside option case.

Proposition 28 *Denote $\widehat{\theta}$ as the equilibrium cut-off when O_H is given exogenously and $\widehat{\theta}_e$ as the equilibrium cut-off when the outside option is determined endogenously with $u(O_L) = 0$. Then, there is a value $0 < \mathcal{K} < u(H_n)$ such that if $u(O_H) > \mathcal{K}$, then $\widehat{\theta} < \widehat{\theta}_e$. However, if $u(O_H) < \mathcal{K}$, then $\widehat{\theta} > \widehat{\theta}_e$. Moreover, $\widehat{\theta} = \widehat{\theta}_e$ if $u(O_H) = \mathcal{K}$.*

Proof 6 *Proof is in the Appendix.*

In words, if the outside option given exogenously for an agent to pursue the high-paying career is greater (lower) than a certain value $\mathcal{K}$, the equilibrium cut-off ability for the endogenous outside option case is greater (lower) than the equilibrium cut-off for the exogenous outside option case. Intuitively, if the outside option for pursuing the high-paying career is high enough, an agent with ability $\widehat{\theta}_e$ strictly prefers the outside option O_H to the opportunity that she can attend the other labor market if she is not successful in her career. This implies $\widehat{\theta} < \widehat{\theta}_e$. On the other hand, if both outside options (O_H and O_L) represent the same status such as "Unemployment" ($u(O_H) = u(O_L) = 0$), an agent with ability $\widehat{\theta}_e$ prefers the case with endogenous outside option case since she has one more chance to get a job, which is better than being unemployed.

4.8 Conclusion

In this paper, I examine the equilibrium behavior in a career choice problem when there are two exclusive labor markets with different payoff distributions. The result shows that there is a cut-off point in ability such that agents are divided into two groups. Members of the high ability group pursue the career where the best outcome is possible. On the other hand, members of the low ability group attend both labor markets with positive probability which depends on their ability. Moreover, if one consider the probability that an agent pursue the high-paying career as a function of agent's ability, this function also sharply divides agents into two groups. In particular, this function has a jump discontinuity at the cut-off point though this function is continuous in each group.

Going back to the high-paying versus the low-paying career example, this equilibrium behavior is consistent with reality in a sense that well-qualified students pursue the career with the best career path without fear of failure. On the other hand, students who are not strong enough spend much time agonizing over which career to pursue. This agony leads different market choice although students are similar. That is, these students use a mixed strategy in the aspect of game theory. Among a number of interpretations of mixed strategy equilibrium, the following interpretation from Osborne and Rubinstein (1994) would be appropriate for this situation:

> A player's action is a response to his guess about the other player's choice; guessing is a psychological operation that is very much deliberate and not random.

One contribution of this paper is to provide a way of introducing multi-markets on the organization side in matching problems. The paper also contributes to the literature of career choice problem in labor economics. Most papers in the literature have focused on uncertainty over the agent's own ability and learning. However, this paper shows that the information about own ability is not sufficient to explain one's career choice. Since people usually have to compete for jobs or positions with competitors they do not know well, their relative ranking, which is uncertain, is really important for their career choice. I believe that incorporating the modelling assumption in this paper into existing literature would give us much more satisfactory explanation.

In future research, more complex type spaces or heterogeneous preference relations can be incorporated into the model in order to reflect a career specific or general skills and students' different preference relations over career paths. Also, one can focus on the behavior of the organization side. Organizations might have some requirements for applicants to elevate overall quality of its members. For example, a market could impose restrictions on applicants' quality by requiring some specialized test score. Moreover, one can consider a model where payoff distribution is endogenously determined. Despite the simple structure of the model in this paper, I believe that this model can be used as a starting point of the future research on more realistic career choice problems with information asymmetry.

BIBLIOGRAPHY

Abdulkadiroglu, Atila and Tayfun Sonmez (2003). "School choice: A mechanism design approach". In: *The American Economic Review* 93.3, pp. 729–747.

Acs, Zoltan J. and David B. Audretsch (1987). "Innovation, Market Structure, and Firm Size". In: *The review of Economics and Statistics*, pp. 567–574.

Altonji, Joseph G, Erica Blom, and Costas Meghir (2012). "Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers". In: *Annual Review of Economics* 4.1, pp. 185–223.

Baker, George, Michael Gibbs, and Bengt Holmstrom (1994). "The Wage Policy of a Firm". In: *The Quarterly Journal of Economics*, pp. 921–955.

Baker, George P. and Brian J. Hall (2004). "CEO Incentives and Firm Size". In: *Journal of Labor Economics* 22.4, pp. 767–798.

Biais, Bruno et al. (2007). "Dynamic security design: Convergence to continuous time and asset pricing implications". In: *The Review of Economic Studies* 74.2, pp. 345–390.

Bilancini, Ennio and Leonardo Boncinelli (2013). "Disclosure of information in matching markets with non-transferable utility". In: *Games and Economic Behavior* 82, pp. 143–156.

Bognanno, Michael L. (2001). "Corporate Tournaments". In: *Journal of Labor Economics* 19.2, pp. 290–315.

Bolton, Patrick and Mathias Dewatripont (2005). *Contract Theory*. MIT press.

Cvitanic, Jakša, Xuhu Wan, and Huali Yang (2013). "Dynamics of contract design with screening". In: *Management Science* 59.5, pp. 1229–1244.

Cvitanić, Jakša, Xuhu Wan, and Jianfeng Zhang (2009). "Optimal compensation with hidden action and lump-sum payment in a continuous-time model". In: *Applied Mathematics and Optimization* 59.1, pp. 99–146.

Dechenaux, Emmanuel, Dan Kovenock, and Roman M. Sheremeta (2015). "A Survey of Experimental Research on Contests, All-pay Auctions and Tournaments". In: *Experimental Economics* 18.4, pp. 609–669.

DeMarzo, Peter M and Yuliy Sannikov (2006). "Optimal Security Design and Dynamic Capital Structure in a Continuous-Time Agency Model". In: *The Journal of Finance* 61.6, pp. 2681–2724.

Edlin, Aaron S. and Chris Shannon (1998). "Strict Monotonicity in Comparative Statics". In: *Journal of Economic Theory* 81.1, pp. 201–219.

Ekinci, Emre and Michael Waldman (2015). "Bonuses and Promotion Tournaments: Theory and Evidence". In: *Working paper*.

Frydman, Carola and Dirk Jenter (2010). “CEO Compensation”. In: *Annual Review of Financial Economics* 2.1, pp. 75–102.

Frydman, Carola and Raven E. Saks (2010). “Executive Compensation: A New View from a Long-Term Perspective, 1936-2005”. In: *Review of Financial Studies* 23.5, pp. 2099–2138.

Gabaix, Xavier and Augustin Landier (2008). “Why has CEO Pay Increased So Much?” In: *The Quarterly Journal of Economics* 123.1, pp. 49–100.

Gale, David and Lloyd S Shapley (1962). “College admissions and the stability of marriage”. In: *American Mathematical Monthly*, pp. 9–15.

Georgiadis, George (2014). “Projects and team dynamics”. In: *The Review of Economic Studies*, rdu031.

Goel, Anand M. and Anjan V. Thakor (2008). “Overconfidence, CEO selection, and Corporate Governance”. In: *The Journal of Finance* 63.6, pp. 2737–2784.

Grossman, Sanford J. and Oliver D. Hart (1983). “An Analysis of the Principal-agent Problem”. In: *Econometrica: Journal of the Econometric Society*, pp. 7–45.

Guo, Yingni (2014). *Dynamic Delegation of Experimentation*. Tech. rep. mimeo, Yale University.

He, Zhiguo (2009). “Optimal executive compensation when firm size follows geometric brownian motion”. In: *Review of Financial Studies* 22.2, pp. 859–892.

Hölmstrom, Bengt (1979). “Moral hazard and observability”. In: *The Bell journal of economics*, pp. 74–91.

Hörner, Johannes and Larry Samuelson (2013). “Incentives for experimenting agents”. In: *The RAND Journal of Economics* 44.4, pp. 632–663.

Jenter, Dirk and Fadi Kanaan (2015). “CEO Turnover and Relative Performance Evaluation”. In: *The Journal of Finance* 70.5, pp. 2155–2184.

Kale, Jayant R., Ebru Reis, and Anand Venkateswaran (2009). “Rank-order Tournaments and Incentive Alignment: The Effect on Firm Performance”. In: *The Journal of Finance* 64.3, pp. 1479–1512.

Ke, Rongzhu, Jin Li, and Michael Powell (2014). “Managing Careers in Organizations”. In: *Working paper*.

Konrad, Kai A. (2009). *Strategy and Dynamics in Contests*. Oxford University Press.

Lazear, Edward P. and Paul Oyer (2012). “Personnel Economics”. In: *The Handbook of Organizational Economics*. Ed. by Robert Gibbons and John Roberts. Princeton University Press, pp. 479–519. ISBN: 0691132798.

Lazear, Edward P. and Sherwin Rosen (1981). “Rank-Order Tournaments as Optimum Labor Contracts”. In: *Journal of Political Economy* 89.5.

Lemieux, Thomas, W. Bentley MacLeod, and Daniel Parent (2009). "Performance Pay and Wage Inequality". In: *The Quarterly Journal of Economics* 124.1, pp. 1–49.

Macera, Rosario (2016). "Present or Future Incentives? On the Optimality of Fixed Wages With Moral Hazard". In: *Working paper*.

Manso, Gustavo (2011). "Motivating innovation". In: *The Journal of Finance* 66.5, pp. 1823–1860.

McCann, Robert J et al. (2015). "Becker meets ricardo: Multisector matching with communication and cognitive skills". In: *Journal of Law, Economics, and Organization*, ewv002.

Milgrom, Paul and Chris Shannon (1994). "Monotone Comparative Statics". In: *Econometrica: Journal of the Econometric Society*, pp. 157–180.

Mirrlees, James A. (1999). "The Theory of Moral Hazard and Unobservable Behaviour: Part I". In: *The Review of Economic Studies* 66.1, pp. 3–21.

Mookherjee, Dilip (1984). "Optimal Incentive Schemes with Many Agents". In: *The Review of Economic Studies* 51.3, pp. 433–446.

Nalebuff, Barry J. and Joseph E. Stiglitz (1983). "Prizes and incentives: Towards a General Theory of Compensation and Competition". In: *The Bell Journal of Economics*, pp. 21–43.

Osborne, Martin J and Ariel Rubinstein (1994). *A course in game theory*. MIT press.

Prendergast, Canice (1999). "The Provision of Incentives in Firms". In: *Journal of Economic Literature* 37.1, pp. 7–63.

– (2002). "The Tenuous Trade-off between Risk and Incentives". In: *Journal of Political Economy* 110.5, pp. 1071–1102.

Rogerson, William P. (1985a). "Repeated Moral Hazard". In: *Econometrica: Journal of the Econometric Society*, pp. 69–76.

– (1985b). "The First-order Approach to Principal-agent Problems". In: *Econometrica: Journal of the Econometric Society*, pp. 1357–1367.

Roth, Alvin E (1984). "The evolution of the labor market for medical interns and residents: a case study in game theory". In: *The Journal of Political Economy* 92.6, p. 991.

– (2008). "Deferred acceptance algorithms: History, theory, practice, and open questions". In: *International Journal of Game Theory* 36.3-4, pp. 537–569.

Sannikov, Yuliy (2008). "A continuous-time version of the principal-agent problem". In: *The Review of Economic Studies* 75.3, pp. 957–984.

– (2013). *Moral hazard and long-run incentives*.

Satterthwaite, Mark and Artyom Shneyerov (2007). “Dynamic Matching, Two-Sided Incomplete Information, and Participation Costs: Existence and Convergence to Perfect Competition”. In: *Econometrica* 75.1, pp. 155–200.

Spear, Stephen E. and Sanjay Srivastava (1987). “On Repeated Moral Hazard with Discounting”. In: *The Review of Economic Studies* 54.4, pp. 599–617.

Spence, Michael (1973). “Job market signaling”. In: *The Quarterly Journal of Economics* 87.3, pp. 355–374.

Stokey, Nancy L (2008). *The economics of inaction: Stochastic control models with fixed costs*. Princeton University Press.

Tervio, Marko (2008). “The Difference That CEOs Make: An Assignment Model Approach”. In: *American Economic Review* 98.3, pp. 642–68.

Waldman, Michael (2012). “Theory and Evidence in Internal Labor Markets”. In: *The Handbook of Organizational Economics*. Ed. by Robert Gibbons and John Roberts. Princeton University Press, pp. 520–572. ISBN: 0691132798.

– (2013). “Classic Promotion Tournaments Versus Market-based Tournaments”. In: *International Journal of Industrial Organization* 31.3, pp. 198–210.

Zabojnik, Jan and Dan Bernhardt (2001). “Corporate Tournaments, Human Capital Acquisition, and the Firm Size-Wage Relation”. In: *The Review of Economic Studies* 68.3, pp. 693–716.

Zhu, John Y (2013a). *Myopic agency*. Tech. rep. Working Paper University of Pennsylvania.

– (2013b). “Optimal contracts with shirking”. In: *The Review of Economic Studies* 80.2, pp. 812–839.

Appendix A

APPENDIX TO CHAPTER 2

A.1 Proofs

A.1.1 Derivation of $P(\mathbf{e}_{-Mi})$

Note that

$$\begin{aligned} P(\mathbf{e}_{-Mi}) &= \sum_{j=0}^{N-1} \frac{1}{j+1} \binom{N-1}{j} s(e_{-Mi})^j (1-s(e_{-Mi}))^{N-1-j} \\ &= \frac{1}{Ns(e_{-Mi})} \sum_{j=0}^{N-1} \binom{N}{j+1} s(e_{-Mi})^{j+1} (1-s(e_{-Mi}))^{N-1-j} \\ &= \frac{1}{Ns(e_{-Mi})} \left(\sum_{s=0}^{N} \binom{N}{s} s(e_{-Mi})^s (1-s(e_{-Mi}))^{N-s} - (1-s(e_{-Mi}))^N \right) \\ &= \frac{1-(1-s(e_{-Mi}))^N}{Ns(e_{-Mi})}. \end{aligned}$$

A.1.2 Proof of Lemma 1

Denote the firm's profit when the CEO's IR condition binds as $\underline{\Pi}$. Then, the firm's optimal profit must be greater than or equal to $\underline{\Pi}$.

First, for an action $(e_C^*, e_M^*) \in [0,1] \times [0,1]$, I show that there exists a solution $(W_C^G, W_C^B, W_M^G, W_M^B)$ to the firm's problem. Note that when the CEO's IR condition binds the firm's problem is reduced to the case of Grossman and Hart (1983). Hence, a solution exists to this restricted problem. Now, I show that I can artificially bound the constraint set of $(W_C^G, W_C^B, W_M^G, W_M^B)$. They are bounded below by two IR conditions. Moreover, they are also bounded above since the firm's optimal profit is lower than $\underline{\Pi}$ and the firm's profit is a strictly decreasing function in all four components in (W_C^G, W_C^B, W_g, W_b) without a lower bound. Also, the constraint set is closed according to two IC and two IR conditions. Hence, there exists a solution by the Extreme value theorem. The remaining proof exactly follows the proof in Grossman and Hart (1983).

A.1.3 Proof of Lemma 2

Suppose this is not true. That is,

$$s(e_M)u(W_M^G) + (1-s(e_M))u(W_M^B) - g(e_M) + s(e_M)P(\mathbf{e}_{-M})V_C > \underline{U}_M.$$

Then, choosing new wage scheme $(\widehat{W}_M^G, \widehat{W}_M^B) = (W_M^G - \epsilon_1, W_M^B - \epsilon_2)$, where $\epsilon_1 > 0$ and $\epsilon_2 > 0$, satisfying

$$
\begin{aligned}
&u(W_M^G) - u(W_M^B) = u(\widehat{W}_M^G) - u(\widehat{W}_M^B) \text{ and} \\
&s(e_M)u(\widehat{W}_M^G) + (1 - s(e_M))u(\widehat{W}_M^B) - g(e_M) + P(\mathbf{e}_{-M})V_C \geq \underline{U}_M
\end{aligned}
$$

gives a higher profit to the firm without affecting other constraints. Hence, the wage scheme (W_M^G, W_M^B) is not optimal.

A.1.4 Proof of Lemma 3

Note that

$$\frac{\partial F(\mathcal{V})}{\partial \mathcal{V}} = -\frac{1}{u'(W_C)} + \frac{1}{u'(W_M)}$$

using the envelope theorem.

Differentiating this with respect to $\mathcal{V}$ gives

$$
\begin{aligned}
\frac{\partial^2 F(\mathcal{V})}{\partial \mathcal{V}^2} &= \frac{u''(W_C)}{u'(W_C)^3} + \frac{1}{N}\frac{u''(W_M)}{u'(W_M)^3} + \frac{u''(W_C)}{u'(W_C)^3}g'(e_C^*)\frac{\partial e_C^*}{\partial \mathcal{V}} - \frac{u''(W_M)}{u'(W_M)^3}g'(e_M^*)\frac{\partial e_M^*}{\partial \mathcal{V}} \\
&= \frac{u''(W_C)}{u'(W_C)^3} - \frac{u''(W_C)}{u'(W_C)^3}g'(e_C^*)\frac{\frac{u''(W_C)}{u'(W_C)^2}g'(e_C^*)}{\frac{u''(W_C)}{u'(W_C)^2}g'(e_C^*)^2 - g''(e_C^*)} \\
&\quad + \frac{1}{N}\frac{u''(W_M)}{u'(W_M)^3} - \frac{u''(W_M)}{u'(W_M)^3}g'(e_M^*)\frac{\frac{1}{N}\frac{u''(W_M)}{u'(W_M)^2}g'(e_M^*)}{\frac{u''(W_M)}{u'(W_M)^2}g'(e_M^*)^2 - g''(e_M^*)} \\
&= \frac{u''(W_C)}{u'(W_C)^3}\left(1 - \frac{\frac{u''(W_C)}{u'(W_C)^2}g'(e_C^*)^2}{\frac{u''(W_C)}{u'(W_C)^2}g'(e_C^*)^2 - g''(e_C^*)}\right) \\
&\quad + \frac{1}{N}\frac{u''(W_M)}{u'(W_M)^3}\left(1 - \frac{\frac{u''(W_M)}{u'(W_M)^2}g'(e_M^*)^2}{\frac{u''(W_M)}{u'(W_M)^2}g'(e_M^*)^2 - g''(e_M^*)}\right) \\
&< 0.
\end{aligned}
$$

That is, $F(\mathcal{V})$ is a strictly concave function.

A.1.5 Proof of Proposition 1

It is enough to show that

$$\frac{\partial^2 F(\mathcal{V})}{\partial N \partial \mathcal{V}} > 0.$$

Notice that

$$\begin{aligned}
\frac{\partial^2 F(\mathcal{V})}{\partial N \partial \mathcal{V}} &= -\frac{1}{N^2}\frac{u''(W_M)}{u'(W_M)^3}\mathcal{V} - \frac{u''(W_M)}{u'(W_M)^3}g'(e_M^*)\left(\frac{-\beta(\mathcal{G}_M - \mathcal{B}_M)\frac{1}{N^2}\frac{u''(W_M)}{u'(W_M)}}{\beta^2(\mathcal{G}_M - \mathcal{B}_M)^2 u''(W_M) - g''(e_M^*)}\right)\mathcal{V} \\
&= -\frac{1}{N^2}\frac{u''(W_M)}{u'(W_M)^3}\left(1 - \frac{\beta(\mathcal{G}_M - \mathcal{B}_M)\frac{u''(W_M)}{u'(W_M)}g'(e_M^*)}{\beta^2(\mathcal{G}_M - \mathcal{B}_M)^2 u''(W_M) - g''(e_M^*)}\right)\mathcal{V} \\
&= -\frac{1}{N^2}\frac{u''(W_M)}{u'(W_M)^3}\left(1 - \frac{\beta^2(\mathcal{G}_M - \mathcal{B}_M)^2 u''(W_M)}{\beta^2(\mathcal{G}_M - \mathcal{B}_M)^2 u''(W_M) - g''(e_M^*)}\right)\mathcal{V} \\
&> 0.
\end{aligned}$$

A.1.6 Proof of Corollary 2

Suppose this is not the case. From the first order condition

$$\beta(\mathcal{G}_C - \mathcal{B}_c) = \frac{g'(e_C)}{u'(W_C)},$$

it can be shown that e_C^* and W_C^* move in the opposite direction since the left hand side is a constant. Therefore, e_C^* should increase if W_C^* decreases. Since $\mathcal{V}^*$ increases as N increases, W_C^* must increase according to

$$u(W_C^*) = \mathcal{V}^* + g(e_C^*).$$

This contradicts the premise that W_C^* decreases. Hence,

$$\frac{\partial W_C^*}{\partial N} > 0.$$

A.1.7 Proof of Proposition 2

Recall that the first order condition is

$$-\frac{1}{u'(W_C^*)} + \frac{1}{u'(W_M^*)} = 0.$$

Hence, W_C^* should be the same as W_M^*.

If $\mathcal{G}_C - \mathcal{B}_C = \mathcal{G}_M - \mathcal{B}_M$, using the previous result and two first order conditions, it can be shown that

$$g'(e_C^*) = g'(e_M^*).$$

That is, $e_C^* = e_M^*$.

Also, this result and the two individual rationality constraints imply that

$$u(W_M) - g(e_M) + \frac{1}{N}\mathcal{V}^* = \mathcal{V}^* + \frac{1}{N}\mathcal{V}^* = \underline{U}_M.$$

Hence,

$$\mathcal{V}^* = \frac{N}{N+1}\underline{U}_M.$$

A.1.8 Proof of Proposition 3

Note that when agents are risk-neutral

$$\begin{aligned}\frac{\partial F(\mathcal{V})}{\partial \mathcal{V}} &= -s(e_C) - (1 - s(e_C)) + N * s(e_M)P(\mathbf{e}_{-M}) \\ &= -1 + (1 - (1 - s(e_M))^N) \\ &< 0\end{aligned}$$

using the envelope theorem.[1] Hence, the firm's profit decreases as the level of $\mathcal{V}$ increases.

A.1.9 Proof of Lemma 4

Using the envelope theorem,

$$\frac{\partial F(\mathcal{V})}{\partial \mathcal{V}} = -\frac{s(e_C)}{u'(W_C^G)} - \frac{1 - s(e_C)}{u'(W_C^B)} + Ns(e_M)P(\mathbf{e}_{-M})\frac{1}{u'(W_M^G)}.$$

Differentiating this with respect to $\mathcal{V}$ gives

$$\frac{\partial^2 F(\mathcal{V})}{\partial \mathcal{V}^2} = s(e_C)\frac{u''(W_C^G)}{u'(W_C^G)^3} + (1 - s(e_C))\frac{u''(W_C^B)}{u'(W_C^B)^3} + Ns(e_M)P(\mathbf{e}_{-M})^2\frac{u''(W_M^G)}{u'(W_M^G)^3} < 0.$$

A.1.10 Proof of Proposition 4

It is enough to show that

$$\frac{\partial^2 F(\mathcal{V})}{\partial N \partial \mathcal{V}} > 0.$$

Note that

$$\begin{aligned}\frac{\partial^2 F(\mathcal{V})}{\partial N \partial \mathcal{V}} &= -(1 - s(e_M))^N \log(1 - s(e_M))\frac{1}{u'(W_M^G)} \\ &\quad + [1 - (1 - s(e_M))^N]\frac{\partial P(\mathbf{e}_{-M})}{\partial N}\mathcal{V}\frac{u''(W_M^G)}{u'(W_M^G)^3} > 0,\end{aligned}$$

where

$$\begin{aligned}\frac{\partial P(\mathbf{e}_{-M})}{\partial N} &= \frac{1}{s(e_M)N^2}\left[-\log(1 - s(e_M))(1 - s(e_M))^N N - 1 + (1 - s(e_M))^N\right] \\ &< 0\end{aligned}$$

since $k(s) \equiv -\log(1-s)(1-s)^N N - 1 + (1-s)^N$ is equal to zero when $s = 0$ and $k'(s) < 0$. Here, I use the condition that $\mathcal{V} \geq 0$.

A.1.11 Proof of Corollary 4

Denote the expected compensation to CEO by $E[W_C]$,

$$E[W_C] = s(e_C)W_C^G + (1 - s(e_C))W_C^B.$$

[1] In equilibrium, $s(e_M)$ is equal to $s(e_{-M})$ since I am considering a symmetric equilibrium.

Since $\frac{\partial \mathcal{V}^*}{\partial N} > 0$, it is enough to show that

$$\frac{\partial E[W_C]}{\partial \mathcal{V}} = \frac{s(e_C)}{u'(W_C^G)} + \frac{1 - s(e_C)}{u'(W_C^B)} > 0.$$

A.1.12 Proof of Corollary 5

I need to show that the wage gap

$$\left[s(e_C)(W_C^G)^* + (1 - s(e_C))(W_C^B)^*\right] - \left[s(e_M)(W_M^G)^* + (1 - s(e_M))(W_M^B)^*\right]$$

widens as N increases. Since $(1 - s(e_M))(W_M^B)^*$ has a fixed value regardless of the number of managers, it is enough to show that

$$\left[s(e_C)(W_C^G)^* + (1 - s(e_C))(W_C^B)^*\right] - s(e_M)(W_M^G)^*$$

is an increasing function in N. When agents have the log utility function, the first order condition with respect to $\mathcal{V}$ is

$$s(e_C)(W_C^G)^* + (1 - s(e_C))(W_C^G)^* = (1 - (1 - s(e_M))^N)(W_M^G)^*.$$

Since the left hand side of the equation is a strictly increasing function in $\mathcal{V}$ and $\frac{\partial \mathcal{V}^*}{\partial N} > 0$, this side strictly increases as N increases. Hence,

$$\frac{\partial}{\partial N}(1 - (1 - s(e_M))^N)(W_M^G)^* = -\log(1 - s(e_M))(1 - s(e_M))^N (W_M^G)^* + (1 - (1 - s(e_M))^N)\frac{\partial (W_M^G)^*}{\partial N} > 0.$$

Since

$$\left[s(e_C)(W_C^G)^* + (1 - s(e_C))(W_C^B)^*\right] - s(e_M)(W_M^G)^* = (1 - s(e_M) - (1 - s(e_M))^N)(W_M^G)^* \tag{A.1}$$

$$\begin{aligned}
&\frac{\partial}{\partial N}\left\{\left[s(e_C)(W_C^G)^* + (1 - s(e_C))(W_C^B)^*\right] - s(e_M)(W_M^G)^*\right\} = \\
&\quad -\log(1 - s(e_M))(1 - s(e_M))^N (W_M^G)^* + (1 - s(e_M) - (1 - s(e_M))^N)\frac{\partial (W_M^G)^*}{\partial N} \\
&> -\log(1 - s(e_M))(1 - s(e_M))^N (W_M^G)^* + (1 - s(e_M) - (1 - s(e_M))^N)\cdot \\
&\quad \frac{\log(1 - s(e_M))(1 - s(e_M))^N (W_M^G)^*}{1 - (1 - s(e_M))^N} \\
&= -\frac{s(e_M)}{1 - (1 - s(e_M))^N}\log(1 - s(e_M))(1 - s(e_M))^N (W_M^G)^* \\
&> 0
\end{aligned}$$

when $N > 1$. Also, when $N = 1$, the wage gap is equal to zero according to A.1. On the other hand, the gap has a positive value when $N = 2$ since $(W_M^G)^* > 0$. Therefore, the expected compensation gap is a strictly increasing function in N.

A.1.13 Proof of Proposition 5

First, I consider a case when $e_C = e_M$.

Note that
$$F(\underline{U}_M|N) = -\frac{s(e_C)}{u'(W_C^G)} - \frac{1-s(e_C)}{u'(W_C^B)} + \frac{1-(1-s(e_M))^N}{u'(W_M^G)},$$
where

$$\begin{aligned}
u(W_C^G) &= \underline{U}_M + g(e_C) + (1-s(e_C))\frac{g'(e_C)}{h'(e_C)},\\
u(W_C^B) &= \underline{U}_M + g(e_C) - s(e_C)\frac{g'(e_C)}{h'(e_C)}, \text{ and}\\
u(W_M^G) &= \underline{U}_M + g(e_M) + (1-s(e_M))\frac{g'(e_M)}{h'(e_M)} - P(\mathbf{e}_{-M})\underline{U}_M.
\end{aligned}$$

Denote the difference between $1/u'(W_C^G)$ and $1/u'(W_C^B)$ by $\mathcal{D}$.

For given $(1-s(e_C))\mathcal{D} > \epsilon > 0$, there is $\widehat{N}$ such that

$$\frac{1}{u'(W_C^G)} - \frac{1-(1-s(e_M))^N}{u'(W_G)} < \epsilon$$

when $N \geq \widehat{N}$ since $P(\mathbf{e}_{-M}) \to 0$ and $(1-s(e_M))^N \to 0$ as $N \to \infty$. Therefore, when $N \geq \widehat{N}$,

$$\begin{aligned}
F(\underline{U}_M|N) &= -\frac{s(e_C)}{u'(W_C^G)} - \frac{1-s(e_C)}{u'(W_C^B)} + \frac{1-(1-s(e_M))^N}{u'(W_g)}\\
&> -\frac{s(e_C)}{u'(W_C^G)} - \frac{1-s(e_C)}{u'(W_C^B)} + \frac{1}{u'(W_C^G)} - \epsilon\\
&= (1-s(e_C))\left(\frac{1}{u'(W_C^G)} - \frac{1}{u'(W_C^B)}\right) - \epsilon\\
&> 0.
\end{aligned}$$

Since $F(\mathcal{V}|N)$ is a strictly concave in $\mathcal{V}$, $\mathcal{V}^* > \underline{U}_M$ when $N \geq \widehat{N}$.

Second, I show that there is N^* such that $\mathcal{V}^* > \underline{U}_M$ if $N > N^*$ when $0 < e_C < e_M < 1$.

There are two possibilities;

$$\frac{s(e_M)}{u'(W_M^G)} \geq \frac{1}{u'(W_C^G)} \text{ or } \frac{s(e_M)}{u'(W_M^G)} < \frac{1}{u'(W_C^G)}$$

when $\mathcal{V} = \underline{U}_M$ and $N = 1$.

1. $\left(\frac{s(e_M)}{u'(W_M^G)} \geq \frac{1}{u'(W_C^G)}\right)$

This condition implies that

$$
\begin{aligned}
F(\underline{U}_M|1) &= -\frac{s(e_C)}{u'(W_C^G)} - \frac{1-s(e_C)}{u'(W_C^B)} + \frac{s(e_M)}{u'(W_M^G)} \\
&\geq (1-s(e_C))\left(\frac{1}{u'(W_C^G)} - \frac{1}{u'(W_C^B)}\right) \\
&> 0.
\end{aligned}
$$

Since $\frac{\partial \mathcal{V}^*}{\partial N} > 0$, $\mathcal{V}^* > \underline{U}_M$ for every N.

2. $\left(\frac{s(e_M)}{u'(W_M^G)} < \frac{1}{u'(W_C^G)}\right)$

Again, denote the difference between $1/u'(W_C^G)$ and $1/u'(W_C^B)$ by $\mathcal{D}$. Then, for given $(1-s(e_C))\mathcal{D} > \epsilon$, there exists $\widehat{N}$ such that

$$
0 \leq \frac{1}{u'(W_C^G)} - \frac{1-(1-s(e_M))^N}{u'(W_M^G)} < \epsilon.
$$

Therefore, when $N \geq \widehat{N}$,

$$
\begin{aligned}
F(\underline{U}_M|N) &= -\frac{s(e_C)}{u'(W_C^G)} - \frac{1-s(e_C)}{u'(W_C^B)} + \frac{1-(1-s(e_M))^N}{u'(W_M^G)} \\
&> -\frac{s(e_C)}{u'(W_C^G)} - \frac{1-s(e_C)}{u'(W_C^B)} + \frac{1}{u'(W_C^G)} - \epsilon \\
&= (1-s(e_C))\left(\frac{1}{u'(W_C^G)} - \frac{1}{u'(W_C^B)}\right) - \epsilon \\
&> 0.
\end{aligned}
$$

Since $F(\mathcal{V}|N)$ is a strictly concave in $\mathcal{V}$, $\mathcal{V}^* > \underline{U}_M$ when $N \geq \widehat{N}$.

A.1.14 Proof of Proposition 6

Under the given assumption, it can be shown that

$$
\begin{aligned}
\frac{\partial^2 F(\mathcal{V})}{\partial \mathcal{V} \partial e_C} &= -\beta\left(\frac{1}{u'(W_C^G)} - \frac{1}{u'(W_C^B)}\right) + s(e_C)(1-s(e_C))\frac{g''(e_C)}{\beta}\left(\frac{u''(W_C^G)}{u'(W_C^G)^3} - \frac{u''(W_C^B)}{u'(W_C^B)^3}\right) \\
&< 0, \\
\frac{\partial^2 F(\mathcal{V})}{\partial \mathcal{V} \partial e_M} &= N\beta\frac{(1-s(e_M))^{N-1}}{u'(W_g)} \\
&\quad - [1-(1-s(e_M))^N]\frac{u''(W_M^G)}{u'(W_M^G)^3}\left((1-s(e_M))\frac{g''(e_M)}{\beta} - \frac{\partial P(\mathbf{e}_{-M})}{\partial e_M}\mathcal{V}\right) > 0.
\end{aligned}
$$

A.1.15 Proof of Proposition 7

First, I show that $\mathcal{V}^*$ increases as $\underline{U}_M$ increases. From the first order condition with respect to $\mathcal{V}^*$:

$$-\frac{s(e_C)}{u'(W_C^G)} - \frac{1-s(e_C)}{u'(W_C^B)} + \frac{1-(1-s(e_M))^N}{u'(W_M^G)} = 0,$$

$$\frac{\partial \mathcal{V}^*}{\partial \underline{U}_M} = \frac{[1-(1-s(e_M))^N]\frac{u''(W_M^G)}{u'(W_M^G)^3}}{s(e_C)\frac{u''(W_C^G)}{u'(W_C^G)^3} + (1-s(e_C))\frac{u''(W_C^B)}{u'(W_C^B)^3} + Ns(e_M)P(\mathbf{e}_{-M})^2\frac{u''(W_M^G)}{u'(W_M^G)^3}} > 0.$$

Note that for a given (N, e_M), $(W_M^G)^* = (W_M^B)^*$ if

$$\mathcal{V}^* = \frac{g'(e_M)}{\beta P(\mathbf{e}_{-M})}$$

since $u(W_M^G) - u(W_M^B) = \frac{g'(e_M)}{\beta} - P(\mathbf{e}_{-M})\mathcal{V}$. For given (N, e_C, e_M), denote $\mathcal{V}$ satisfying $(W_M^G)^* = (W_M^B)^*$ by $\widehat{\mathcal{V}}$. That is,

$$\widehat{\mathcal{V}} = \frac{g'(e_M)}{\beta P(\mathbf{e}_{-M})} > 0.$$

Then,

$$\left.\frac{\partial F(\mathcal{V}|\underline{U}_M)}{\partial \mathcal{V}}\right|_{\mathcal{V}=\widehat{\mathcal{V}}} = -\frac{s(e_C)}{u'(W_C^G)} - \frac{1-s(e_C)}{u'(W_C^B)} + \frac{1-(1-s(e_M))^N}{u'(W_M^G)}$$

There are two possible cases, $\left.\frac{\partial F(\mathcal{V}|\underline{U}_M)}{\partial \mathcal{V}}\right|_{\mathcal{V}=\widehat{\mathcal{V}}} \geq 0$ and $\left.\frac{\partial F(\mathcal{V}|\underline{U}_M)}{\partial \mathcal{V}}\right|_{\mathcal{V}=\widehat{\mathcal{V}}} < 0$.

If $\left.\frac{\partial F(\mathcal{V}|\underline{U}_M)}{\partial \mathcal{V}}\right|_{\mathcal{V}=\widehat{\mathcal{V}}} \geq 0$, then $\mathcal{V}^* > \widehat{\mathcal{V}}$. This implies that $(W_M^G)^* \leq (W_M^B)^*$. Suppose that $\left.\frac{\partial F(\mathcal{V}|\underline{U}_M)}{\partial \mathcal{V}}\right|_{\mathcal{V}=\widehat{\mathcal{V}}} < 0$. Since $\mathcal{V}$ is fixed at $\widehat{\mathcal{V}}$,

$$\lim_{\underline{U}_M \to \infty} \frac{1-(1-s(e_M))^N}{u'(W_M^G)} = \infty.$$

Hence, there is $0 < \underline{U}_M^* < \infty$ such that

$$\left.\frac{\partial F(\mathcal{V}|\underline{U}_M)}{\partial \mathcal{V}}\right|_{\mathcal{V}=\widehat{\mathcal{V}}} = 0.$$

Since $\frac{\partial \mathcal{V}^*}{\partial \underline{U}_M} > 0$, $(W_M^G)^* \leq (W_M^B)^*$ if $\underline{U}_M \geq \underline{U}_M^*$.

A.1.16 Proof of Corollary 6

It is enough to show that there is $\underline{U}_M$ such that $W_C^B \geq W_M^G$ when $\mathcal{V} = 0$ since this implies that

$$\left.\frac{\partial F(\mathcal{V})}{\partial \mathcal{V}}\right|_{\mathcal{V}=0} < 0.$$

Note that

$$u(W_C^B) - u(W_M^G) = g(e_C) - s(e_C)\frac{g'(e_C)}{\beta} - \underline{U}_M - g(e_M) - (1 - s(e_M))\frac{g'(e_M)}{\beta}$$

when $\mathcal{V} = 0$. Hence, if

$$\underline{U}_M \leq g(e_C) - s(e_C)\frac{g'(e_C)}{\beta} - g(e_M) - (1 - s(e_M))\frac{g'(e_M)}{\beta},$$

$W_C^B \geq W_M^G$.

Since $\frac{\partial \mathcal{V}^*}{\partial \underline{U}_M} > 0$ and there is $\underline{U}_M$ such that $\mathcal{V}^* > 0$ according to **Proposition 7**, there exists $\widetilde{\underline{U}}_M$ such that $\mathcal{V}^* = 0$ with $\frac{\partial F(\mathcal{V})}{\partial \mathcal{V}}\Big|_{\mathcal{V}=\mathcal{V}^*} = 0$. Also, if $\underline{U}_M$ is less than $\widetilde{\underline{U}}_M$, the solution is $\mathcal{V}^* = 0$ with $\frac{\partial F(\mathcal{V})}{\partial \mathcal{V}}\Big|_{\mathcal{V}=\mathcal{V}^*} < 0$.

A.1.17 Proof of Proposition 8

Before I prove the proposition, I show that $\mathcal{V}^*$ is bounded for every N. The first order condition with respect to $\mathcal{V}$ implies that $(W_M^G)^* > (W_C^B)^*$ for every N. Therefore,

$$\underline{U}_M + g(e_M) + (1 - s(e_M))\frac{g'(e_M)}{\beta} - P(\mathbf{e}_{-M})\mathcal{V}^* \geq \mathcal{V}^* + g(e_C) - s(e_C)\frac{g'(e_C)}{\beta}.$$

Since $0 \leq P(\mathbf{e}_{-M}) \leq 1$,

$$\underline{U}_M + g(e_M) + (1 - s(e_M))\frac{g'(e_M)}{\beta} - g(e_C) + s(e_C)\frac{g'(e_C)}{\beta} > \mathcal{V}^*,$$

where the left hand side does not depend on N.

First, I denote the optimal compensations by $(W_C^G(N), W_C^B(N), W_M^G(N), W_M^B)$ for a given N.[2]

Then, the difference between the two profit per agent is

$$\begin{aligned}2(N+1)(\Pi(\mathcal{V}^*|N) - \Pi(\mathcal{V}^*|1)) &= (N-1)(s(e_M) - s(e_C))(\mathcal{G} - \mathcal{B}) \\ &\quad - 2W_C(N) + (N+1)W_C(1) - 2NW_M(N) + (N+1)W_M(1),\end{aligned}$$

where

$$\begin{aligned}W_C(N) &= s(e_C)W_C^G(N) + (1 - s(e_C))W_C^B(N) \\ W_M(N) &= s(e_M)W_M^G(N) + (1 - s(e_M))W_M^B(N).\end{aligned}$$

Since $\mathcal{V}^*$ is bounded, optimal compensations $(W_C(N), W_M(N), W_C(1), W_M(1))$ are also bounded. Also, they are not depend on $\mathcal{G}$ and $\mathcal{B}$. Hence, there is $\mathcal{G}^* - \mathcal{B}^*$ such that the difference has a positive value for a given N since $s(e_M) > s(e_C)$.

[2]Note that W_M^B does not depend on the number of candidates.

For the second part, I impose a restriction on e_M. Namely, for a given e_C. e_M satisfies the condition that

$$\mathcal{V}^*(e_C, e_M | N = 1) + \frac{g'(e_C)}{\beta} - \frac{g'(e_M)}{\beta} \geq 0,$$

where $\mathcal{V}^*(e_C, e_M | N = 1)$ is the optimal $\mathcal{V}$ when $N = 1$ for a given (e_C, e_M). Since $\mathcal{V}^*(e_C, e_M | N = 1) + \frac{g'(e_C)}{\beta} - \frac{g'(e_M)}{\beta} > 0$ if $e_M = e_C$, there is $\bar{e}_M$ such that $\mathcal{V}^*(e_C, e_M | N = 1) + \frac{g'(e_C)}{\beta} - \frac{g'(e_M)}{\beta} \geq 0$ if $e_M \in (e_C, \bar{e}_M]$.

Note that

$$\begin{aligned}
\Pi(\mathcal{V}^*|N) - \Pi(\mathcal{V}^*|1) = &\frac{N-1}{2(N+1)}(s(e_M) - s(e_C))(\mathcal{G} - \mathcal{B}) \\
&- \frac{1}{2}s(e_M)(W_M^G(N) - W_M^G(1)) \\
&+ \frac{1}{2(N+1)}[-(N-1)W_M(N) - 2W_C(N) + (N+1)W_C(1)] \\
< &\frac{N-1}{2(N+1)}(s(e_M) - s(e_C))(\mathcal{G} - \mathcal{B}) \\
&- \frac{1}{2}s(e_M)(W_M^G(N) - W_M^G(1)) \\
&+ \frac{1}{2(N+1)}[-(N-1)W_M(N) - 2W_C(N) + (N-1)W_M(1) + 2W_C(N)] \\
= &\frac{N-1}{2(N+1)}(s(e_M) - s(e_C))(\mathcal{G} - \mathcal{B}) \\
&- \frac{N}{N+1}s(e_M)(W_M^G(N) - W_M^G(1)).
\end{aligned}$$

The inequality holds since $W_C(N) > W_C(1)$ and $W_M(1) > W_C(1)$. Note that $W_C(N) > W_C(1)$ is true because $\frac{\partial \mathcal{V}^*}{\partial N} > 0$. On the other hand, the condition imposed on e_M guarantees that $W_M(1) > W_C(1)$.[3] Here, I show that $W_M(1) > W_C(1)$ if $e_M \in (e_C, \bar{e}_M]$. The first order condition with respect to $\mathcal{V}$ when $N = 1$ is

$$\frac{s(e_C)}{u'(W_C^G(1))} + \frac{1 - s(e_C)}{u'(W_C^B(1))} = \frac{s(e_M)}{u'(W_M^G(1))},$$

which implies that $u(W_M^G(1)) > u(W_C^G(1))$. Therefore,

$$\underline{U}_M + g(e_M) + (1 - s(e_M))\frac{g'(e_M)}{\beta} > 2\mathcal{V}^* + g(e_C) + (1 - s(e_C))\frac{g'(e_C)}{\beta}.$$

This inequality and the condition on e_M imply that

$$\begin{aligned}
u(W_M^B) &= \underline{U}_M + g(e_M) - s(e_M)\frac{g'(e_M)}{\beta} \\
&> \mathcal{V}^* + g(e_C) - s(e_C)\frac{g'(e_C)}{\beta} + \mathcal{V}^* + \frac{g'(e_C)}{\beta} - \frac{g'(e_M)}{\beta} \\
&\geq \mathcal{V}^* + g(e_C) - s(e_C)\frac{g'(e_C)}{\beta} \\
&= u(W_C^B(1)).
\end{aligned}$$

[3] If agents have the log utility function, the condition is not needed.

Therefore, $W_M(1) > W_C(1)$.

Hence, if

$$\mathcal{G} - \mathcal{B} < \frac{2N}{N-1}\frac{s(e_M)}{s(e_M)-s(e_C)}\left[W_M^G(N) - W_M^G(1)\right],$$
$$\Pi(\mathcal{V}^*|N) < \Pi(\mathcal{V}^*|1).$$

Notice that there is $\widehat{N}$ such that $W_M^G(N) > W_M^G(1)$ if $N > \widehat{N}$ since $P(\mathbf{e}_{-M})$ converges to zero as N approaches infinity and $\mathcal{V}^*$ is bounded. Let $\overline{O}$ denote

$$\inf_{N\in[\widehat{N},\infty)} \frac{2N}{N-1}\frac{s(e_M)}{s(e_M)-s(e_C)}\left[W_M^G(N) - W_M^G(1)\right].$$

Then, $\Pi(\mathcal{V}^*|N) < \Pi(\mathcal{V}^*|1)$ if $N > \widehat{N}$ and $\mathcal{G} - \mathcal{B} \leq \overline{O}$.

A.1.18 Proof of Proposition 9

Proof 7 *First, note that*

$$\Pi(\mathcal{V}^*|N) - \Pi(\mathcal{V}^*|1) = \frac{(N+1)W_C(1) - 2W_C(N) + (N+1)W_M(1) - 2NW_M(N)}{2(N+1)},$$

where $W_C(k)$ and $W_M(k)$ represent CEO's and managers' expected compensation when the firm hires k managers[4] *, respectively. When agents have the log utility function, the first order condition with respect to $\mathcal{V}$ is*

$$s(e)W_C^G(N) + (1-s(e))W_C^B(N) = (1-(1-s(e))^N)W_M^G(N).$$

Using this condition, it can be shown that

$$\begin{aligned}\Pi(\mathcal{V}^*|N) - \Pi(\mathcal{V}^*|1) &= s(e)[W_M^G(1) - W_M^G(N)] \\ &\quad - \frac{(1-s(e))(1-(1-s(e))^{N-1})W_M^G(N)}{N+1} - \frac{N-1}{2(N+1)}(1-s(e))W_M^B.\end{aligned}$$

This indicates that $\Pi(\mathcal{V}^*|N) - \Pi(\mathcal{V}^*|1) < 0$ if $W_M^G(N) > W_M^G(1)$.

Also, when $u(x) = \log(x)$ and $e_C = e_M = e$, $\mathcal{V}(N)$ is[5]

$$\begin{aligned}\mathcal{V}^*(N) &= -\frac{1}{1+P(\boldsymbol{e}_{-M})}\cdot \\ &\quad \log\left[\frac{s(e_C)\exp\left[g(e_C) + (1-s(e_C))\frac{g'(e_C)}{\beta}\right] + (1-s(e_C))\exp\left[g(e_C) - s(e_C)\frac{g'(e_C)}{\beta}\right]}{(1-(1-s(e_M))^N)\exp\left[\underline{U}_M + g(e_M) + (1-s(e_M))\frac{g'(e_M)}{\beta}\right]}\right] \\ &= -\frac{1}{1+P(\boldsymbol{e}_{-M})}\log\left[\frac{s(e) + (1-s(e))\exp\left[-\frac{g'(e)}{\beta}\right]}{(1-(1-s(e))^N)\exp\left[\underline{U}_M\right]}\right] \\ &= \frac{1}{1+P(\boldsymbol{e}_{-M})}\underline{U}_M + \frac{1}{1+P(\boldsymbol{e}_{-M})}\log\left[\frac{(1-(1-s(e))^N)}{s(e) + (1-s(e))\exp\left[-\frac{g'(e)}{\beta}\right]}\right].\end{aligned}$$

[4]These are defined in A.1.17.

[5]In this proof, I explicitly indicate the dependency of the variable on N.

Note that when $\underline{U}_M = -\log\left[\frac{s(e)}{s(e)+(1-s(e))\exp\left[-\frac{g'(e)}{\beta}\right]}\right]$, $\mathcal{V}^*(1) = 0$, *and*

$$\mathcal{V}^*(N) = \frac{1}{1+P(\boldsymbol{e}_{-M})}\log\left[\frac{(1-(1-s(a))^N)}{s(a)}\right].$$

Therefore, $\Pi(\mathcal{V}^*|N) - \Pi(\mathcal{V}^*|1)$ *is greater than* 0 *if*

$$s(e) - \left[\frac{Ns(e)+1-(1-s(a))^N}{N+1}\right]\exp[-P(\boldsymbol{e}_{-M}|N)\mathcal{V}^*(N)] \\ - \frac{N-1}{2(N+1)}(1-s(e))\exp\left(-\frac{g'(e)}{\beta}\right) > 0.$$

This is equivalent to

$$1 > \frac{N}{N+1}\frac{1+P(\boldsymbol{e}_{-M}|N)}{[NP(\boldsymbol{e}_{-M}|N)]^{\frac{P(\boldsymbol{e}_{-M}|N)}{1+P(\boldsymbol{e}_{-M}|N)}}} + \frac{N-1}{2(N+1)}\frac{1-s(e)}{s(e)}\exp\left(-\frac{g'(a)}{\beta}\right).$$

For a fixed N, the first term on the right hand side is increasing in the agents' effort level e since

$$\frac{\partial}{\partial e}\left(\frac{1+P(\boldsymbol{e}_{-M}|N)}{[NP(\boldsymbol{e}_{-M}|N)]^{\frac{P(\boldsymbol{e}_{-M}|N)}{1+P(\boldsymbol{e}_{-M}|N)}}}\right) = -\frac{\log[NP(\boldsymbol{e}_{-M}|N)]}{(1+P(\boldsymbol{e}_{-M}|N))[NP(\boldsymbol{e}_{-M}|N)]^{\frac{P(\boldsymbol{e}_{-M}|N)}{1+P(\boldsymbol{e}_{-M}|N)}}}\frac{\partial P(\boldsymbol{e}_{-M}|N)}{\partial e} \\ > 0.$$

Also, the first term on the right hand side is bounded above by 1. *On the other hand, the second term on the right hand side is decreasing in a and converges to zero as a goes to* 1.[6] *Therefore, there is* $e^*(N) \in (0,1)$ *such that*

$$1 > \frac{N}{N+1}\frac{1+P(\boldsymbol{e}_{-M}|N)}{[NP(\boldsymbol{e}_{-M}|N)]^{\frac{P(\boldsymbol{e}_{-M}|N)}{1+P(\boldsymbol{e}_{-M}|N)}}} + \frac{N-1}{2(N+1)}\frac{1-s(e)}{s(e)}\exp\left(-\frac{g'(e)}{\beta}\right)$$

holds if $e \geq e^*(N)$. *That is,* $\Pi(\mathcal{V}^*|N) - \Pi(\mathcal{V}^*|1)$ *is greater than zero.*

However, for fixed e and N, there is $\underline{U}_M^*$ *such that* $\Pi(\mathcal{V}^*|N) - \Pi(\mathcal{V}^*|1) < 0$ *if* $\underline{U}_M > \underline{U}_M^*(N)$. *Notice that*

$$W_M^G(1) - W_M^G(N) = \exp\left[g(e) + (1-s(e))\frac{g'(e)}{\beta}\right]\cdot \\ \left[\exp\left[-\mathcal{V}^*(1)\right] - \exp\left[-P(\boldsymbol{e}_{-M}|N)\mathcal{V}^*(N)\right]\right],$$

which is less than zero if $\mathcal{V}^*(1) > P(\boldsymbol{e}_{-M}|N)\mathcal{V}^*(N)$. *The difference between these two terms is*

$$\mathcal{V}^*(1) - P(\boldsymbol{e}_{-M}|N)\mathcal{V}^*(N) = \left(\frac{1}{2} - \frac{1}{1+P(\boldsymbol{e}_{-M}|N)}\right)\underline{U}_M + R(N,e), \qquad \text{(A.2)}$$

[6] This result relies on the condition that $\lim_{e\to 1} g'(e) = \infty$.

where $R(N,e)$ is a constant only depending on N and e not $\underline{U}_M$. Since $P(\boldsymbol{e}_{-M}|N)$ is strictly less than 1 when $N \geq 2$, there is $\underline{U}_M^*(N)$ such that $W_M^G(1) < W_M^G(N)$ if $\underline{U}_M > \underline{U}_M^*(N)$. This implies that $\Pi(\mathcal{V}^*|N) - \Pi(\mathcal{V}^*|1) < 0$ when $\underline{U}_M > \underline{U}_M^*(N)$.

Now, I show that $\frac{\partial^2}{\partial(\underline{U}_M)^2}[\Pi(\mathcal{V}^*|N) - \Pi(\mathcal{V}^*|1)] < 0$ if $\frac{\partial}{\partial \underline{U}_M}[\Pi(\mathcal{V}^*|N) - \Pi(\mathcal{V}^*|1)] \leq 0$.

Suppose that

$$\frac{\partial}{\partial \underline{U}_M}[\Pi(\mathcal{V}^*|N) - \Pi(\mathcal{V}^*|1)] = \frac{s(e)}{2}W_M^G(1) - \frac{N}{N+1}s(e)W_M^G(N) - \frac{N-1}{2(N+1)}(1-s(e))W_M^B$$

has a negative value. Then,

$$\begin{aligned}\frac{\partial^2}{\partial(\underline{U}_M)^2}[\Pi(\mathcal{V}^*|N)-\Pi(\mathcal{V}^*|1)] &= \frac{s(e)}{4}W_M^G(1) - \frac{N}{N+1}\frac{s(e)}{1+P(\boldsymbol{e}_{-M}|N)}W_M^G(N) \\ &\quad - \frac{N-1}{2(N+1)}(1-s(a))W_M^B \\ &< \frac{s(e)}{4}W_M^G(1) - \frac{N}{N+1}\frac{s(e)}{2}W_M^G(N) \\ &\quad - \frac{N-1}{4(N+1)}(1-s(a))W_M^B \\ &= \frac{1}{2}\left[\frac{s(e)}{2}W_M^G(1) - \frac{N}{N+1}s(e)W_M^G(N) - \frac{N-1}{2(N+1)}(1-s(e))W_M^B\right] \\ &= \frac{1}{2}\frac{\partial}{\partial \underline{U}_M}[\Pi(\mathcal{V}^*|N) - \Pi(\mathcal{V}^*|1)] \leq 0.\end{aligned}$$

When $\Pi(\mathcal{V}^*|N) - \Pi(\mathcal{V}^*|1) > 0$ as $\underline{U}_M = -\log\left[\frac{s(a)}{s(a)+(1-s(a))\exp\left[-\frac{g'(a)}{\beta}\right]}\right] \equiv \underline{U}_M^0$, there are two possible cases.

1. *$\left(\frac{\partial}{\partial \underline{U}_M}[\Pi(\mathcal{V}^*|N) - \Pi(\mathcal{V}^*|1)] \leq 0\right)$*
   Since there is $\underline{U}_M^*(N)$ such that $\Pi(\mathcal{V}^*|N) - \Pi(\mathcal{V}^*|1) < 0$ when $\underline{U}_M > \underline{U}_M^*(N)$, there is a unique $\widehat{\underline{U}}_M(N)$ such that $\Pi(\mathcal{V}^*|N) - \Pi(\mathcal{V}^*|1) > 0$ if $\underline{U}_M \in \left[\underline{U}_M^0, \widehat{\underline{U}}_M(N)\right)$.

2. *$\left(\frac{\partial}{\partial \underline{U}_M}[\Pi(\mathcal{V}^*|N) - \Pi(\mathcal{V}^*|1)] > 0\right)$.*
   The condition that $\frac{\partial^2}{\partial(\underline{U}_M)^2}[\Pi(\mathcal{V}^*|N) - \Pi(\mathcal{V}^*|1)] < 0$ if $\frac{\partial}{\partial \underline{U}_M}[\Pi(\mathcal{V}^*|N) - \Pi(\mathcal{V}^*|1)] = 0$ implies that there is a unique $\underline{U}_M$ such that $\frac{\partial}{\partial \underline{U}_M}[\Pi(\mathcal{V}^*|N) - \Pi(\mathcal{V}^*|1)] = 0$. Hence, there is a unique $\widehat{\underline{U}}_M(N)$ such that $\Pi(\mathcal{V}^*|N) - \Pi(\mathcal{V}^*|1) > 0$ if $\underline{U}_M \in \left[\underline{U}_M^0, \widehat{\underline{U}}_M(N)\right)$.

Lastly, I show that there is N^* *such that* $W_M^G(1) < W_M^G(N)$ *if* $N \geq N^*$ *for a given* $(\underline{U}_M, e)$*. This implies that* $\Pi(\mathcal{V}^*|N) < \Pi(\mathcal{V}^*|1) > 0$*. In the equation (A.2),* $R(N, e)$ *is equal to*

$$R(N,e) = \frac{1}{2}\log\left(\frac{s(e)}{s(e) + (1-s(e))\exp\left[-\frac{g'(a)}{\beta}\right]}\right) - \frac{P(\boldsymbol{e}_{-M}|N)}{1+P(\boldsymbol{e}_{-M}|N)}\log\left(\frac{1-(1-s(e))^N}{s(e) + (1-s(e))\exp\left[-\frac{g'(e)}{\beta}\right]}\right).$$

Since $\left(\frac{1}{2} - \frac{1}{1+P(\boldsymbol{e}_{-M}|N)}\right)\underline{U}_M > 0$*, it is enough to show that* $R(N, e) > 0$ *if* $N \geq N^*$*. The first term of* $R(N, e)$ *does not depend on* N *and has a strictly positive number. Denote this number by* C*. On the other hand, the second term is always less than*

$$\frac{P(\boldsymbol{e}_{-M}|N)}{1+P(\boldsymbol{e}_{-M}|N)}\log\left(\frac{1}{s(e) + (1-s(e))\exp\left[-\frac{g'(a)}{\beta}\right]}\right), \tag{A.3}$$

which is strictly decreasing function in N *and converges to zero. Hence, there is* N^* *such that (A.3) is less than* C *if* $N \geq N^*$*. This implies that* $R(N, e) > 0$*.*

A.1.19 Proof of Proposition 10

Note that

$$\begin{aligned}
\frac{\partial^2 F}{\partial \mathcal{V} \partial N} &= -(1-s(e_M))^N \log(1-s(e_M))\frac{1}{u'(W_g)} > 0, \\
\frac{\partial^2 F}{\partial(-e_C)\partial N} &= 0, \text{ and} \\
\frac{\partial^2 F}{\partial \mathcal{V} \partial(-e_C)} &= s(e_C)\frac{\partial^2 W_C^G}{\partial \mathcal{V} \partial e_C} + (1-s(e_C))\frac{\partial^2 W_C^B}{\partial \mathcal{V} \partial e_C} + \beta\left(\frac{\partial W_C^G}{\partial \mathcal{V}} - \frac{\partial W_C^B}{\partial \mathcal{V}}\right) \\
&= -s(e_C)(1-s(e_C))\frac{g''(e_C)}{\beta}\left(\frac{u''(W_C^G)}{u'(W_C^G)^3} - \frac{u''(W_C^B)}{u'(W_C^B)^3}\right) \\
&\quad + \beta\left(\frac{1}{u'(W_C^G)} - \frac{1}{u'(W_C^B)}\right).
\end{aligned}$$

This result shows that $\frac{\partial \mathcal{V}^*}{\partial N} \geq 0$ and $\frac{\partial e_C^*}{\partial N} \leq 0$ according to Milgrom and Shannon (1994). First, consider e_C^* as a function of $\mathcal{V}$. Then, under the condition that $u''(x)/u'(x)^3$ is a decreasing function in x,

$$\begin{aligned}\frac{\partial^2 F}{\partial \mathcal{V} \partial N} &= \left[-\beta\left(\frac{1}{u'(W_C^G)} - \frac{1}{u'(W_C^B)}\right) + s(e_C^*)\frac{u''(W_C^G)}{u'(W_C^G)^2}\frac{\partial W_C^G}{\partial e_C^*}\right. \\ &\qquad \left. +(1 - s(e_C^*))\frac{u''(W_C^B)}{u'(W_C^B)^2}\frac{\partial W_C^B}{\partial e_C^*}\right]\frac{\partial e_C^*}{\partial N} \\ &\qquad - (1 - s(e_M))^N \log(1 - s(e_M))\frac{1}{u'(W_g)} \\ &> 0\end{aligned}$$

since $\frac{\partial e_C^*}{\partial N} \leq 0$. This implies that $\frac{\partial \mathcal{V}^*}{\partial N} > 0$. Now, consider the first order condition with respect to e_C:

$$\begin{aligned}\beta(\mathcal{G} - \mathcal{B}) &= \beta(W_C^G - W_C^B) + s(e_C^*)\frac{\partial W_C^G}{\partial e_C} + (1 - s(e_C))\frac{\partial W_C^B}{\partial e_C} \\ &= \beta(W_C^G - W_C^B) + s(e_C^*)(1 - s(e_C^*))\frac{g''(e_C^*)}{\beta}\left(\frac{1}{u'(W_C^G)} - \frac{1}{u'(W_C^B)}\right).\end{aligned}$$

The right hand side of the equation is a strictly increasing function in e_C^* if $g'''(e_C) \geq 0$ and a strictly decreasing function in $\mathcal{V}$. Hence, $\frac{\partial \mathcal{V}^*}{\partial N} > 0$ indicates that $\frac{\partial e_C^*}{\partial N} < 0$.

A.1.20 Proof of Corollary 7

Note that

$$\begin{aligned}\frac{\partial}{\partial N}[(W_C^G)^* - (W_C^B)^*] &= \frac{\partial \mathcal{V}}{\partial N}\left[\frac{1}{u'((W_C^G)^*)} - \frac{1}{u'((W_C^B)^*)} + \left(\frac{1 - s(e_C^*)}{u'((W_C^G)^*} + \frac{s(e_C^*)}{u'((W_C^B)^*)}\right)\frac{g''(e_C^*)}{\beta}\frac{\partial e_C^*}{\partial \mathcal{V}}\right] \\ &= \frac{\partial \mathcal{V}}{\partial N}\left[(W_C^G)^* - (W_C^B)^* + \left((1 - s(e_C^*))(W_C^G)^* + s(e_C^*)(W_C^B)^*\right)\frac{g''(e_C^*)}{\beta}\frac{\partial e_C^*}{\partial \mathcal{V}}\right]\end{aligned}$$

when agents have the log utility function. Also, it can be shown that

$$\frac{\partial e_C^*}{\partial \mathcal{V}} = -\frac{\left(\beta + s(e_C^*)(1 - s(e_C^*))\frac{g''(e_C^*)}{\beta}\right)((W_C^G)^* - (W_C^B)^*)}{D_1 + D_2},$$

where

$$\begin{aligned}D_1 &= \left(\beta(1 - 2s(e_C^*))\frac{g''(e_C^*)}{\beta} + s(e_C^*)(1 - s(e_C^*))\frac{g'''(e_C^*)}{\beta}\right)((W_C^G)^* - (W_C^B)^*) \\ D_2 &= \left(\beta + s(e_C^*)(1 - s(e_C^*))\frac{g''(e_C^*)}{\beta}\right)\frac{g''(e_C^*)}{\beta}((1 - s(e_C^*))(W_C^G)^* + s(e_C^*)(W_C^B)^*).\end{aligned}$$

Hence,

$$\frac{\partial}{\partial N}[(W_C^G)^* - (W_C^B)^*] = \frac{(W_C^G)^* - (W_C^B)^*}{D_1 + D_2}D_1 > 0.$$

A.1.21 Proof of Proposition 11

Note that when $\mathcal{V} = 0$, $e_C^* = e_M^*$ by two first order conditions. Also, this implies that $W_C^G = W_M^G$. Therefore,

$$\begin{aligned}\frac{\partial F(\mathcal{V})}{\partial \mathcal{V}}\bigg|_{\mathcal{V}=0} &= -\frac{s(e_C^*)}{u'(W_C^G)} - \frac{1-s(e_C^*)}{u'(W_C^B)} + \frac{1-(1-s(e_M^*))^N}{u'(W_M^G)} \\ &= (1-s(e_C^*))\left(\frac{1-(1-s(e_C^*))^{N-1}}{u'(W_C^G)} - \frac{1}{u'(W_C^B)}\right).\end{aligned}$$

Since $W_C^G > W_C^B$ and $(1-s(e_C^*))^{N-1} \to 0$ as $N \to \infty$, there is $\widehat{N}$ such that $\frac{\partial F(\mathcal{V})}{\partial \mathcal{V}}\big|_{\mathcal{V}=0} > 0$ if $N > \widehat{N}$. This implies that $\mathcal{V}^* > \underline{U}_M = 0$ when $N > \widehat{N}$. As a next step, I show that the promotion incentive is bounded regardless of the firm size.

Claim 5 *The optimal promotion incentive $\mathcal{V}^*$ is bounded for any N.*

Proof 8 *First, fix $(e_M) \in (0,1)$ for a given N. Denote the optimal $\mathcal{V}$ by $\mathcal{V}^*(e_C)$ for a given e_C. Recall that $\frac{\partial \mathcal{V}^*(e_C)}{\partial e_C} < 0$. Therefore,*

$$\mathcal{V}^*(e_C) \leq \mathcal{V}^*(0).$$

When $e_C = 0$, the first order condition with respect to $\mathcal{V}$ is

$$-\frac{1}{u'(W_C^F)} + \frac{1-(1-s(e))^N}{u'(W_M^G)} = 0,$$

where $u(W_C^F) = \mathcal{V}^*(0)$. It can be easily shown that

$$\mathcal{V}^*(0) = u(W_C^F) < u(W_M^G) = \underline{U}_M + g(e_M) + (1-s(e_M))\frac{g'(e_M)}{\beta} - P(\boldsymbol{e}_{-M})\mathcal{V}^*(0),$$

implying that

$$\mathcal{V}^*(0) < \frac{1}{1+P(\boldsymbol{e}_{-M})}\left[\underline{U}_M + g(e_M) + (1-s(e_M))\frac{g'(e_M)}{\beta}\right].$$

Notice that this bound does not depend on e_C. Now, suppose that

$$\mathcal{V} = \frac{1}{1+P(\boldsymbol{e}_{-M})}\left[\underline{U}_M + g(e_M) + (1-s(e_M))\frac{g'(e_M)}{\beta}\right].$$

Then, the manager's wage for good performance satisfies

$$\begin{aligned}u(W_M^G) &= \frac{1}{1+P(\boldsymbol{e}_{-M})}\left[\underline{U}_M + g(e_M) + (1-s(e_M))\frac{g'(e_M)}{\beta}\right] \\ &\geq \frac{1}{2}\left[\underline{U}_M + g(e_M) + (1-s(e_M))\frac{g'(e_M)}{\beta}\right].\end{aligned}$$

The bound for $u(W_M^G)$ does not depend on N. This result means that the firm has to pay the wage satisfying the lower bound if it requires an effort level e_M form its managers. Since the wage approaches infinity as e_M converges to one, there is $\bar{e}_M < 1$ such that $e_C^* \leq \bar{e}_M$ regardless of N and e_C. Hence, $\mathcal{V}^*$ is bounded by $\frac{1}{1+P(\bar{\boldsymbol{e}}_{-M})}\left[\underline{U}_M + g(\bar{e}_M) + (1-s(\bar{e}_M))\frac{g'(\bar{e}_M)}{\beta}\right]$, which is less than $\left[\underline{U}_M + g(\bar{e}_M) + (1-s(\bar{e}_M))\frac{g'(\bar{e}_M)}{\beta}\right]$. I denote this bound by $\bar{\mathcal{V}}$. This upper bound does not depend on N.

Based on this result, I show the following result.

Claim 6 *There is $\bar{N}$ such that $(W_M^G)^* > (W_M^B)^*$ if $N > \bar{N}$.*

Proof 9 *Recall that*

$$(W_M^G)^* \leq (W_M^B)^*$$

if and only if

$$\frac{g'(e_M^*)}{\beta} \leq P(\boldsymbol{e}_{-M})\mathcal{V}^*.$$

Since $e_M^* \geq \underline{e}_M$, where $\underline{e}_M$ is the firm's optimal effort choice when $\mathcal{V} = 0$, and $P(\boldsymbol{e}_{-M}) \to 0$ as $N \to \infty$, there exists $\bar{N}$ such that

$$\frac{g'(e_M^*)}{\beta} \geq \frac{g'(\underline{e}_M)}{\beta} > P(\underline{\boldsymbol{e}}_{-M})\bar{\mathcal{V}} \geq P(\boldsymbol{e}_{-M})\bar{\mathcal{V}}$$

if $N > \bar{N}$. Therefore, $(W_M^G)* > (W_M^B)*$ if $N > \bar{N}$.

Now, I show that $e_M^* > e_C^*$ if $N > N^* \equiv \max\{\widehat{N}, \bar{N}\}$.

Note that two optimal effort levels e_C^* and e_M^* are decided by two first order conditions for a given $\mathcal{V}$:

$$\begin{aligned}\beta(\mathcal{G} - \mathcal{B}) &= \beta(W_C^G - W_C^B) + s(e_C^*)\frac{\partial W_C^G}{\partial e_C} + (1 - s(e_C))\frac{\partial W_C^B}{\partial e_C} \\ &= \beta(W_C^G - W_C^B) + s(e_C^*)(1 - s(e_C^*))\frac{g''(e_C^*)}{\beta}\left(\frac{1}{u'(W_C^G)} - \frac{1}{u'(W_C^B)}\right), \text{ and}\end{aligned} \tag{A.4}$$

$$\begin{aligned}\beta(\mathcal{G} - \mathcal{B}) &= \beta(W_M^G - W_M^B) + s(e_M^*)\frac{\partial W_M^G}{\partial e_M} + (1 - s(e_M^*))\frac{\partial W_M^B}{\partial e_M} \\ &= \beta(W_M^G - W_M^B) + s(e_M^*)(1 - s(e_M^*))\frac{g''(e_M^*)}{\beta}\left(\frac{1}{u'(W_M^G)} - \frac{1}{u'(W_M^B)}\right) \\ &\quad - s(e_M^*)\frac{\partial P(\mathbf{e}_{-M}^*)}{\partial e_M}\frac{1}{u'(W_M^G)}\mathcal{V}.\end{aligned} \tag{A.5}$$

If $\mathcal{V}$ is equal to $\underline{U}_M = 0$, two conditions yield $e_C^* = e_M^*$. The right hand side of (A.4) is a strictly increasing function in $\mathcal{V}$ while that of (A.5) is a strictly decreasing function in $\mathcal{V}$ since

$$\frac{\partial J_C(\mathcal{V},e_C)}{\partial \mathcal{V}} = \beta\left(\frac{1}{u'(W_C^G)} - \frac{1}{u'(W_C^B)}\right) - s(e_C)(1-s(e_C))\frac{g''(e_C)}{\beta}\left(\frac{u''(W_C^G)}{u'(W_C^G)^3} - \frac{u''(W_C^B)}{u'(W_C^B)^3}\right) > 0,$$

$$\begin{aligned}\frac{\partial J_M(\mathcal{V},e_M)}{\partial \mathcal{V}} &= -\beta P(\mathbf{e}_{-M})\frac{1}{u'(W_M^G)} + s(e_M)(1-s(e_M))\frac{g''(e_M)}{\beta}P(\mathbf{e}_{-M})\frac{u''(W_M^G)}{u'(W_M^G)^3} \\ &\quad - s(e_M)\frac{\partial P(\mathbf{e}_{-M})}{\partial e_M}\frac{1}{u'(W_M^G)} - s(e_M)\frac{\partial P(\mathbf{e}_{-M})}{\partial e_M}P(\mathbf{e}_{-M})\frac{u''(W_M^G)}{u'(W_M^G)^3}\mathcal{V} \\ &= s(e_M)(1-s(e_M))\frac{g''(e_M)}{\beta}P(\mathbf{e}_{-M})\frac{u''(W_M^G)}{u'(W_M^G)^3} - \beta(1-s(e_M))^{N-1}\frac{1}{u'(W_M^G)} \\ &\quad - s(e_M)\frac{\partial P(\mathbf{e}_{-M})}{\partial e_M}P(\mathbf{e}_{-M})\frac{u''(W_M^G)}{u'(W_M^G)^3}\mathcal{V} < 0,\end{aligned}$$

where

$$J_C(\mathcal{V},e_C) = \beta(W_C^G - W_C^B) + s(e_C)(1-s(e_C))\frac{g''(e_C)}{\beta}\left(\frac{1}{u'(W_C^G)} - \frac{1}{u'(W_C^B)}\right), \text{ and}$$

$$\begin{aligned}J_M(\mathcal{V},e_M) &= \beta(W_M^G - W_M^B) + s(e_M)(1-s(e_M))\frac{g''(e_M)}{\beta}\left(\frac{1}{u'(W_M^G)} - \frac{1}{u'(W_M^B)}\right) \\ &\quad - s(e_M)\frac{\partial P(\mathbf{e}_{-M})}{\partial e_M}\frac{1}{u'(W_M^G)}\mathcal{V}, \text{and}\end{aligned}$$

$$\frac{\partial P(\mathbf{e}_{-M})}{\partial e_M} = \frac{\beta}{s(e_M)}[(1-s(e_M))^{N-1} - P(\mathbf{e}_{-M})] < 0.$$

In addition, the following results

$$\begin{aligned}\frac{\partial J_C(\mathcal{V},e_C)}{\partial e_C} &= 2g''(e_C)\left(\frac{1-s(e_C)}{u'(W_C^G)} + \frac{s(e_C)}{u'(W_C^B)}\right) - g''(e_C)\left(\frac{s(e_C)}{u'(W_C^G)} + \frac{1-s(e_C)}{u'(W_C^B)}\right) \\ &\quad - s(e_C)(1-s(e_C))\frac{g''(e_C)^2}{\beta^2}\left[(1-s(e_C))\frac{u''(W_C^G)}{u'(W_C^B)^3} + s(e_C)\frac{u''(W_C^B)}{u'(W_C^B)^3}\right] \\ &\quad + s(e_C)(1-s(e_C))\frac{g'''(e_C)}{\beta}\left(\frac{1}{u'(W_C^G)} - \frac{1}{u'(W_C^B)}\right) > 0,\end{aligned}$$

$$\begin{aligned}\frac{\partial J_M(\mathcal{V},e_M)}{\partial e_M} &= \left.\frac{\partial J_C(\mathcal{V},e_C)}{\partial e_C}\right|_{e_C=e_M} - 2\beta\frac{\partial P(\mathbf{e}_{-M})}{\partial e_M}\frac{1}{u'(W_M^G)}\mathcal{V} \\ &\quad + \left[2s(e_M)(1-s(e_M))\frac{g''(e_M)}{\beta}\frac{\partial P(\mathbf{e}_{-M})}{\partial e_M}\mathcal{V} - s(e_M)\left(\frac{\partial P(\mathbf{e}_{-M})}{\partial e_M}\right)^2\mathcal{V}^2\right]\frac{u''(W_M^G)}{u'(W_M^G)^3} \\ &\quad - s(e_M)\frac{\partial^2 P(\mathbf{e}_{-M})}{\partial (e_M)^2}\frac{1}{u'(W_M^G)}\mathcal{V} \\ &= \left.\frac{\partial J_C(\mathcal{V},e_C)}{\partial e_C}\right|_{e_C=e_M} + \beta^2(N-1)(1-s(e_M))^{N-2}\frac{1}{u'(W_M^G)}\mathcal{V} \\ &\quad + \left[2s(e_M)(1-s(e_M))\frac{g''(e_M)}{\beta}\frac{\partial P(\mathbf{e}_{-M})}{\partial e_M}\mathcal{V} - s(e_M)\left(\frac{\partial P(\mathbf{e}_{-M})}{\partial e_M}\right)^2\mathcal{V}^2\right]\frac{u''(W_M^G)}{u'(W_M^G)^3} \\ &> 0,\end{aligned}$$

where

$$\frac{\partial^2 P(\mathbf{e}_{-M})}{\partial (e_M)^2} = \frac{\beta^2}{Ns(e_M)^3}\Big[-N(N-1)s(e_M)^2(1-s(e_M))^{N-2}$$
$$+2-2(1-s(e_M))^{N-1}-2(N-1)s(e_M)(1-s(e_M))^{N-1}\Big]$$
$$= -\frac{\beta^2}{s(e_M)}(N-1)(1-s(e_M))^{N-2} - 2\frac{\beta}{s(e_M)}\frac{\partial P(\mathbf{e}_{-M})}{\partial e_M} \geq 0,$$

validate the first order approach.

Hence, $e_C^* < e_M^*$ since $\mathcal{V}^* > \underline{U}_M = 0$ when $N > \widehat{N}$.

A.1.22 Proof of Proposition 12

For a given $(\mathcal{V}, N)$, U_{M2}^* is determined by the equation (2.9). I denote this by $U_{M2}^*(\mathcal{V}, N)$ to explicitly express the dependency. The first order condition with respect to $\mathcal{V}$ and the equation (2.9) imply that $\mathcal{V}^*$ and $U_{M2}^*(\mathcal{V}^*, N)$ satisfy

$$\frac{s(e_C)}{u'((W_C^G)^*)} + \frac{1-s(e_C)}{u'((W_C^B)^*)} + \frac{(1-s(e_{M1}))^N}{u'((W_M^G)^*)} = \delta\left[\frac{s(e_{M2})}{u'((W_M^{GG})^*)} + \frac{1-s(e_{M2})}{u'((W_M^{GB})^*)}\right],$$

where the first order condition with respect to $\mathcal{V}$ is

$$\frac{\partial F(\mathcal{V})}{\partial \mathcal{V}} = -\frac{s(e_C)}{u'(W_C^G)} - \frac{1-s(e_C)}{u'(W_C^B)} + Ns(e_{M1})\cdot$$
$$\left[\frac{1}{u'(W_M^G)}\left(P(\mathbf{e}_{-M}) + (1-P(\mathbf{e}_{-M}))\frac{\partial U_{M2}(\mathcal{V})}{\partial \mathcal{V}}\right)\right]$$
$$-\delta\left(Ns(e_{M1}) - 1 + (1-s(e_{M1}))^N\right)\left(\frac{s(e_{M2})}{u'(W_M^{GG})} + \frac{1-s(e_{M2})}{u'(W_M^{GB})}\right)\frac{\partial U_{M2}(\mathcal{V})}{\partial \mathcal{V}}$$
$$= -\frac{s(e_C)}{u'(W_C^G)} - \frac{1-s(e_C)}{u'(W_C^B)} + \frac{1-(1-s(e_{M1}))^N}{u'(W_M^G)}$$

by (2.9).

Note that $\frac{(1-s(e_{M1}))^N}{u'(W_M^G)}$ approaches zero as N goes to infinity. Also, when $e_C = e_{M2}$, $\mathcal{V} = U_{M2}$, and $\delta = 1$, the condition is equal to

$$\frac{s(e_C)}{u'(W_C^G)} + \frac{1-s(e_C)}{u'(W_C^B)} = \frac{s(e_{M2})}{u'(W_M^{GG})} + \frac{1-s(e_{M2})}{u'(W_M^{GB})}.$$

Since the left hand side of this equation is a strictly increasing function in e_C and $\mathcal{V}$, there is $\widehat{N}$, $\hat{e}_C < e_{M2}$ such that

$$\frac{s(\hat{e}_C)}{u'(W_C^G)} + \frac{1-s(\hat{e}_C)}{u'(W_C^B)} + \frac{(1-s(e_{M1}))^{\widehat{N}}}{u'(W_M^G)} < \delta\left[\frac{s(e_{M2})}{u'(W_M^{GG})} + \frac{1-s(e_{M2})}{u'(W_M^{GB})}\right]$$

when $\mathcal{V} = \underline{U}_M$ for a sufficiently large δ. This implies that for given $(\delta, e_C, e_{M1}, e_{M2}, N)$, where $e_C \leq \hat{e}_C < e_{M2}$ and $N \geq \widehat{N}$, $\mathcal{V}^* > U_{M2}^*(\mathcal{V}^*, N)$.

Now, I show that $\mathcal{V}^*$ is an increasing function in $N > \widehat{N}$ when $\mathcal{V}^* \geq U^*_{M2}(\mathcal{V}^*, \widehat{N})$.

First, note that for a given $\mathcal{V}$ and N,

$$\frac{\partial U^*_{M2}(\mathcal{V},N)}{\partial N} = -\frac{\frac{u''(W_M^G)}{u'(W_M^G)^3}\frac{\partial P(\mathbf{e}_{-M1})}{\partial N}(\mathcal{V} - U^*_{M2}(\mathcal{V},N))}{\frac{u''(W_M^G)}{u'(W_M^G)^3}(1 - P(\mathbf{e}_{-M1})) + \delta\left[s(e_{M2})\frac{u''(W_M^{GG})}{u'(W_M^{GG})^3} + (1 - s(e_{M2}))\frac{u''(W_M^{GB})}{u'(W_M^{GB})^3}\right]}.$$

Therefore,

$$\begin{aligned}
\frac{\partial^2 F(\mathcal{V})}{\partial N \partial \mathcal{V}} &= -(1 - s(e_{M1}))^N \log(1 - s(e_{M1}))\frac{1}{u'(W_M^G)} \\
&\quad + [1 - (1 - s(e_{M1}))^N]\frac{u''(W_M^G)}{u'(W_M^G)^3}\cdot \\
&\quad \left[\frac{\partial P(\mathbf{e}_{-M1})}{\partial N}(\mathcal{V}^* - U^*_{M2}) + (1 - P(\mathbf{e}_{-M1}))\frac{\partial U^*_{M2}(\mathcal{V},N)}{\partial N}\right] \\
&= -(1 - s(e_{M1}))^N \log(1 - s(e_{M1}))\frac{1}{u'(W_M^G)} \\
&\quad - [1 - (1 - s(e_{M1}))^N]\left[s(e_{M2})\frac{u''(W_M^{GG})}{u'(W_M^{GG})^3} + (1 - s(e_{M2}))\frac{u''(W_M^{GB})}{u'(W_M^{GB})^3}\right] > 0.
\end{aligned}$$

Moreover,

$$\begin{aligned}
\frac{\partial^2 F(\mathcal{V})}{\partial \mathcal{V}^2} &= s(e_C)\frac{u''(W_C^G)}{u'(W_C^G)^3} + (1 - s(e_C))\frac{u''(W_C^B)}{u'(W_C^B)^3} \\
&\quad + (1 - (1 - s(e_{M1}))^N)\frac{u''(W_M^G)}{u'(W_M^G)^3}\left[P(\mathbf{e}_{-M1}) + (1 - P(\mathbf{e}_{-M1}))\frac{\partial U^*_{M2}(\mathcal{V},N)}{\partial \mathcal{V}}\right] \\
&= s(e_C)\frac{u''(W_C^G)}{u'(W_C^G)^3} + (1 - s(e_C))\frac{u''(W_C^B)}{u'(W_C^B)^3} \\
&\quad + (1 - (1 - s(e_{M1}))^N)\frac{u''(W_M^G)}{u'(W_M^G)^3} \\
&\quad \cdot\left[P(\mathbf{e}_{-M1})\frac{\delta\left[s(e_{M2})\frac{u''(W_M^{GG})}{u'(W_M^{GG})^3} + (1 - s(e_{M2}))\frac{u''(W_M^{GB})}{u'(W_M^{GB})^3}\right]}{(1 - P(\mathbf{e}_{-M1}))\frac{u''(W_M^G)}{u'(W_M^G)^3} + \delta\left[s(e_{M2})\frac{u''(W_M^{GG})}{u'(W_M^{GG})^3} + (1 - s(e_{M2}))\frac{u''(W_M^{GB})}{u'(W_M^{GB})^3}\right]}\right] \\
&< 0,
\end{aligned}$$

where I exploit

$$\frac{\partial U^*_{M2}(\mathcal{V},N)}{\partial \mathcal{V}} = -\frac{P(\mathbf{e}_{-M1})\frac{u''(W_M^G)}{u'(W_M^G)^3}}{(1 - P(\mathbf{e}_{-M1}))\frac{u''(W_M^G)}{u'(W_M^G)^3} + \delta\left[s(e_{M2})\frac{u''(W_M^{GG})}{u'(W_M^{GG})^3} + (1 - s(e_{M2}))\frac{u''(W_M^{GB})}{u'(W_M^{GB})^3}\right]}$$

using the implicit function theorem.

A.1.23 Derivation of the Firm's Problem in Section 2.7.2

The firm's problem can be written as

$$\max_{C,M_1,M_2} E_0\left[\sum_{t=1}^{\infty}\delta^{t-1}\mathcal{P}_t(C,M_1,M_2|\mathcal{H}_{t-1})\right]$$
$$\text{subject to } (IR_C),\ (IC_C),\ (IR_{M1}),\ (IC_{M1}),\ (IR_{M2}),\ (IC_{M2}),$$

where

$$\mathcal{P}_1(C,M_1,M_2|\mathcal{H}_0) = \mathcal{P}_C(C) + \mathcal{P}_M(M_1)$$
$$\mathcal{P}_t(C,M_1,M_2|\mathcal{H}_{t-1}) = \begin{cases}\mathcal{P}_C(C) + N\mathcal{P}_M(M_1) & \text{if } \mathcal{H}_{t-1}\in\mathcal{S}_1\\ N\mathcal{P}_M(M_2) & \text{if } \mathcal{H}_{t-1}\in\mathcal{S}_2\end{cases}$$

with

$$\begin{aligned}\mathcal{P}_C(C) &= s(e_{C1})(\mathcal{G}_C - W_C^G) + (1-s(e_{C1}))(\mathcal{B}_C - W_C^B)\\ &\quad + \delta s(e_{C1})[s(e_{C2})(\mathcal{G}_C - W_C^{GG}) + (1-s(e_{C2}))(\mathcal{B}_C - W_C^{GB})]\\ \mathcal{P}_M(M_i) &= s(e_{Mi})(\mathcal{G}_M - W_{Mi}^G) + (1-s(e_{Mi}))(\mathcal{B}_M - W_{Mi}^B)\\ C &= (W_C^G, W_C^B, W_C^{GG}, W_C^{GB})\\ M_i &= (W_{Mi}^G, W_{Mi}^B).\end{aligned}$$

Also, $\mathcal{H}_t$ denotes the CEO's seniority and outcome at time t. Hence,

$$\mathcal{H}_t \in \{(C_1,\mathcal{G}_C),\ (C_1,\mathcal{B}_C),\ (C_2,\mathcal{G}_C),\ (C_2,\mathcal{B}_C)\},$$

where C_i is equal to C_1 (C_2) if the CEO is her first period (second period) in the position. For brevity, I use two terms, $\mathcal{S}_1$ and $\mathcal{S}_2$, in order to represent

$$\begin{aligned}\mathcal{S}_1 &= \{(C_1,\mathcal{B}_C),\ (C_2,\mathcal{G}_C),\ (C_2,\mathcal{B}_C)\}\\ \mathcal{S}_2 &= \{(C_1,\mathcal{G}_C)\},\end{aligned}$$

respectively.

Then,

$$\begin{aligned}E_0\left[\sum_{t=1}^{\infty}\delta^{t-1}\mathcal{P}_t(C,M_1,M_2|\mathcal{H}_{t-1})\right] &= \mathcal{P}_C(C) + N\mathcal{P}_M(M_1)\\ &\quad + \delta s(e_{C1})\left\{N\mathcal{P}_M(M_2) + \delta E_0\left[\sum_{t=1}^{\infty}\delta^{t-1}\mathcal{P}_t(C,M_1,M_2|\mathcal{H}_{t-1})\right]\right\}\\ &\quad + \delta(1-s(e_{C1}))E_0\left[\sum_{t=1}^{\infty}\delta^{t-1}\mathcal{P}_t(C,M_1,M_2|\mathcal{H}_{t-1})\right],\end{aligned}$$

where I exploit

$$E_0\left[\sum_{t=1}^{\infty}\delta^{t-1}\mathcal{P}_t(C,M_1,M_2|\mathcal{H}_{t-1})\right] = E_s\left[\sum_{t=s+1}^{\infty}\delta^{t-(s+1)}\mathcal{P}_t(C,M_1,M_2|\mathcal{H}_{t-1})\right]$$

if $\mathcal{H}_s \in \mathcal{S}_1$. Therefore,

$$E_0\left[\sum_{t=1}^{\infty}\delta^{t-1}\mathcal{P}_t(C,M_1,M_2|\mathcal{H}_{t-1})\right] = \frac{1}{(1-\delta)(1+\delta s(e_{C1}))}\left[\mathcal{P}_C(C) + N\mathcal{P}_M(M_1) + \delta s(e_{C1})N\mathcal{P}_M(M_2)\right].$$

Since I treat e_{C1} as an exogenous variable, the firm's problem is to choose (C, M_1, M_2) maximizing

$$\mathcal{P}_C(C) + N\mathcal{P}_M(M_1) + \delta s(e_{C1})N\mathcal{P}_M(M_2).$$

A.1.24 Proof of Proposition 13

The firm's problem for guaranteed situation is to choose $\widehat{\mathcal{V}} \in [0,\infty)$ maximizing $\widehat{F}(\widehat{\mathcal{V}})$ defined by[7]

$$\begin{aligned}
\widehat{F}(\widehat{\mathcal{V}}) \equiv \max_{\widehat{\mathcal{A}}}\ & s(e_C)(\mathcal{G}_C - \widehat{W}_C^G) + (1-s(e_C))(\mathcal{B}_C - \widehat{W}_C^B) \\
& + \delta s(e_C)[s(e_C)(\mathcal{G}_C - \widehat{W}_C^{GG}) + (1-s(e_C))(\mathcal{B}_C - \widehat{W}_C^{GB})] \\
& + \delta(1-s(e_C))[s(e_C)(\mathcal{G}_C - \widehat{W}_C^{BG}) + (1-s(e_C))(\mathcal{B}_C - \widehat{W}_C^{BB})] \\
& + N\left[s(e_M)(\mathcal{G}_M - \widehat{W}_{M1}^G) + (1-s(e_M))(\mathcal{B}_M - \widehat{W}_{M1}^B)\right] \\
& + \delta N\left[s(e_M)(\mathcal{G}_M - \widehat{W}_{M2}^G) + (1-s(e_M))(\mathcal{B}_M - \widehat{W}_{M2}^B)\right]
\end{aligned}$$

subject to

$$\begin{aligned}
u(\widehat{W}_C^G) &= \mathcal{V} + g(e_C) + (1-s(e_C))\frac{g'(e_C)}{\beta} - V_2^G, \\
u(\widehat{W}_C^B) &= \mathcal{V} + g(e_C) - s(e_C)\frac{g'(e_C)}{\beta} - V_2^B, \\
u(\widehat{W}_C^{GG}) &= V_2^G + g(e_C) + (1-s(e_C))\frac{g'(e_C)}{\beta}, \\
u(\widehat{W}_C^{GB}) &= V_2^G + g(e_C) - s(e_C)\frac{g'(e_C)}{\beta}, \\
u(\widehat{W}_C^{BG}) &= V_2^B + g(e_C) + (1-s(e_C))\frac{g'(e_C)}{\beta}, \\
u(\widehat{W}_C^{BB}) &= V_2^B + g(e_C) - s(e_C)\frac{g'(e_C)}{\beta}, \\
u(\widehat{W}_{M1}^G) &= \underline{U}_M + g(e_M) + (1-s(e_M))\frac{g'(e_M)}{\beta}, \\
u(\widehat{W}_{M1}^B) &= \underline{U}_M + g(e_M) - s(e_M)\frac{g'(e_M)}{\beta}, \\
u(\widehat{W}_{M2}^G) &= \underline{U}_M + g(e_M) + (1-s((e_M))\frac{g'((e_M)}{\beta} - P(\mathbf{e}_{-M})\mathcal{V}, \text{ and} \\
u(\widehat{W}_{M2}^B) &= \underline{U}_M + g(e_M) - s(e_M)\frac{g'(e_M)}{\beta},
\end{aligned}$$

[7] I use the hat notation to indicate guaranteed job security case.

where

$$
\begin{aligned}
\widehat{\mathcal{A}} &= \{(\widehat{W}_C^G, \widehat{W}_C^B, \widehat{W}_C^{GG}, \widehat{W}_C^{GB}, \widehat{W}_C^{BG}, \widehat{W}_C^{BB}), (\widehat{W}_{M1}^G, \widehat{W}_{M2}^B), (\widehat{W}_{M2}^G, \widehat{W}_{M2}^B)\}, \\
V_2^G &= s(e_C)u(\widehat{W}_C^{GG}) + (1 - s(e_C))u(\widehat{W}_C^{GB}) - g(e_C), \text{ and} \\
V_2^B &= s(e_C)u(\widehat{W}_C^{BG}) + (1 - s(e_C))u(\widehat{W}_C^{BB}) - g(e_C).
\end{aligned}
$$

Then, the first order condition with respect to $\mathcal{V}$ for unguaranteed situation is

$$
-\frac{s(e_C)}{u'(W_C^G)} - \frac{1 - s(e_C)}{u'(W_C^B)} + \delta(1 - s(e_C))\frac{1 - (1 - s(e_M))^N}{u'(W_{M1}^G)} + \delta s(e_C)\frac{1 - (1 - s(e_M))^N}{u'(W_{M2}^G)} = 0.
$$

On the other hand, the condition for guaranteed case is

$$
-\frac{s(e_C)}{u'(\widehat{W}_C^G)} - \frac{1 - s(e_C)}{u'(\widehat{W}_C^B)} + \delta\frac{1 - (1 - s(e_M))^N}{u'(\widehat{W}_{M2}^G)} = 0.
$$

First, I show that there is $\delta^* \in (0,1)$ such that $(V_2^B)^* < \widehat{\mathcal{V}}$ for a given $\widehat{\mathcal{V}} \in [0, \infty)$. Note that $\widehat{\mathcal{V}}$ and $(V_2^B)^*$ satisfy

$$
\frac{1}{u'(\widehat{W}_C^B)} - \delta\left[\frac{s(e_C)}{u'(\widehat{W}_C^{BG})} + \frac{1 - s(e_C)}{u'(\widehat{W}_C^{BB})}\right] = 0.
$$

Suppose that $\delta = 1$. Then, the equation cannot hold if $\widehat{\mathcal{V}} \leq (V_2^B)^*$ since this inequality implies that $(\widehat{W}_C^{BG})^* > (\widehat{W}_C^{BB})^* \geq (\widehat{W}_C^G)^*$. Here, the first inequality holds since $e_C > 0$ and the last inequality holds as a strict inequality unless $\mathcal{V} = (V_2^B)^* = 0$. Since this is true for all $\widehat{\mathcal{V}} \in [0, \infty)$, there is $\delta^* \in (0,1)$ such that $(V_2^B)^* < \widehat{\mathcal{V}}$.

There are two possible cases when $e_C = 0$: 1) $\widehat{\mathcal{V}}^*(0) > 0$, and 2) $\widehat{\mathcal{V}}^*(0) = 0$, where $\widehat{\mathcal{V}}^*(e_C)$ is the optimal promotion incentive for a given CEO's effort level e_C. First, I show that $\frac{\partial (V_2^B)^*}{\partial \widehat{\mathcal{V}}} > 0$ and $\frac{\partial \widehat{\mathcal{V}}^*(e_C)}{\partial e_C} < 0$ when $\widehat{\mathcal{V}}^*(e_C) > 0$ for $e_C \in (0,1)$. Notice that, by the implicit function theorem,

$$
\frac{\partial (V_2^B)^*}{\partial \widehat{\mathcal{V}}} = \frac{\frac{u''(\widehat{W}_C^B)}{u'(\widehat{W}_C^B)^3}}{\frac{u''(\widehat{W}_C^B)}{u'(\widehat{W}_C^B)^3} + \delta\left[s(e_C)\frac{u''(\widehat{W}_C^{BG})}{u'(\widehat{W}_C^{BG})^3} + (1 - s(e_C))\frac{u''(\widehat{W}_C^{BB})}{u'(\widehat{W}_C^{BB})^3}\right]} > 0.
$$

Also, this means that $\frac{\partial (V_2^B)^*}{\partial \widehat{\mathcal{V}}}$ is less than 1. Likewise, $0 < \frac{\partial (V_2^G)^*}{\partial \widehat{\mathcal{V}}} < 1$. Hence,

$$
\frac{\partial \widehat{\mathcal{V}}^*(e_C)}{\partial e_C} = -\frac{\frac{\partial^2 \widehat{F}(\widehat{\mathcal{V}}^*)}{\partial e_C \partial \widehat{\mathcal{V}}}}{\frac{\partial^2 \widehat{F}(\widehat{\mathcal{V}}^*)}{\partial \widehat{\mathcal{V}}^2}} < 0,
$$

where

$$\frac{\partial^2 \widehat{F}(\widehat{\mathcal{V}}^*)}{\partial e_C \partial \widehat{\mathcal{V}}} = -\beta\left[\frac{1}{u'((\widehat{W}_C^G)^*)} - \frac{1}{u'((\widehat{W}_C^B)^*)}\right] + s(e_C)\frac{u''((\widehat{W}_C^G)^*)}{u'((\widehat{W}_C^G)^*)^3}\left[(1-s(e_C))\frac{g''(e_C)}{\beta} - \frac{\partial (V_2^G)^*}{\partial e_C}\right]$$
$$+ (1-s(e_C))\frac{u''((\widehat{W}_C^B)^*)}{u'((\widehat{W}_C^B)^*)^3}\left[-s(e_C)\frac{g''(e_C)}{\beta} - \frac{\partial (V_2^B)^*}{\partial e_C}\right]$$
$$= \beta\left[\frac{1}{u'((\widehat{W}_C^G)^*)} - \frac{1}{u'((\widehat{W}_C^B)^*)}\right]$$
$$- s(e_C)\frac{u''((\widehat{W}_C^G)^*)}{u'((\widehat{W}_C^G)^*)^3}\cdot$$
$$\left[\frac{\delta\left[\beta\left(\frac{1}{u'((\widehat{W}_C^{GG})^*)} - \frac{1}{u'((\widehat{W}_C^{GB})^*)}\right) - s(e_c)(1-s(e_C))\frac{g''(e_C)}{\beta}\left(\frac{u''((\widehat{W}_C^{GG})^*)}{u'((\widehat{W}_C^{GG})^*)^3} - \frac{u''((\widehat{W}_C^{GB})^*)}{u'((\widehat{W}_C^{GB})^*)^3}\right)\right]}{\frac{u''((\widehat{W}_C^G)^*)}{u'((\widehat{W}_C^G)^*)^3} + \delta\left[s(e_C)\frac{u''((\widehat{W}_C^{GG})^*)}{u'((\widehat{W}_C^{GG})^*)^3} + (1-s(e_C))\frac{u''((\widehat{W}_C^{GB})^*)}{u'((\widehat{W}_C^{GB})^*)^3}\right]}\right]$$
$$- (1-s(e_C))\frac{u''((\widehat{W}_C^B)^*)}{u'((\widehat{W}_C^B)^*)^3}\cdot$$
$$\left[\frac{\delta\left[\beta\left(\frac{1}{u'((\widehat{W}_C^{BG})^*)} - \frac{1}{u'((\widehat{W}_C^{BB})^*)}\right) - s(e_c)(1-s(e_C))\frac{g''(e_C)}{\beta}\left(\frac{u''((\widehat{W}_C^{BG})^*)}{u'((\widehat{W}_C^{BG})^*)^3} - \frac{u''((\widehat{W}_C^{BB})^*)}{u'((\widehat{W}_C^{BB})^*)^3}\right)\right]}{\frac{u''((\widehat{W}_C^B)^*)}{u'((\widehat{W}_C^B)^*)^3} + \delta\left[s(e_C)\frac{u''((\widehat{W}_C^{BG})^*)}{u'((\widehat{W}_C^{BG})^*)^3} + (1-s(e_C))\frac{u''((\widehat{W}_C^{BB})^*)}{u'((\widehat{W}_C^{BB})^*)^3}\right]}\right]$$
$$+ \delta s(e_C)(1-s(e_C))\frac{g''(e_C)}{\beta}\cdot$$
$$\frac{1}{\left(\frac{u''((\widehat{W}_C^G)^*)}{u'((\widehat{W}_C^G)^*)^3} + \delta\left[s(e_C)\frac{u''((\widehat{W}_C^{GG})^*)}{u'((\widehat{W}_C^{GG})^*)^3} + (1-s(e_C))\frac{u''((\widehat{W}_C^{GB})^*)}{u'((\widehat{W}_C^{GB})^*)^3}\right]\right)}\cdot$$
$$\frac{1}{\left(\frac{u''((\widehat{W}_C^B)^*)}{u'((\widehat{W}_C^B)^*)^3} + \delta\left[s(e_C)\frac{u''((\widehat{W}_C^{BG})^*)}{u'((\widehat{W}_C^{BG})^*)^3} + (1-s(e_C))\frac{u''((\widehat{W}_C^{BB})^*)}{u'((\widehat{W}_C^{BB})^*)^3}\right]\right)}\cdot$$
$$\left\{\frac{u''((\widehat{W}_C^G)^*)}{u'((\widehat{W}_C^G)^*)^3}\frac{u''((\widehat{W}_C^B)^*)}{u'((\widehat{W}_C^B)^*)^3}\left[s(e_C)\left(\frac{u''((\widehat{W}_C^{GG})^*)}{u'((\widehat{W}_C^{GG})^*)^3} - \frac{u''((\widehat{W}_C^{BG})^*)}{u'((\widehat{W}_C^{BG})^*)^3}\right)\right.\right.$$
$$\left.+(1-s(e_C))\left(\frac{u''((\widehat{W}_C^{GB})^*)}{u'((\widehat{W}_C^{GB})^*)^3} - \frac{u''((\widehat{W}_C^{BB})^*)}{u'((\widehat{W}_C^{BB})^*)^3}\right)\right]$$
$$+ \delta\left(s(e_C)\frac{u''((\widehat{W}_C^{GG})^*)}{u'((\widehat{W}_C^{GG})^*)^3} + (1-s(e_C))\frac{u''((\widehat{W}_C^{GB})^*)}{u'((\widehat{W}_C^{GB})^*)^3}\right)\cdot$$
$$\left.\left(s(e_C)\frac{u''((\widehat{W}_C^{BG})^*)}{u'((\widehat{W}_C^{BG})^*)^3} + (1-s(e_C))\frac{u''((\widehat{W}_C^{BB})^*)}{u'((\widehat{W}_C^{BB})^*)^3}\right)\left(\frac{u''((\widehat{W}_C^G)^*)}{u'((\widehat{W}_C^G)^*)^3} - \frac{u''((\widehat{W}_C^B)^*)}{u'((\widehat{W}_C^B)^*)^3}\right)\right\}$$
$$< 0,$$

and

$$\frac{\partial^2 \widehat{F}(\widehat{\mathcal{V}}^*)}{\partial \widehat{\mathcal{V}}^2} = s(e_C)\frac{u''((\widehat{W}_C^G)^*)}{u'((\widehat{W}_C^G)^*)^3}\left[1 - \frac{\partial (V_2^G)^*}{\partial \widehat{\mathcal{V}}}\right] + (1-s(e_C))\frac{u''((\widehat{W}_C^B)^*)}{u'((\widehat{W}_C^B)^*)^3}\left[1 - \frac{\partial (V_2^B)^*}{\partial \widehat{\mathcal{V}}}\right]$$
$$+ \delta(1-(1-s(e_M))^N)P(\mathbf{e}_{-M})\frac{u''((\widehat{W}_{M2}^G)^*)}{u'((\widehat{W}_{M2}^G)^*)^3} < 0,$$

where I use the following two results

$$\begin{aligned}\frac{\partial (V_2^G)^*}{\partial e_C} = &\frac{1}{\frac{u''(\widehat{W}_C^G)}{u'(\widehat{W}_C^G)^3} + \delta\left[s(e_C)\frac{u''(\widehat{W}_C^{GG})}{u'(\widehat{W}_C^{GG})^3} + (1-s(e_C))\frac{u''(\widehat{W}_C^{GB})}{u'(\widehat{W}_C^{GB})^3}\right]}\cdot \\ &\left\{(1-s(e_C))\frac{g''(e_C)}{\beta}\frac{u''(\widehat{W}_C^G)}{u'(\widehat{W}_C^G)^3} + \delta\left[\beta\left(\frac{1}{u'(\widehat{W}_C^{GG})} - \frac{1}{u'(\widehat{W}_C^{GB})}\right)\right.\right. \\ &\left.\left.-s(e_c)(1-s(e_C))\frac{g''(e_C)}{\beta}\left(\frac{u''(\widehat{W}_C^{GG})}{u'(\widehat{W}_C^{GG})^3} - \frac{u''(\widehat{W}_C^{GB})}{u'(\widehat{W}_C^{GB})^3}\right)\right]\right\}, \text{ and} \\ \frac{\partial (V_2^B)^*}{\partial e_C} = &\frac{1}{\frac{u''(\widehat{W}_C^B)}{u'(\widehat{W}_C^B)^3} + \delta\left[s(e_C)\frac{u''(\widehat{W}_C^{BG})}{u'(\widehat{W}_C^{BG})^3} + (1-s(e_C))\frac{u''(\widehat{W}_C^{BB})}{u'(\widehat{W}_C^{BB})^3}\right]}\cdot \\ &\left\{-s(e_C)\frac{g''(e_C)}{\beta}\frac{u''(\widehat{W}_C^B)}{u'(\widehat{W}_C^B)^3} + \delta\left[\beta\left(\frac{1}{u'(\widehat{W}_C^{BG})} - \frac{1}{u'(\widehat{W}_C^{BB})}\right)\right.\right. \\ &\left.\left.-s(e_c)(1-s(e_C))\frac{g''(e_C)}{\beta}\left(\frac{u''(\widehat{W}_C^{BG})}{u'(\widehat{W}_C^{BG})^3} - \frac{u''(\widehat{W}_C^{BB})}{u'(\widehat{W}_C^{BB})^3}\right)\right]\right\}\end{aligned}$$

based on the implicit function theorem.

For the previous results, I exploit the condition $(\widehat{W}_C^G)^* > (\widehat{W}_C^B)^*$, $(\widehat{W}_C^{GG})^*) > (\widehat{W}_C^{BG})^*$, and $(\widehat{W}_C^{GB})^*) > (\widehat{W}_C^{BB})^*$, which all hold since $(V_2^G)^* > (V_2^B)^*$. These imply that

$$\begin{aligned}\frac{1}{u'((\widehat{W}_C^G)^*)} &= \delta\left[\frac{s(e_C)}{u'((\widehat{W}_C^{GG})^*)} + \frac{1-s(e_C)}{u'((\widehat{W}_C^{GB})^*)}\right] \\ &> \delta\left[\frac{s(e_C)}{u'((\widehat{W}_C^{BG})^*)} + \frac{1-s(e_C)}{u'((\widehat{W}_C^{BB})^*)}\right] \\ &= \frac{1}{u'((\widehat{W}_C^B)^*)}.\end{aligned}$$

Here, I show that why the condition, $(V_2^G)^* > (V_2^B)^*$, holds. Suppose $(V_2^G)^* \leq (V_2^B)^*$. Then $(\widehat{W}_C^G)^* \leq (\widehat{W}_C^B)^*$ according to the same logic above. Note that $u((\widehat{W}_C^G)^*) + (V_2^G)^*$ must be strictly greater than $u((\widehat{W}_C^B)^*) + (V_2^B)^*$ in order to induce managers to exert a positive effort. However, two conditions, $(V_2^G)^* \leq (V_2^B)^*$ and $(\widehat{W}_C^G)^*) \leq (\widehat{W}_C^B)^*$, yield $u((\widehat{W}_C^G)^*) + (V_2^G)^* \leq u((\widehat{W}_C^B)^*) + (V_2^B)^*$. Therefore, $(V_2^G)^*$ must be strictly greater than $(V_2^B)^*$.

The next step is to show that there is $\bar{e}_C \in (0,1)$ such that $\widehat{\mathcal{V}}^*(e_C) = 0$ if $e_C \in [\bar{e}_C, 1)$ and $\widehat{\mathcal{V}}^*(e_C) > 0$ if $e_C \in [0, \bar{e}_C)$. Since $\frac{\partial \widehat{\mathcal{V}}^*(e_C)}{\partial e_C} < 0$ when $\widehat{\mathcal{V}}^*(e_C) > 0$, it is enough to show that there is $\bar{e}_C$ such that $\widehat{\mathcal{V}}^*(e_C) = 0$. Recall that

$$\frac{\partial \widehat{F}(\widehat{\mathcal{V}})}{\partial \widehat{\mathcal{V}}} = -\frac{s(e_C)}{u'(\widehat{W}_C^G)} - \frac{1-s(e_C)}{u'(\widehat{W}_C^B)} + \delta\frac{1-(1-s(e_M))^N}{u'(\widehat{W}_{M2}^G)}.$$

When $\widehat{\mathcal{V}} = 0$, the last term is a positive constant regardless of the value of e_C. On the other hand, from the condition

$$\frac{1}{u'(\widehat{W}_C^G)} - \delta\left[\frac{s(e_C)}{u'(\widehat{W}_C^{GG})} + \frac{1-s(e_C)}{u'(\widehat{W}_C^{GB})}\right] = 0$$

it can be shown that the first term approaches negative infinity as e_C approaches one because $(V_2^G)^* \geq \frac{\widehat{\mathcal{V}}}{2} = 0$ and $u((W_C^{GG})^*)$ approaches positive infinity as e_C converges to one.

Hence, there is $\bar{e}_C$ supporting the optimal choice of zero promotion incentive. This result yields that there is $\hat{e}_C \in [0, \bar{e}_C)$ such that $(V_2^B)^* \leq 0$ if $e_C \in [\hat{e}_C, 1)$ since $(V_2^B)^* < \widehat{\mathcal{V}}^*(e_C)$.

The remaining proof is to show that $\mathcal{V}^*(e_C) \geq \widehat{\mathcal{V}}^*(e_C)$ when $e_C \in [\hat{e}_C, 1)$. Notice that, when $\mathcal{V} = \widehat{\mathcal{V}} \in [0, \widehat{\mathcal{V}}^*(e_C)]$ for $e_C \in [\hat{e}_C, 1)$,

$$\frac{\partial F(\mathcal{V})}{\partial \mathcal{V}} > \frac{\partial \widehat{F}(\widehat{\mathcal{V}})}{\partial \widehat{\mathcal{V}}}$$

since $(W_C^G)^* = (\widehat{W}_C^G)^*$, $(W_C^B)^* < (\widehat{W}_C^G)^*$, and $(W_{M1}^G)^* > (\widehat{W}_{M2}^G)^*$. Hence, $\mathcal{V}^*(e_C) \geq \widehat{\mathcal{V}}^*(e_C)$ when $e_C \in [\hat{e}_C, 1)$. Moreover, when $e_C \in [\hat{e}_C, \bar{e}_C]$, $\mathcal{V}^*(e_C) > \widehat{\mathcal{V}}^*(e_C)$ since $\left.\frac{\partial \widehat{F}(\widehat{\mathcal{V}})}{\partial \widehat{\mathcal{V}}}\right|_{\widehat{\mathcal{V}} = \widehat{\mathcal{V}}^*(e_C)} = 0$.

Consider the second case, $\widehat{\mathcal{V}}^*(0) = 0$. In this case, $\widehat{\mathcal{V}}^*(e_C) = 0$ for every $e_C \in (0, 1)$. Hence, $\mathcal{V}^*(e_C) \geq \widehat{\mathcal{V}}^*(e_C)$ regardless of the value of $e_C \in (0, 1)$.

A.1.25 Proof of Proposition 14

For brevity, I denote $e_{M11} = e_{M21}$ by e_{M1} and $e_{M12} = e_{M22}$ by e_{M2}. Suppose that $(U_{M1}^2)^* \leq (U_{M2}^2)^*$. Note that, for a given $\mathcal{V}$, the expected utility for the second period, U_{Mi}^2, is determined according to the equation

$$\frac{1}{u'(W_{Mi}^G)} = \delta\left[\frac{s(e_{Mi2})}{u'(W_{Mi}^{GG})} + \frac{1-s(e_{Mi2})}{u'(W_{Mi}^{GB})}\right],$$

$i = 1,\ 2$. This equation and the condition that $(U_{M1}^2)^* \leq (U_{M2}^2)^*$ imply that $(W_{M1}^G)^* \leq (W_{M2}^G)^*$. In order for the inequality to hold, the following must hold:

$$\begin{aligned}(1 - s(e_{C1}))P(\mathbf{e}_{-M1})\mathcal{V} + &(1 - (1 - s(e_{C1}))P(\mathbf{e}_{-M1}))(U_{M1}^2)^* \\ &\geq P(\mathbf{e}_{-M1})\mathcal{V} + (1 - P(\mathbf{e}_{-M1}))(U_{M2}^2)^*,\end{aligned}$$

which implies that

$$(1 - P(\mathbf{e}_{-M1}))((U_{M1}^2)^* - (U_{M2}^2)^*) + s(e_{C1})P(\mathbf{e}_{-M1})(U_{M1}^2)^* \geq s(e_C)P(\mathbf{e}_{-M})\mathcal{V}.$$

Then, $(U_{M1}^2)^*$ must be greater than $\mathcal{V}$ since $(U_{M1}^2)^* \leq (U_{M2}^2)^*$. This contradicts to the given condition. Hence, $(U_{M1}^2)^* > (U_{M2}^2)^*$ if $\mathcal{V}^* > (U_{M1}^2)^*$. Moreover, the difference between $u(W_{M1}^G)$ and $u(W_{M2}^G)$ is

$$u(W_{M1}^G) - u(W_{M2}^G) = s(e_{C1})P(\mathbf{e}_{-M})\mathcal{V} - \left[(1 - P(\mathbf{e}_{-M}))(U_{M1}^2 - U_{M2}^2) + s(e_{C1})P(\mathbf{e}_{-M1})U_{M1}^2\right],$$

which has a positive value when $\mathcal{V}^* > (U_{M1}^2)^* > (U_{M2}^2)^*$. That is, $(W_{M1}^G)^* > (W_{M2}^G)^*$.

A.1.26 Proof of Proposition 15

First, I show that there is a constant $\widehat{N}$ such that $\mathcal{V}^*(N+1) - \mathcal{V}^*(N) \leq 0$ if $N > \widehat{N}$.

For a given $\mathcal{V}$ and N, $e_C(N)$ is determined by

$$E\left[f\left(\sum_{i=1}^{N} X_i\right)\right]\beta(\mathcal{G}_C - \mathcal{B}_C) = \beta(W_C^G(N) - W_C^B(N)) + s(e_C(N))(1 - s(e_C(N)))\frac{g''(e_C(N))}{\beta}\left[\frac{1}{u'(W_C^G(N))} - \frac{1}{u'(W_C^B(N))}\right]. \tag{A.6}$$

Claim 7 *There is $\mathcal{M}$ such that*

$$E\left[f\left(\sum_{i=1}^{N+1} X_i\right)\right] - E\left[f\left(\sum_{i=1}^{N} X_i\right)\right] > \mathcal{M}$$

for every N.

Proof 10 *For brevity, denote*

$$E\left[f\left(\sum_{i=1}^{N} X_i\right)\right] = \sum_{i=0}^{N}\binom{N}{i}s(e_M)^i(1 - s(e_M))^{N-i}f(i\mathcal{G}_M + (N-i)\mathcal{B}_M)$$

by $I(N)$. Also, I denote

$$\min_i \left[f(i\mathcal{G}_M + (N+1-i)\mathcal{B}_M) - f(i\mathcal{G} + (N-i)\mathcal{B}_M)\right]$$

by $\underline{f}$. Note that there is $\mathcal{M}_f > 0$ such that $\underline{f} \geq \mathcal{M}_f$ for every N since $f'(x) > 0$ for every $x \geq 0$.

Then,

$$I(N+1)-I(N) = \sum_{i=0}^{N} \binom{N}{i} s(e_M)^i (1-s(e_M))^{N-i} \cdot$$
$$\left[\frac{N+1}{N+1-i}(1-s(e_M)) f(i\mathcal{G}_M + (N+1-i)\mathcal{B}_M) - f(i\mathcal{G}_M + (N-i)\mathcal{B}_M)\right]$$
$$+ s(e_M)^{N+1} f((N+1)\mathcal{G}_M)$$
$$\geq \sum_{i=0}^{N} \binom{N}{i} s(e_M)^i (1-s(e_M))^{N-i} \cdot$$
$$\left[\frac{N+1}{N+1-i}(1-s(e_M)) - 1\right] f(i\mathcal{G}_M + (N-i)\mathcal{B}_M)$$
$$+ s(e_M)^{N+1} f((N+1)\mathcal{G}_M) + \underline{f}.$$

Denote $\lceil (N+1)s(e_M) \rceil$ *by* $\hat{s}$. *Then,*

$$\left[\frac{N+1}{N+1-i}(1-s(e_M)) - 1\right] \begin{cases} > 0 & \text{if } i \geq \hat{s} \\ \leq 0 & \text{otherwise.} \end{cases}$$

There are two possible cases.

1. *(*$\hat{s} = N+1$*)*

   Then,

$$I(N+h) - I(h) \geq \sum_{i=0}^{N} \binom{N}{i} s(e_M)^i (1-s(e_M))^{N-i} \cdot$$
$$\left[\frac{N+1}{N+1-i}(1-s(e_M)) - 1\right] f(i\mathcal{G}_M + (N-i)\mathcal{B}_M)$$
$$+ s(e_M)^{N+1} f(N\mathcal{G}_M)$$
$$+ s(e_M)^{N+1} \left[f((N+1)\mathcal{G}_M) - f(N\mathcal{G}_M)\right] + \underline{f}$$
$$> s(e_M)^{N+1} [f((N+h)\mathcal{G}_M) - f(N\mathcal{G}_M)] + \underline{f},$$

   where the last inequality holds since

$$\sum_{i=0}^{N} \binom{N}{i} s(e_M)^i (1-s(e_M))^{N-i} \left| \left[\frac{(N+h)!(N-i)!}{N!(N+h-i)!}(1-s(e_M))^h - 1\right] \right| = s(e_M)^{N+1}.$$

2. *(*$\hat{s} < N+1$*)*

   First, notice that if $\frac{N}{N+1} > s(e_M)$, *then* $\hat{s} < N+1$. *That is, if* N *is sufficiently large,* $\hat{s} < N+1$. *In this case,*

$$
\begin{aligned}
I(N+1) - I(N) \geq & \sum_{i=0}^{\hat{s}-1} \binom{N}{i} s(e_M)^i (1 - s(e_M))^{N-i} \cdot \\
& \left[\frac{N+1}{N+1-i} (1 - s(e_M)) - 1 \right] f(i\mathcal{G}_M + (N-i)\mathcal{B}_M) \\
& + \sum_{i=\hat{s}}^{N} \binom{N}{i} s(e_M)^i (1 - s(e_M))^{N-i} \cdot \\
& \left[\frac{N+1}{N+1-i} (1 - s(e_M)) - 1 \right] f(i\mathcal{G}_M + (N-i)\mathcal{B}_M) \\
& + s(e_M)^{N+1} f(N\mathcal{G}_M) + \underline{f} \\
& > \sum_{i=\hat{s}}^{N} \binom{N}{i} s(e_M)^i (1 - s(e_M))^{N-i} \left[\frac{N+1}{N+1-i} (1 - s(e_M)) - 1 \right] \cdot \\
& [f(\hat{s}\mathcal{G}_M + (N-\hat{s})\mathcal{B}_M) - f((\hat{s}-1)\mathcal{G}_M + (N-\hat{s}+1)\mathcal{B}_M)] \\
& + s(e_M)^{N+1} f((N+1)\mathcal{G}_M) + \underline{f} \\
& > \sum_{i=\hat{s}}^{N} \binom{N}{i} s(e_M)^i (1 - s(e_M))^{N-i} \left[\frac{N+1}{N+1-i} (1 - s(e_M)) - 1 \right] \cdot \\
& [f(\hat{s}\mathcal{G}_M + (N-\hat{s})\mathcal{B}_M) - f((\hat{s}-1)\mathcal{G}_M + (N-\hat{s}+1)\mathcal{B}_M)] + \underline{f} \\
& > \underline{f}
\end{aligned}
$$

Hence

$$
E\left[f\left(\sum_{i=1}^{N+1} X_i \right) \right] - E\left[f\left(\sum_{i=1}^{N} X_i \right) \right] > \underline{f} \geq \mathcal{M}_f > 0.
$$

This result means that $e_C(N+1) > e_C(N)$ if $\mathcal{V}$ is fixed.

Now, I show that there is $\widetilde{N}$ such that $\frac{\partial F(\mathcal{V})}{\partial \mathcal{V}} < 0$ for every $\mathcal{V} \in [0, \bar{\mathcal{V}}]$. Note that

$$
\begin{aligned}
\frac{\partial F(\mathcal{V}|N)}{\partial \mathcal{V}} &= -\frac{s(e_C(\mathcal{V}, N))}{u'(W_C^G(\mathcal{V}, e_C(\mathcal{V}, N))} - \frac{1 - s(e_C(\mathcal{V}, N))}{u'(W_C^B(\mathcal{V}, e_C(\mathcal{V}, N))} + \frac{1 - (1 - s(e_M))^N}{u'(W_M^G)} \\
&\leq -\frac{s(e_C(\bar{\mathcal{V}}, N))}{u'(W_C^G(0, e_C(\bar{\mathcal{V}}, N))} - \frac{1 - s(e_C(\bar{\mathcal{V}}, N))}{u'(W_C^B(0, e_C(\bar{\mathcal{V}}, N))} + \frac{1 - (1 - s(e_M))^N}{u'(W_M^G)},
\end{aligned}
$$

where

$$
\begin{aligned}
W_C^G(\mathcal{V}_1, e_C(\mathcal{V}_2, N)) &= \mathcal{V}_1 + g(e_C(\mathcal{V}_2, N)) + (1 - s(e_C(\mathcal{V}_2, N))) \frac{g'(e_C(\mathcal{V}_2, N))}{\beta} \\
W_C^B(\mathcal{V}_1, e_C(\mathcal{V}_2, N)) &= \mathcal{V}_1 + g(e_C(\mathcal{V}_2, N)) - s(e_C(\mathcal{V}_2, N)) \frac{g'(e_C(\mathcal{V}_2, N))}{\beta}
\end{aligned}
$$

and $e_C(\mathcal{V}_2, N)$ satisfies

$$E\left[f\left(\sum_{i=1}^{N} X_i\right)\right]\beta(\mathcal{G}_C - \mathcal{B}_C) = \beta(W_C^G(\mathcal{V}_2, e_C(\mathcal{V}_2, N)) - W_C^B(\mathcal{V}_2, e_C(\mathcal{V}_2, N)) + s(e_C(\mathcal{V}_2, N))(1 - s(e_C(\mathcal{V}_2, N)))\frac{g''(e_C(\mathcal{V}_2, N))}{\beta} \cdot \left[\frac{1}{u'(W_C^G(\mathcal{V}_2, e_C(\mathcal{V}_2, N)))} - \frac{1}{u'(W_C^B(\mathcal{V}_2, e_C(\mathcal{V}_2, N)))}\right]$$

since

$$\frac{\partial}{\partial \mathcal{V}_1}\left[\frac{s(e_C(\mathcal{V}_2, N))}{u'(W_C^G(\mathcal{V}_1, e_C(\mathcal{V}_2, N))} + \frac{1 - s(e_C(\mathcal{V}_2, N))}{u'(W_C^B(\mathcal{V}_1, e_C(\mathcal{V}_2, N))}\right] > 0,$$

$$\frac{\partial}{\partial \mathcal{V}_2}\left[\frac{s(e_C(\mathcal{V}_2, N))}{u'(W_C^G(\mathcal{V}_1, e_C(\mathcal{V}_2, N))} + \frac{1 - s(e_C(\mathcal{V}_2, N))}{u'(W_C^B(\mathcal{V}_1, e_C(\mathcal{V}_2, N))}\right] < 0.$$

Since $\mathcal{V}$ is bounded $\frac{1-(1-s(e_M))^N}{u'(W_M^G)} \leq \frac{1}{u'(\underline{W}_M^G)}$, where $\underline{W}_M^G$ satisfies

$$u(\underline{W}_M^G) = \underline{U}_M + g(e_M) + (1 - s(e_M))\frac{g'(e_M)}{\beta}.$$

Moreover, there is $\hat{e}_C \in (0, 1)$ such that

$$\frac{s(e_C)}{u'(W_C^G(0, e_C))} + \frac{1 - s(e_C)}{u'(W_C^B(0, e_C)} > \frac{1}{u'(\underline{W}_M^G)}$$

if $e_C \geq \hat{e}_C$ since $\lim_{e_C \to 1} W_C^G(0, e_C) = \infty$. Since there is N_1 such that $e_C(\bar{\mathcal{V}}, N) \geq \hat{e}_C$ if $N \geq N_1$,

$$\begin{aligned}\frac{\partial F(\mathcal{V}|N)}{\partial \mathcal{V}} &\leq -\frac{s(e_C(\bar{\mathcal{V}}, N))}{u'(W_C^G(0, e_C(\bar{\mathcal{V}}, N))} - \frac{1 - s(e_C(\bar{\mathcal{V}}, N))}{u'(W_C^B(0, e_C(\bar{\mathcal{V}}, N))} + \frac{1 - (1 - s(e_M))^N}{u'(W_M^G)} \\ &< -\frac{s(e_C(\bar{\mathcal{V}}, N))}{u'(W_C^G(0, e_C(\bar{\mathcal{V}}, N))} - \frac{1 - s(e_C(\bar{\mathcal{V}}, N))}{u'(W_C^B(0, e_C(\bar{\mathcal{V}}, N))} + \frac{1}{u'(\underline{W}_M^G)} \\ &< 0\end{aligned}$$

when $N \geq N_1$.

Hence, $\mathcal{V}^* = 0$ if $N \geq N_1$. Also, this implies that there is $N_2 < N_1$ such that $\mathcal{V}^*(N+1) - \mathcal{V}^*(N) < 0$ when $N \in [N_2, N_1 - 1]$ if there is $N^* < N_1$ such that $\mathcal{V}^*(N^*) > 0$.

Moreover, $e_C^*(N+1) - e_C^*(N) > 0$ when $N \geq N_2$ since

$$\frac{\partial E(\mathcal{V}, e_C)}{\partial \mathcal{V}} > 0 \text{ and } \frac{\partial E(\mathcal{V}, e_C)}{\partial e_C} > 0,$$

where

$$E(\mathcal{V}, e_C) \equiv \beta(W_C^G(\mathcal{V}, e_C) - W_C^B(\mathcal{V}, e_C)) + s(e_C)(1 - s(e_C))\frac{g''(e_C)}{\beta}\left[\frac{1}{u'(W_C^G(\mathcal{V}, e_C))} - \frac{1}{u'(W_C^B(\mathcal{V}, e_C))}\right]$$

comes from the right hand side of (A.6).

A.1.27 Proof of Proposition 16

First, note that constraints regarding managers give the following results:

$$
\begin{aligned}
u(W_C) &= \mathcal{V} \\
u(W_M^G) &= \underline{U}_M + g(e_L) + (1 - s_L(e_L))\frac{g'(e_L)}{\underline{\beta}} - P(\mathbf{e}_{-M})\mathcal{V}, \\
u(W_M^B) &= \underline{U}_M + g(e_L) - s_L(e_L)\frac{g'(e_L)}{\underline{\beta}} - R(\mathbf{e}_{-M})\mathcal{V}, \text{ and} \\
e_H &= e_L\frac{\bar{\beta}}{\underline{\beta}}
\end{aligned}
$$

for a given $(e_L, \mathcal{V})$. From now on, I use subscript 1 for promotion rule 1 and subscript 2 for promotion 2 in order to distinguish two problems. For a given e_{L1} and e_{L2}, denote the optimal $\mathcal{V}$ by $\mathcal{V}_1^*(e_{L1})$ under promotion rule 1 and $\mathcal{V}_2^*(e_{L1})$ under promotion rule 2, Also, denote the firm's objective function under promotion rule 1 and promotion rule 2 by $F_1(e_{L1}|\gamma)$ and $F_2(e_{L2}|\gamma)$, respectively, for a given γ. That is,

$$
\begin{aligned}
F_1(e_L|\gamma) = \gamma\left[(\bar{\beta} - \underline{\beta})H_1(e_L) + \underline{\beta}\right] &- W_{C1} + 2[qs_H(e_H) + (1-q)s_L(e_L)](\mathcal{G} - W_{M1}^G) \\
&+ 2[1 - qs_H(e_H) - (1-q)s_L(e_L)](\mathcal{B} - W_{M1}^B), \text{ and} \\
F_2(e_L|\gamma) = \gamma\left[(\bar{\beta} - \underline{\beta})H_2(e_L) + \underline{\beta}\right] &- W_{C2} + 2[qs_H(e_H) + (1-q)s_L(e_L)](\mathcal{G} - W_{M2}^G) \\
&+ 2[1 - qs_H(e_H) - (1-q)s_L(e_L)](\mathcal{B} - W_{M2}^B),
\end{aligned}
$$

where

$$
\begin{aligned}
H_1(e_{L1}) = {} & \frac{qs_H(e_{H1})}{qs_H(e_{H1}) + (1-q)s_L(e_{L1})}\left[2(qs_H(e_{H1}) + (1-q)s_L(e_{L1})) - (qs_H(e_{H1})\right. \\
& \left. +(1-q)s_L(e_{L1}))^2\right] \\
& + q\left[1 - 2(qs_H(e_{H1}) + (1-q)s_L(e_{L1})) + (qs_H(e_{H1}) + (1-q)s_L(e_{L1}))^2\right], \text{ and} \\
H_2(e_{L2}) = {} & q + q(1-q)(s_H(e_{H2}) - s_L(e_{L2})).
\end{aligned}
$$

Then, the firm's problem is to choose e_L in order to maximize its objective function. Notice that, for a given e_{Lj}, $\mathcal{V}_j^*(e_{Lj})$, $j = 1$ and 2, satisfies

$$
\begin{aligned}
\frac{\partial F_1(e_{L1}|\gamma)}{\partial \mathcal{V}_1} &= -\frac{1}{u'(W_{C1})} + \frac{2(qs_H(e_{H1}) + (1-q)s_L(e_{L1}))}{u'(W_{M1}^G)}P(\mathbf{e}_{-M1}) = 0, \text{ and} \\
\frac{\partial F_2(e_{L2}|\gamma)}{\partial \mathcal{V}_2} &= -\frac{1}{u'(W_{C2})} + \frac{2(qs_H(e_{H2}) + (1-q)s_L(e_{L2}))}{u'(W_{M2}^G)}P(\mathbf{e}_{-M2}) \\
&\quad + \frac{2(1 - qs_H(e_{H2}) - (1-q)s_L(e_{L2}))}{u'(W_{M2}^B)}R(\mathbf{e}_{-M2}) = 0.
\end{aligned}
$$

Also, the first order conditions with respect to e_{Lj} are

$$\begin{aligned}
\frac{\partial F_1(e_{L1}|\gamma)}{\partial e_{L1}} &= \gamma(\bar{\beta}-\underline{\beta})\frac{\partial H_1(e_{L1})}{\partial e_{L1}} + 2\left(q\frac{\bar{\beta}^2}{\underline{\beta}} + (1-q)\underline{\beta}\right)[\mathcal{G}-\mathcal{B}-(W^G_{M1}-W^B_{M1})] \\
&\quad - 2\frac{\kappa}{\underline{\beta}}\left[(1-s_L(e_{L1}))\frac{qs_H(e_{H1})+(1-q)s_L(e_{L1})}{u'(W^G_{M1})}\right. \\
&\quad \left. - s_L(e_{L1})\frac{1-qs_H(e_{H1})-(1-q)s_L(e_{L1})}{u'(W^B_{M1})}\right] \\
&\quad + 2\frac{qs_H(e_{H1})+(1-q)s_L(e_{L1})}{u'(W^G_{M1})}\frac{\partial P(\mathbf{e}_{-M1})}{\partial e_{L1}}\mathcal{V}^*_1(e_{L1}), \text{ and} \\
\frac{\partial F_2(e_{L2}|\gamma)}{\partial e_{L2}} &= \gamma(\bar{\beta}-\underline{\beta})\frac{\partial H_2(e_{L2})}{\partial e_{L2}} + 2\left(q\frac{\bar{\beta}^2}{\underline{\beta}} + (1-q)\underline{\beta}\right)[\mathcal{G}-\mathcal{B}-(W^G_{M2}-W^B_{M2})] \\
&\quad - 2\frac{\kappa}{\underline{\beta}}\left[(1-s_L(e_{L2}))\frac{qs_H(e_{H2})+(1-q)s_L(e_{L2})}{u'(W^G_{M2})}\right. \\
&\quad \left. - s_L(e_{L2})\frac{1-qs_H(e_{H2})-(1-q)s_L(e_{L2})}{u'(W^B_{M2})}\right] \\
&\quad + 2\frac{qs_H(e_{H2})+(1-q)s_L(e_{L2})}{u'(W^G_{M2})}\frac{\partial P(\mathbf{e}_{-M2})}{\partial e_{L2}}\mathcal{V}^*_2(e_{L2}) \\
&\quad + 2\frac{1-qs_H(e_{H2})-(1-q)s_L(e_{L2})}{u'(W^B_{M2})}\frac{\partial R(\mathbf{e}_{-M2})}{\partial e_{L2}}\mathcal{V}^*_2(e_{L2}).
\end{aligned}$$

According to Milgrom and Shannon (1994), $\frac{\partial e^*_{L1}}{\partial \gamma} \geq 0$ and $\frac{\partial e^*_{L1}}{\partial \gamma} \geq 0$ since

$$\begin{aligned}
\frac{\partial^2 F_1(e_{L1}|\gamma)}{\partial e_{L1}\partial\gamma} &= (\bar{\beta}-\underline{\beta})\frac{\partial H_1(e_{L1})}{\partial e_{L1}} \\
&= 2(\bar{\beta}-\underline{\beta})q(1-q))\left(\frac{\bar{\beta}^2}{\underline{\beta}}-\underline{\beta}\right)[1-qs_H(e_{H1})-(1-q)s_L(e_{L1})] > 0, \text{ and} \\
\frac{\partial^2 F_2(e_{L2}|\gamma)}{\partial e_{L2}\partial\gamma} &= (\bar{\beta}-\underline{\beta})\frac{\partial H_2(e_{L2})}{\partial e_{L2}} \\
&= (\bar{\beta}-\underline{\beta})q(1-q)\left(\frac{\bar{\beta}^2}{\underline{\beta}}-\underline{\beta}\right) > 0.
\end{aligned}$$

Moreover, these inequalities imply that $\frac{\partial e_{L1}}{\partial \gamma} > 0$ and $\frac{\partial e_{L1}}{\partial \gamma} > 0$ if $e^*_{L1} \in \left(0, \frac{\beta}{\bar{\bar{\beta}}}\right)$ and $e^*_{L2} \in \left(0, \frac{\beta}{\bar{\bar{\beta}}}\right)$, respectively, according to Edlin and Shannon (1998).

Also, there is γ^*_1 such that $e^*_{L1} = \frac{\beta}{\bar{\bar{\beta}}}$, which means that $e_{H1} = 1$, if $\gamma \geq \gamma^*_1$ since $\left.\frac{\partial F_1(e_L|\gamma)}{\partial e_L}\right|_{e_{L1}=1}$ is a strictly increasing function in γ and $\lim_{\gamma\to\infty} \left.\frac{\partial F_1(e_L|\gamma)}{\partial e_L}\right|_{e_{L1}=1} = \infty$.

Moreover, by the envelope theorem,

$$\frac{\partial F_1(e^*_{L1}|\gamma)}{\partial \gamma} = (\bar{\beta}-\underline{\beta})H_1(e^*_{L1}) + \underline{\beta}, \text{ and } \frac{\partial F_2(e^*_{L2}|\gamma)}{\partial \gamma} = (\bar{\beta}-\underline{\beta})H_2(e^*_{L2}) + \underline{\beta}.$$

Since $H_1(e_{L1}) > H_2(e_{L1})$ when $e_{L1} = e_{L2}$, there is $\hat{e}_{L1} \in \left(0, \frac{\beta}{\bar{\bar{\beta}}}\right)$ such that $H_1(\hat{e}_{L1}) > H_2\left(\frac{\beta}{\bar{\bar{\beta}}}\right)$. Hence, there is $\hat{\gamma}$ such that $F_1(e^*_{L1}|\gamma) > F_2(e^*_{L2}|\gamma)$ if $\gamma \geq \hat{\gamma}$.

Now, I show that $F_1(e^*_{L1}|\gamma) < F_2(e^*_{L2}|\gamma)$ when $\gamma = 0$. This is true since

$$F_2(e^*_{L2}|\gamma) \geq F_2(e^*_{L1}|\gamma) > F_1(e^*_{L1}|\gamma).$$

The second inequality holds since $(W^G_{M1})^* = W^G_{M2}$ and $(W^B_{M1})^* > W^B_{M2}$ if $(e_{L2}, \mathcal{V}_2) = (e^*_{L1}, \mathcal{V}^*_1)$. Hence, $\hat{\gamma} > 0$.

A.2 Firm's Problems in Detail

A.2.1 The Firm's Problem in Section 2.7.1

Under this extension, the firm's problem is to choose $\mathcal{V} \in [0, \infty)$ maximizing $F(\mathcal{V})$ defined by

$$\begin{aligned} F(\mathcal{V}) \equiv & \max_{\{(W^G_C, W^B_C), (W^G_M, W^B_M), (W^{GG}_M, W^{GB}_M),\}} s(e_C)(\mathcal{G}_C - W^G_C) + (1 - s(e_C))(\mathcal{B}_C - W^B_C) \\ & + N\left[s(e_{M1})(\mathcal{G}_M - W^G_M) + (1 - s(e_{M1}))(\mathcal{B}_M - W^B_M)\right] \\ & + \delta(Ns(e_{M1}) - 1 + (1 - s(e_{M1}))^N)[s(e_{M2})(\mathcal{G}_M - W^{GG}_M) + (1 - s(e_{M2}))(\mathcal{B}_M - W^{GB}_M)] \\ & \text{subject to} \end{aligned}$$

$$\begin{aligned} & E[\mathcal{U}(W^G_C, W^B_C, e_{C1})] = \mathcal{V}, \\ & E[\mathcal{U}(W^G_M, W^B_M, e_{M1})] + s(e_{M1})\{P(\mathbf{e}_{-M1})\mathcal{V} + (1 - P(\mathbf{e}_{-M1}))E[\mathcal{U}(W^{GG}_M, W^{GB}_M, e_{M2})]\} = \underline{U}_M, \\ & e_C \in \arg\max_{\hat{e}} E[\mathcal{U}(W^G_C, W^B_C, \hat{e}], \\ & e_{M1} \in \arg\max_{\hat{e}} E[\mathcal{U}(W^G_M, W^B_M, \hat{e})] + s(\hat{e})\{P(\mathbf{e}_{-M1})\mathcal{V} + (1 - P(\mathbf{e}_{-M1}))E[\mathcal{U}(W^{GG}_M, W^{GB}_M, e_{M2})]\}, \\ & e_{M2} \in \arg\max_{\hat{e}} E[\mathcal{U}(W^{GG}_M, W^{GB}_M, \hat{e})], \end{aligned}$$

where

$$E[\mathcal{U}(W^G, W^B, e)] = s(e)u(W^G) + (1 - s(e))u(W^B) - g(e).$$

Note that the expected number of senior managers in the second period is

$$\begin{aligned} \sum_{k=1}^{N} \binom{N}{k} s(e_{M1})^k (1 - s(e_{M1}))^{N-k}(k-1) &= \sum_{k=0}^{N} \binom{N}{k} k s(e_{M1})^k (1 - s(e_{M1}))^{N-k} \\ & \quad - \sum_{k=0}^{N} \binom{N}{k} s(e_{M1})^k (1 - s(e_{M1}))^{N-k} + (1 - s(e_{M1}))^N \\ &= Ns(e_{M1}) - 1 + (1 - s(e_{M1}))^N. \end{aligned}$$

A.2.2 The Firm's Problem in Section 2.7.2

The objective of the firm is to choose $\mathcal{V} \in [0, \infty)$ maximizing $F(\mathcal{V})$ defined by

$$
\begin{aligned}
F(\mathcal{V}) \equiv \max_{\mathcal{A}} \; & s(e_{C1})(\mathcal{G}_C - W_C^G) + (1 - s(e_{C1}))(\mathcal{B}_C - W_C^B) \\
& + \delta s(e_{C1})[s(e_{C2})(\mathcal{G}_C - W_C^{GG}) + (1 - s(e_{C2}))(\mathcal{B}_C - W_C^{GB})] \\
& + N\left[s(e_{M1})(\mathcal{G}_M - W_{M1}^G) + (1 - s(e_{M1}))(\mathcal{B}_M - W_{M1}^B)\right] \\
& + \delta N s(e_{C1})\left[s(e_{M2})(\mathcal{G}_M - W_{M2}^G) + (1 - s(e_{M2}))(\mathcal{B}_M - W_{M2}^B)\right]
\end{aligned}
$$

subject to

$$
\begin{aligned}
E[\mathcal{U}(W_C^G, W_C^B, e_{C1})] + s(e_{C1})E[\mathcal{U}(W_C^{GG}, W_C^{GB}, e_{C2})] = \mathcal{V} \; (IR_C) \\
E[\mathcal{U}(W_{M1}^G, W_{M1}^B, e_{M1})] + s(e_{M1})(1 - s(e_{C1}))P(\mathbf{e}_{-M1})\mathcal{V} = \underline{U}_M \; (IR_{M1}), \\
E[\mathcal{U}(W_{M2}^G, W_{M2}^B, e_{M2})] + s(e_{M2})P(\mathbf{e}_{-M2})\mathcal{V} = \underline{U}_M \; (IR_{M2}) \\
e_{C1} \in \arg\max_{\hat{e}} E[\mathcal{U}(W_C^G, W_C^B, \hat{e}] + s(\hat{e})E[\mathcal{U}(W_C^{GG}, W_C^{GB}, e_{C2})] \; (IC_{C1}) \\
e_{C2} \in \arg\max_{\hat{e}} E[\mathcal{U}(W_C^{GG}, W_C^{GB}, \hat{e})] \; (IC_{C2}) \\
e_{M1} \in \arg\max_{\hat{e}} E[\mathcal{U}(W_{M1}^G, W_{M1}^B, \hat{e})] + s(\hat{e})(1 - s(e_{C1}))P(\mathbf{e}_{-M1})\mathcal{V} \; (IC_{M1}), \\
e_{M2} \in \arg\max_{\hat{e}} E[\mathcal{U}(W_{M2}^G, W_{M2}^B, \hat{e})] + s(\hat{e})P(\mathbf{e}_{-M2})\mathcal{V} \; (IC_{M2}),
\end{aligned}
$$

where

$$
\begin{aligned}
\mathcal{A} = \{(W_C^G, W_C^B, W_C^{GG}, W_C^{GB}), (W_{M1}^G, W_{M2}^B), (W_{M2}^G, W_{M2}^B)\}, \text{ and} \\
E[\mathcal{U}(W^G, W_B, e)] = s(e)u(W^G) + (1 - s(e))u(W^B) - g(e).
\end{aligned}
$$

When the CEO's individual rationality constraint binds at $\mathcal{V}$, the compensation scheme $(W_C^G, W_C^B, W_C^{GG}, W_C^{GB})$ for the CEO satisfies

$$
\begin{aligned}
u(W_C^G) &= \mathcal{V} + g(e_{C1}) + (1 - s(e_{C1}))\frac{g'(e_{C1})}{\beta} - V_2, \\
u(W_C^B) &= \mathcal{V} + g(e_{C1}) - s(e_{C1})\frac{g'(e_{C1})}{\beta}, \\
u(W_C^{GG}) &= V_2 + g(e_{C2}) + (1 - s(e_{C2}))\frac{g'(e_{C2})}{\beta}, \\
u(W_C^{GB}) &= V_2 + g(e_{C2}) - s(e_{C2})\frac{g'(e_{C2})}{\beta},
\end{aligned}
$$

where

$$
V_2 = s(e_{C2})u(W_C^{GG}) + (1 - s(e_{C2}))u(W_C^{GB}) - g(e_{C2})
$$

is the successful CEO's expected utility in the second period.

On the other hand, the compensation schemes for managers are characterized by

$$u(W^G_{M1}) = \underline{U}_M + g(e_{M1}) + (1 - s(e_{M1}))\frac{g'(e_{M1})}{\beta} - (1 - s(e_{C1}))P(\mathbf{e}_{-M1})\mathcal{V},$$
$$u(W^B_{M1}) = \underline{U}_M + g(e_{M1}) - s(e_{M1})\frac{g'(e_{M1})}{\beta},$$
$$u(W^G_{M2}) = \underline{U}_M + g(e_{M2}) + (1 - s((e_{M2}))\frac{g'((e_{M2})}{\beta} - P(\mathbf{e}_{-M2})\mathcal{V}, \text{ and}$$
$$u(W^B_{M2}) = \underline{U}_M + g(e_{M2}) - s(e_{M2})\frac{g'((e_{M2})}{\beta}.$$

A.2.3 The Firm's Problem in Section 2.9

The expected profit from two managers is

$$\begin{aligned} E[\Pi_M|(e_H, e_L, W^G_M, W^B_M)] = {} & q^2 2[s_H(e_H)(\mathcal{G}_M - W^G_M) + (1 - s_H(e_H))(\mathcal{B}_M - W^B_M)] \\ & + 2q(1-q)[s_H(e_H)(\mathcal{G}_M - W^G_M) + (1 - s_H(e_H))(\mathcal{B}_M - W^B_M) \\ & + s_L(e_L)(\mathcal{G}_M - W^G_M) + (1 - s_L(e_L))(\mathcal{B}_M - W^B_M)] \\ & + (1-q)^2 2[s_L(e_L)(\mathcal{G}_M - W^G_M) + (1 - s_L(e_L))(\mathcal{B}_M - W^B_M)]. \end{aligned}$$

Also, the choice of promotion rule determines the expected β when the firm requires e_H and e_L from high-type and low-type managers according to 1) when the firm uses promotion rule 1, $E[\beta|e_H, e_L]$ is equal to

$$\begin{aligned} E^{P1}[\beta|e_H, e_L] = {} & \frac{qs_H(e_H)}{qs_H(e_H) + (1-q)s_L(e_L)}\left[2(qs_H(e_H) + (1-q)s_L(e_L)) \right. \\ & \left. -(qs_H(e_H) + (1-q)s_L(e_L))^2\right] \\ & + q\left[1 - 2(qs_H(e_H) + (1-q)s_L(e_L)) + (qs_H(e_H) + (1-q)s_L(e_L))^2\right], \end{aligned}$$

2) while the expectation has the following value

$$E^{P2}[\beta|e_H, e_L] = q + q(1-q)(s_H(e_H) - s_L(e_L))$$

if the firm adopts promotion rule 2.

Appendix B

APPENDIX TO CHAPTER 3

B.1 Proofs

B.1.1 Proof of Proposition 17

If the agent chooses $a = 0$,

$$\rho W^k dt = \phi dt + \frac{1}{2}\sigma^2 \frac{\partial W^k}{\partial I^2} dt.$$

On the other hand, choosing $a = \mu$ yields

$$\rho W^k dt = \mu \frac{\partial W^k}{\partial I} dt + \frac{1}{2}\sigma^2 \frac{\partial W^k}{\partial I^2} dt.$$

Therefore, the contract is incentive-compatible if

$$\mu \frac{\partial W^k}{\partial I} dt \geq \phi dt$$

for every I and k.

B.1.2 Proof of Proposition 18

By the **Proposition 17**, the incentive-compatible contract must satisfy

$$\frac{\partial W^k}{\partial I} \geq \frac{\phi}{\mu} \text{ for all } I \in (\underline{I}_{k-1}, I_{u,k}).$$

Also, the limited liability condition impies that $W(I) \geq 0$ for all $I \in [\underline{I}, \bar{I}]$. Note that the payment $\mathcal{B}_u$ is the same with $W(\bar{I})$. Suppose that $W(\bar{I}) = 0$. The condition $\partial W^k(I)/\partial I \geq \phi/\mu > 0$, for all $I \in (\underline{I}_{k-1}, I_{u,k})$ and k, implies that $W(I) < 0$ for $I < \bar{I}$. This violates the limited liability condition.

B.1.3 Proof of Proposition 19.

It is enough to show that the principal can't incentivize the agent by transfering all output from the project to the agent since it is the maximum transfer for the principal to the agent without loss. First, suppose that $\underline{I}_k = -\infty$, then every $-\infty < I_t < I_{u,k+1}$ can be reached with a positive probability. Hence, if there exists I_t such that

$$E_t^{\varepsilon}\left[e^{-\rho\tau}\frac{\kappa}{\rho} \middle| I_t \right] < E_t^{\hat{\varepsilon}}\left[\int_t^{\tau} e^{-\rho(s-t)}\phi ds + e^{-\rho\tau}\frac{\kappa}{\rho} \middle| I_t \right],$$

where $\varepsilon = \{a_s = \mu\}_{t \le s \le \tau}$ and $\hat{\varepsilon} = \{a_s = 0\}_{t \le s \le \tau}$, then the proof is done. The difference between them is

$$E_t^{\varepsilon}\left[e^{-\rho\tau}\frac{\kappa}{\rho}\middle| I_t\right] - E_t^{\hat{\varepsilon}}\left[\int_t^{\tau} e^{-\rho(s-t)}\phi ds + e^{-\rho\tau}\frac{\kappa}{\rho}\middle| I_t\right] = \underbrace{e^{-\eta^+(\bar{I}-I_t)}\frac{\kappa}{\rho} - e^{-\frac{\sqrt{2\rho}}{\sigma}(\bar{I}-I_t)}\left(\frac{\kappa-\phi}{\rho}\right) - \frac{\phi}{\rho}}_{\equiv f(I_t)}.$$

Since $\lim_{I_t \to -\infty} f(I_t) = -\phi/\rho < 0$ and $f(\bar{I}) = 0$, there is I_t such that $f(I_t) < 0$.

B.1.4 Proof of Lemma 5

Since

$$\frac{\eta^- \exp(\eta^- I_t + \eta^+ \bar{I}) - \eta^+ \exp(\eta^- \bar{I} + \eta^+ I_t)}{\exp(\eta^- \underline{I} + \eta^+ \bar{I}) - \exp(\eta^- \bar{I} + \eta^+ \underline{I})}\mathcal{B}_d < 0$$

for $\mathcal{B}_d > 0$ and $\underline{I} \le I_t \le \bar{I}$,

$$\left.\frac{\partial W_t(I_t, \mathcal{B}_u, \mathcal{B}_d)}{\partial I_t}\right|_{\mathcal{B}_d=0} > \left.\frac{\partial W_t(I_t, \mathcal{B}_u, \mathcal{B}_d)}{\partial I_t}\right|_{\mathcal{B}_d>0}.$$

Therfore, by setting $\mathcal{B}_d = 0$, the principal can achieve the incentive compatibility condition with strictly lower $\mathcal{B}_u$ for every $\underline{I} \le I_t \le \bar{I}$.

B.1.5 Proof of Lemma 6

For given $\underline{I}$, the principal has to set

$$\mathcal{B}_u \ge \frac{\exp(\eta^- \underline{I} + \eta^+ \bar{I}) - \exp(\eta^- \bar{I} + \eta^+ \underline{I})}{\eta^+ \exp(\eta^- \underline{I} + \eta^+ I^*) - \eta^- \exp(\eta^- I^* + \eta^+ \underline{I})}\frac{\phi}{\mu},$$

where

$$I^* = \min\left[\frac{1}{\eta^+ - \eta^-}\ln\left(\frac{(\eta^-)^2}{(\eta^+)^2}\right) + \underline{I}, \bar{I}\right],$$

in order to satisfy the incentive compatibility condition for all $\underline{I} \le I_t \le \bar{I}$. Notice that

$$\frac{\partial^2 W}{\partial I^2} = \frac{(\eta^+)^2 \exp(\eta^- \underline{I} + \eta^+ I) - (\eta^-)^2 \exp(\eta^- I + \eta^+ \underline{I})}{\exp(\eta^- \underline{I} + \eta^+ \bar{I}) - \exp(\eta^- \bar{I} + \eta^+ \underline{I})}\mathcal{B}_u.$$

This second derivative is equal to zero if

$$I = \frac{1}{\eta^+ - \eta^-}\ln\left(\frac{(\eta^-)^2}{(\eta^+)^2}\right) + \underline{I} \equiv I^{**}.$$

Also, the second derivative is strictly greater (less) than 0 if $I > (<) I^{**}$.

Since the agent's expected discounted profit

$$E\left[e^{-r\tau}\left(\frac{\kappa}{r}\mathbb{1}_{\{\tau=\tau_u\}} - \mathcal{B}_u\mathbb{1}_{\{\tau=\tau_u\}}\right)\right] = \frac{\exp(\nu^-\underline{I} + \nu^+ I_0) - \exp(\nu^- I_0 + \nu^+\underline{I})}{\exp(\nu^-\underline{I} + \nu^+\bar{I}) - \exp(\nu^-\bar{I} + \nu^+\underline{I})}\left(\frac{\kappa}{r} - \mathcal{B}_u\right)$$

is a strictly decreasing function in $\mathcal{B}_u$, it is optimal for him to set $\mathcal{B}_u$ as the minimum value in the set of $\mathcal{B}_u$'s satisfying the incentive compatibility condition.

B.1.6 Proof of Proposition 20

Recall that the principal's problem is

$$\max_{\underline{I}} \mathcal{P}(\underline{I})\left(\frac{\kappa}{r} - \mathcal{B}_u(\underline{I})\right),$$

where

$$\mathcal{P}(\underline{I}) = \frac{\exp(\nu^-\underline{I} + \nu^+ I_0) - \exp(\nu^- I_0 + \nu^+\underline{I})}{\exp(\nu^-\underline{I} + \nu^+\bar{I}) - \exp(\nu^-\bar{I} + \nu^+\underline{I})}, \text{ and}$$

$$\mathcal{B}_u(\underline{I}) = \frac{\exp(\eta^-\underline{I} + \eta^+\bar{I}) - \exp(\eta^-\bar{I} + \eta^+\underline{I})}{\eta^+\exp(\eta^-\underline{I} + \eta^+ I^*) - \eta^-\exp(\eta^- I^* + \eta^+\underline{I})}\frac{\phi}{\mu}.$$

First, I show that there is a unique I^{**} such that if $\underline{I} < I^{**}$, the principal's discounted expected profit is less than 0. There are two possibilities. When $I^* = \frac{1}{\eta^+-\eta^-}\ln\left(\frac{(\eta^-)^2}{(\eta^+)^2}\right) + \underline{I}$, I can rewrite $\mathcal{B}_u(\underline{I})$ by

$$\mathcal{B}_u(\underline{I}) = \frac{\exp(\eta^+(\bar{I} - \underline{I})) - \exp(\eta^-(\bar{I} - \underline{I}))}{(\eta^- - \eta^+)\frac{\eta^-}{\eta^+}\left(\frac{(\eta^-)^2}{(\eta^+)^2}\right)^{\frac{\eta^-}{\eta^+-\eta^-}}}\frac{\phi}{\mu}.$$

Notice that this is a strictly decreasing function in $\underline{I}$ and it is equal to zero when $\underline{I} = \bar{I}$. Therefore, there is a unique I^{**} such that $\mathcal{B}_u(I^{**}) = \frac{\kappa}{r}$. Therefore, choosing $\underline{I}$ lower than I^{**} gives a negative profit to the principal. On the other hand, if $I^* = \bar{I}$,

$$\mathcal{B}_u(\underline{I}) = \frac{1 - \exp(-(\eta^+ - \eta^-)(\bar{I} - \underline{I}))}{\eta^+ - \eta^-\exp(-(\eta^+ - \eta^-)(\bar{I} - \underline{I}))}\frac{\phi}{\mu}$$

is a strictly decreasing function in $\underline{I}$ and $\mathcal{B}_u = 0$ when $\bar{I} = \underline{I}$. Therefore, $\mathcal{B}_u$ has the maximum value when

$$\underline{I} = \bar{I} - \frac{1}{\eta^+ - \eta^-}\ln\left(\frac{(\eta^-)^2}{(\eta^+)^2}\right).$$

After some algebra, I can obtain the maximum value of $\mathcal{B}_u(\underline{I})$

$$\mathcal{B}_u = \frac{\phi}{\rho}.$$

Hence, when $I^* = \bar{I}$, $\mathcal{B}_u(\underline{I})$ is always strictly less than $\frac{\kappa}{r}$ since I assume that $\kappa > \phi$ and $\rho > r$.

Second, $\mathcal{B}_u(\underline{I})$ must be greater than $\mathcal{B}_u(I_0)$ since $\underline{I}$ can't be greater than I_0 and $\mathcal{B}_u(\underline{I})$ is a strictly decreasing function in $\underline{I}$. From now on, I focus on the range $I^{**} \leq \underline{I} \leq I_0$ without loss of generality.

There are three possible cases.

1. $\left(\frac{\kappa}{r} - \mathcal{B}_u(I_0) \leq 0\right)$

   In this case, it is impossible for the principal to achieve a positive utility regardless of the choice of $\underline{I}$ since

$$\frac{\partial \mathcal{B}_u(\underline{I})}{\partial \underline{I}} = \begin{cases} -\frac{\phi}{\mu} \frac{\eta^+ \exp(\eta^- \underline{I} + \eta^+ \bar{I}) - \eta^- \exp(\eta^- \bar{I} + \eta^+ \underline{I})}{\eta^+ \exp(\eta^- \underline{I} + \eta^+ I^*) - \eta^- \exp(\eta^- I^* + \eta^+ \underline{I})} < 0 & \text{if } I^* \neq \bar{I} \\ -\frac{\phi}{\mu} e^{\left(-\frac{2\mu}{\sigma^2}(\bar{I} + \underline{I})\right)} \frac{4(\mu^2 + 2\rho\sigma^2)}{\sigma^4} \frac{1}{[\eta^+ \exp(\eta^- \underline{I} + \eta^+ \bar{I}) - \eta^- \exp(\eta^- \bar{I} + \eta^+ \underline{I})]^2} < 0 & \text{if } I^* = \bar{I} \end{cases}.$$

2. $\left(\frac{\kappa}{r} - \mathcal{B}_u(I_0) > 0 \text{ and } \bar{I} - \frac{1}{\eta^+ - \eta^-} \ln\left(\frac{(\eta^-)^2}{(\eta^+)^2}\right) \geq I_0\right)$

   In this case, $I^* = \underline{I} + \frac{1}{\eta^+ - \eta^-} \ln\left(\frac{(\eta^-)^2}{(\eta^+)^2}\right)$ for all $\underline{I} \in [I^{**}, I_0]$. Under this circumstance, there is a unique $\underline{I}$ maximizing the principal's discounted expected profit since

$$\frac{\partial U_0(\underline{I})}{\partial \underline{I}} = \mathcal{P}'(\underline{I}) \left(\frac{\kappa}{r} - \mathcal{B}_u(\underline{I})\right) - \mathcal{P}(\underline{I}) \frac{\partial \mathcal{B}_u(\underline{I})}{\partial \underline{I}} = \begin{cases} \mathcal{P}'(I_0) \left(\frac{\kappa}{r} - \mathcal{B}_u(I_0)\right) < 0 & \text{if } \underline{I} = I_0 \\ -\mathcal{P}(I^{**}) \left.\frac{\partial \mathcal{B}_u(\underline{I})}{\partial \underline{I}}\right|_{\underline{I} = I^{**}} > 0 & \text{if } \underline{I} = I^{**} \end{cases}$$

   and

$$\frac{\partial^2 U_0(\underline{I})}{\partial \underline{I}^2} = \mathcal{P}''(\underline{I}) \left(\frac{\kappa}{r} - \mathcal{B}_u(\underline{I})\right) - 2\mathcal{P}'(\underline{I}) \frac{\partial \mathcal{B}_u(\underline{I})}{\partial \underline{I}} - \mathcal{P}(\underline{I}) \frac{\partial^2 \mathcal{B}_u(\underline{I})}{\partial \underline{I}^2} < 0$$

   for $I^{**} < \underline{I} < I_0$, where

$$U_0(\underline{I}) = \mathcal{P}(\underline{I}) \left(\frac{\kappa}{r} - \mathcal{B}_u(\underline{I})\right),$$

$$\mathcal{P}'(\underline{I}) = -\frac{2\sqrt{\mu^2 + 2r\sigma^2}}{\sigma^2} \exp\left(-\frac{2\mu}{\sigma^2}\underline{I}\right) \frac{\exp(\nu^- I_0 + \nu^+ \bar{I}) - \exp(\nu^- \bar{I} + \nu^+ I_0)}{[\exp(\nu^- \underline{I} + \nu^+ \bar{I}) - \exp(\nu^- \bar{I} + \nu^+ \underline{I})]^2},$$

$$\mathcal{P}''(\underline{I}) = -4\frac{\mu^2 + 2r\sigma^2}{\sigma^4} \exp\left(-\frac{2\mu}{\sigma^2}\underline{I}\right) \frac{\exp(\nu^- \underline{I} + \nu^+ \bar{I}) + \exp(\nu^- \bar{I} + \nu^+ \underline{I})}{[\exp(\nu^- \underline{I} + \nu^+ \bar{I}) - \exp(\nu^- \bar{I} + \nu^+ \underline{I})]^3} \cdot \left(\exp(\nu^- I_0 + \nu^+ \bar{I}) - \exp(\nu^- \bar{I} + \nu^+ I_0)\right), \text{ and}$$

$$\frac{\partial^2 \mathcal{B}_u(\underline{I})}{\partial \underline{I}^2} = -\frac{\phi}{\mu} \frac{(\eta^-)^2 \exp(\eta^- \bar{I} + \eta^+ \underline{I}) - (\eta^+)^2 \exp(\eta^- \underline{I} + \eta^+ \bar{I})}{\eta^+ \exp(\eta^- \underline{I} + \eta^+ I^*) - \eta^- \exp(\eta^- I^* + \eta^+ \underline{I})}.$$

3. $\left(\frac{\kappa}{r} - \mathcal{B}_u(I_0) > 0 \text{ and } \bar{I} - \frac{1}{\eta^+ - \eta^-}\ln\left(\frac{(\eta^-)^2}{(\eta^+)^2}\right) < I_0\right)$

Now, I^* could be $\underline{I} + \frac{1}{\eta^+ - \eta^-}\ln\left(\frac{(\eta^-)^2}{(\eta^+)^2}\right)$ or $\bar{I}$ depending on $\underline{I}$. In order to analyze this situation, consider the following Lemma.

Lemma 7 *Denote* $\frac{1}{\eta^+ - \eta^-}\ln\left(\frac{(\eta^-)^2}{(\eta^+)^2}\right)$ *by* $\mathcal{T}$. *When* $\frac{\kappa}{r} - \mathcal{B}_u(I_0) > 0$, *if there exists* $\tilde{I} \in (\bar{I} - \mathcal{T}, I_0)$ *such that*

$$\left.\frac{\partial U_0(\underline{I})}{\partial \underline{I}}\right|_{\underline{I}=\tilde{I}} = 0,$$

then

$$\left.\frac{\partial^2 U_0(\underline{I})}{\partial \underline{I}^2}\right|_{\underline{I}=\tilde{I}} < 0.$$

Proof 11 *By the condition* $\left.\frac{\partial U_0(\underline{I})}{\partial \underline{I}}\right|_{\underline{I}=\tilde{I}} = 0$,

$$\mathcal{P}'(\tilde{I})\left(\frac{\kappa}{r} - \mathcal{B}_u(\tilde{I})\right) - \mathcal{P}(\tilde{I})\mathcal{B}_u'(\tilde{I}) = 0.$$

Therefore,

$$\begin{aligned}
\left.\frac{\partial^2 U_0(\underline{I})}{\partial \underline{I}^2}\right|_{\underline{I}=\tilde{I}} &= \mathcal{P}''(\tilde{I})\left(\frac{\kappa}{r} - \mathcal{B}_u(\tilde{I})\right) - 2\mathcal{P}'(\tilde{I})\mathcal{B}_u'(\tilde{I}) - \mathcal{P}(\tilde{I})\mathcal{B}_u''(\tilde{I}) \\
&= \mathcal{P}''\frac{\mathcal{P}(\tilde{I})\mathcal{B}_u'(\tilde{I})}{\mathcal{P}'(\tilde{I})} - 2\mathcal{P}'(\tilde{I})\mathcal{B}_u'(\tilde{I}) - \mathcal{P}(\tilde{I})\mathcal{B}_u''(\tilde{I}) \\
&= -2\mathcal{P}(\tilde{I})\mathcal{B}_u'(\tilde{I})\frac{1}{\sigma^2}\left(\sqrt{\mu^2 + 2r\sigma^2}\frac{1 + \exp((\nu^+ - \nu^-)(I_0 - \tilde{I}))}{1 - \exp((\nu^+ - \nu^-)(I_0 - \tilde{I}))}\right. \\
&\qquad \left. + \sqrt{\mu^2 + 2\rho\sigma^2}\frac{\eta^+ + \eta^- \exp(-(\eta^+ - \eta^-)(\bar{I} - \tilde{I}))}{\eta^+ - \eta^- \exp(-(\eta^+ - \eta^-)(\bar{I} - \tilde{I}))}\right).
\end{aligned}$$

I need to show that

$$\begin{aligned}
&\sqrt{\mu^2 + 2r\sigma^2}\frac{1 + \exp((\nu^+ - \nu^-)(I_0 - \tilde{I}))}{1 - \exp((\nu^+ - \nu^-)(I_0 - \tilde{I}))} \\
&\qquad + \sqrt{\mu^2 + 2\rho\sigma^2}\frac{\eta^+ + \eta^- \exp(-(\eta^+ - \eta^-)(\bar{I} - \tilde{I}))}{\eta^+ - \eta^- \exp(-(\eta^+ - \eta^-)(\bar{I} - \tilde{I}))} < 0.
\end{aligned}$$

Since the LHS is a strictly decreasing function in $\tilde{I}$. *it is enough to show that*

$$\begin{aligned}
&\sqrt{\mu^2 + 2r\sigma^2}\frac{1 + \exp((\nu^+ - \nu^-)(I_0 - \underline{I} + \mathcal{T}))}{1 - \exp((\nu^+ - \nu^-)(I_0 - \underline{I} + \mathcal{T}))} \\
&\qquad + \sqrt{\mu^2 + 2\rho\sigma^2}\frac{\eta^+ + \eta^- \exp(-(\eta^+ - \eta^-)\mathcal{T})}{\eta^+ - \eta^- \exp(-(\eta^+ - \eta^-)\mathcal{T})} \\
&\qquad = \sqrt{\mu^2 + 2r\sigma^2}\frac{1 + \exp((\nu^+ - \nu^-)(I_0 - \underline{I} + \mathcal{T}))}{1 - \exp((\nu^+ - \nu^-)(I_0 - \underline{I} + \mathcal{T}))} + \mu \\
&\qquad < 0,
\end{aligned}$$

where the last inequality holds since

$$\sqrt{\mu^2 + 2r\sigma^2} > \mu \text{ and } \frac{1 + \exp((\nu^+ - \nu^-)(I_0 - \underline{I} + \mathcal{T}))}{1 - \exp((\nu^+ - \nu^-)(I_0 - \underline{I} + \mathcal{T}))} < -1.$$

Since $\frac{\partial U_0(\underline{I})}{\partial \underline{I}}\Big|_{\underline{I}=I_0} = \mathcal{P}'(I_0)\left(\frac{\kappa}{r} - \mathcal{B}_u(I_0)\right) < 0$ and $U_0(\underline{I})$ is a strict convex function for $\underline{I} \in [I^{**}, \bar{I} - \mathcal{T}]$, if $\frac{\partial U_0(\underline{I})}{\partial \underline{I}}\Big|_{\underline{I}=I_0} \leq 0$, there exists a unique $\underline{I} \in [I^{**}, \bar{I} - \mathcal{T}]$ maximizing the principal's profit. Note that $\frac{\partial U_0(\underline{I})}{\partial \underline{I}}\Big|_{\underline{I}=I_0} \leq 0$ implies $\frac{\partial U_0(\underline{I})}{\partial \underline{I}} < 0$ for $\underline{I} \in (\bar{I} - \mathcal{T}, I_0]$ by **Lemma 7**. Also, if $\frac{\partial U_0(\underline{I})}{\partial \underline{I}}\Big|_{\underline{I}=I_0} > 0$, there exists a unique $\underline{I} \in (\bar{I} - \mathcal{T}, I_0]$ maximizing the principal's profit by **Lemma 7**.

B.1.7 Proof of Corollary 8

Denote the project level right after the final intermediate payment by I_t. Then, the bonus payment $\mathcal{B}_u$ must be less than $\mathcal{B}_u^*$ which is the optimal bonus payment without any intermediate payment. (If not, the compensation scheme with intermedate payments makes the principal worse off.) Recall that there is a $I^* \in (\underline{I}, \bar{I})$ such that $\frac{\partial W(I, \mathcal{B}_u^*, \mathcal{A})}{\partial I}\Big|_{I=I^*} = \frac{\phi}{\mu}$. Hence, $\frac{\partial W(I, \mathcal{B}_u, \mathcal{A})}{\partial I}\Big|_{I=I^*} < \frac{\phi}{\mu}$, where $\mathcal{A} = \{a_t = \mu\}_{0 \leq t < \tau}$. Since I^* is reached with a positive probability, the compensation scheme with intermediate payments is not incentive-compatible.

B.1.8 Proof of Comparative Statics

Lemma 8 *For given parameters, the optimal termination level $\underline{I}^*$ is a strictly increasing function in I_0. That is,*

$$\frac{\partial \underline{I}^*}{\partial I_0} > 0.$$

Proof 12 *Recall that*

$$\frac{\partial U_0(\underline{I})}{\partial \underline{I}}\Big|_{\underline{I}=I^*} = \mathcal{P}'(\underline{I}^*)\left(\frac{\kappa}{r} - \mathcal{B}_u(\underline{I}^*)\right) - \mathcal{P}(\underline{I}^*)\mathcal{B}_u'(\underline{I}^*) = 0.$$

By the implicit function theorem,

$$\frac{\partial \underline{I}^*}{\partial I_0} = -\left(\frac{\partial^2 U_0(\underline{I})}{\partial \underline{I}^2}\Big|_{\underline{I}=\underline{I}^*}\right)^{-1}\left[\frac{\partial \mathcal{P}'(\underline{I}^*)}{\partial I_0}\left(\frac{\kappa}{r} - \mathcal{B}_u(\underline{I}^*)\right) - \frac{\partial \mathcal{P}(\underline{I}^*)}{\partial I_0}\mathcal{B}_u'(\underline{I}^*)\right] > 0$$

since

$$\frac{\partial \mathcal{P}(\underline{I}^*)}{\partial I_0} = \frac{\nu^+ \exp(\nu^- \underline{I}^* + \nu^+ I_0) - \nu^- \exp(\nu^- I_0 + \nu^+ \underline{I}^*)}{\exp(\nu^- \underline{I}^* + \nu^+ \bar{I}) - \exp(\nu^- \bar{I} + \nu^+ \underline{I}^*)} > 0, \text{ and}$$

$$\frac{\partial \mathcal{P}'(\underline{I}^*)}{\partial I_0} = -2\frac{\sqrt{\mu^2 + 2r\sigma^2}}{\sigma^2} \exp\left(-\frac{2\mu}{\sigma^2}\underline{I}\right) \frac{\nu^- \exp(\nu^- I_0 + \nu^+ \bar{I}) - \nu^+ \exp(\nu^- \bar{I} + \nu^+ I_0)}{(\exp(\nu^- \underline{I}^* + \nu^+ \bar{I}) - \exp(\nu^- \bar{I} + \nu^+ \underline{I}^*))^2} > 0.$$

Note that this implies that

$$\frac{\partial \mathcal{B}_u(\underline{I}^*)}{\partial I_0} < 0$$

since $\mathcal{B}_u(\underline{I})$ is a strictly decreasing function in $\underline{I}$.

Lemma 9 *For given parameters, the optimal termination time $\underline{I}^*$ is a strictly increasing function in ϕ. That is,*

$$\frac{\partial \underline{I}^*}{\partial \phi} > 0.$$

Proof 13 *Again, by the implicit function theorem,*

$$\frac{\partial \underline{I}^*}{\partial \phi} = \left(\left.\frac{\partial^2 U_0(\underline{I})}{\partial \underline{I}^2}\right|_{\underline{I}=\underline{I}^*}\right)^{-1} \left[\mathcal{P}'(\underline{I}^*)\frac{\partial \mathcal{B}_u(\underline{I}^*)}{\partial \phi} + \mathcal{P}(\underline{I}^*)\frac{\partial \mathcal{B}_u'(\underline{I}^*)}{\partial \phi}\right] > 0$$

since

$$\frac{\partial \mathcal{B}_u(\underline{I}^*)}{\partial \phi} = \frac{\mathcal{B}_u(\underline{I}^*)}{\phi} > 0$$
$$\frac{\partial \mathcal{B}_u'(\underline{I}^*)}{\partial \phi} = \frac{\mathcal{B}_u'(\underline{I}^*)}{\phi} < 0.$$

Lemma 10 *For given parameters, the principal's discounted expected utility is a strictly increasing function in I_0. That is,*

$$\frac{\partial U_0(\underline{I}^*)}{\partial I_0} > 0.$$

Proof 14 *By the envelope theorem,*

$$\frac{\partial U_0(\underline{I}^*)}{\partial I_0} = \frac{\partial \mathcal{P}(\underline{I}^*)}{\partial I_0}\left(\frac{\kappa}{r} - \mathcal{B}_u(\underline{I}^*)\right) - \mathcal{P}(\underline{I}^*)\frac{\partial \mathcal{B}_u(\underline{I}^*)}{\partial I_0} > 0.$$

*The inequality holds by **Lemma 8**.*

Lemma 11 *Consider the case where* $I^* = \frac{1}{\eta^+ - \eta^-}\log\left(\frac{(\eta^-)^2}{(\eta^+)^2}\right) + \underline{I}^* < \bar{I}$. *Under this circumstance, the agent's discounted expected utility at time 0 is a strictly increasing (decreasing) function in* I_0 *if*

$$\frac{\partial \underline{I}^*}{\partial I_0} < (>) 1.$$

On the other hand, if $I^* = \bar{I}$*, the agent's discounted expected utility at time 0 is a strictly increasing (decreasing) function in* I_0 *if*

$$\frac{\partial \underline{I}^*}{\partial I_0} < (>) \frac{(\eta^- e^{\eta^- I_0 + \eta^+ \underline{I}^*} - \eta^+ e^{\eta^- \underline{I}^* + \eta^+ I_0})(\eta^+ e^{\eta^- \underline{I}^* + \eta^+ \bar{I}} - \eta^- e^{\eta^- \bar{I} + \eta^+ \underline{I}^*})}{(\eta^+ - \eta^-)(\eta^- e^{\eta^-(\bar{I} + \underline{I}^*) + \eta^+(I_0 + \underline{I}^*)} - \eta^+ e^{\eta^-(I_0 + \underline{I}^*) + \eta^+(\bar{I} + \underline{I}^*)})} \leq 1.$$

Proof 15 *Note that when* $I^* = \frac{1}{\eta^+ - \eta^-}\log\left(\frac{(\eta^-)^2}{(\eta^+)^2}\right) + \underline{I}^* < \bar{I}$,

$$\mathcal{B}_u(\underline{I}^*) = \frac{\exp(\eta^- \underline{I} + \eta^+ \bar{I}) - \exp(\eta^- \bar{I} + \eta^+ \underline{I})}{\eta^+ \exp(\eta^- \underline{I} + \eta^+ I^*) - \eta^- \exp(\eta^- I^* + \eta^+ \underline{I})} \frac{\phi}{\mu}.$$

Therefore, the agent's discounted expected utility at time 0 is

$$W_0(I_0) = \frac{\exp(\eta^- \underline{I}^* + \eta^+ I_0) - \exp(\eta^- I_0 + \eta^+ \underline{I}^*)}{\eta^+ \exp(\eta^- \underline{I}^* + \eta^+ I^*) - \eta^- \exp(\eta^- I^* + \eta^+ \underline{I}^*)} \frac{\phi}{\mu}.$$

Differentiation this with respect to I_0 *yields*

$$\frac{\partial W_0(I_0)}{\partial I_0} = \underbrace{\frac{-\eta^+ \exp(\eta^- \underline{I}^* + \eta^+ I_0) + \eta^- \exp(\eta^- I_0 + \eta^+ \underline{I}^*)}{\eta^+ \exp(\eta^- \underline{I}^* + \eta^+ I^*) - \eta^- \exp(\eta^- I^* + \eta^+ \underline{I}^*)}}_{<0} \left(\frac{\partial \underline{I}^*}{\partial I_0} - 1\right) \frac{\phi}{\mu}.$$

Hence,

$$\frac{\partial W_0(I_0)}{\partial I_0} > (<) 0 \text{ if and only if } \frac{\partial \underline{I}^*}{\partial I_0} < (>) 1.$$

When $I^* = \underline{I}$,

$$\frac{\partial W_0(I_0)}{\partial I_0} = \frac{1}{\eta^+ e^{\eta^- \underline{I}^* + \eta^+ \bar{I}} - \eta^- e^{\eta^- \bar{I} + \eta^+ \underline{I}^*}} \frac{\phi}{\mu} \cdot \left[\frac{(\eta^+ - \eta^-)(\eta^- e^{\eta^-(\bar{I} + \underline{I}^*) + \eta^+(I_0 + \underline{I}^*)} - \eta^+ e^{\eta^-(I_0 + \underline{I}^*) + \eta^+(\bar{I} + \underline{I}^*)})}{\eta^+ \exp(\eta^- \underline{I}^* + \eta^+ \bar{I}) - \eta^- \exp(\eta^- \bar{I} + \eta^+ \underline{I}^*)} \frac{\partial \underline{I}^*}{\partial I_0} + \eta^+ \exp(\eta^- \underline{I}^* + \eta^+ I_0) - \eta^- \exp(\eta^- I_0 + \eta^+ \underline{I}^*)\right].$$

Hence, $\partial W_0(I_0)/\partial I_0 > 0$ *if the term in the square bracket is positive, and this holds if the condition in Lemma holds. Also, the difference between the numerator and denominator in the condition in Lemma is :*

Numerator−Denominator =

$$(e^{\eta^- \underline{I}^* + \eta^+ I_0} - e^{\eta^- I_0 + \eta^+ \underline{I}^*}) e^{\eta^- \underline{I}^* + \eta^+ \bar{I}} [(\eta^-)^2 e^{-(\eta^+ - \eta^-)(\bar{I} - \underline{I}^*)} - (\eta^+)^2].$$

Since $\bar{I} - \underline{I}^* < \frac{1}{\eta^+ - \eta^-} \log\left(\frac{(\eta^-)^2}{(\eta^+)^2}\right)$, *the difference is greater than 0. Hence, the threshold is less than 1. (Note that both numerator and denominator are negative.)*

Lemma 12 *For a given parameters, the principal's discounted expected utility is a strictly decreasing function in* ϕ. *That is,*

$$\frac{\partial U_0(\underline{I}^*)}{\partial \phi} < 0.$$

Proof 16 *By the envelope theorem,*

$$\begin{aligned}\frac{\partial U_0(\underline{I}^*)}{\partial \phi} &= -\mathcal{P}(\underline{I}^*)\frac{\partial \mathcal{B}_u(\underline{I}^*)}{\partial \phi} \\ &= -\mathcal{P}(\underline{I}^*)\frac{\exp(\eta^-\underline{I} + \eta^+\bar{I}) - \exp(\eta^-\bar{I} + \eta^+\underline{I})}{\eta^+\exp(\eta^-\underline{I} + \eta^+ I^*) - \eta^-\exp(\eta^- I^* + \eta^+\underline{I})}\frac{1}{\mu} < 0.\end{aligned}$$

Lemma 13 *The agent's discounted expected utility at time* 0 *can be increasing or decreasing in* ϕ, *depending on the parameter values.*

Proof 17 *Note that if* $\phi = 0$, *the agent's discounted expected utility is 0 since* $\mathcal{B}_u(\underline{I}) = 0$ *regardless of the value of* $\underline{I}$. *On the other hand, there is* $\bar{\phi}$ *such that* $\mathcal{B}_u(I_0) > \frac{\kappa}{r}$ *if* $\phi \geq \bar{\phi}$ *since* $\mathcal{B}_u(I_0)$ *is a strictly increasing function in* ϕ. *Hence, the agent's discounted expected utility is 0 if* $\phi \geq \bar{\phi}$. *That is, the principal does not hire the agent or sets* $\underline{I} = I_0$. *Combining this with the limited liability gives us the desired result.*

Lemma 14 *The completion probability can be increasing or decreasing in* ϕ, *depending on the parameter values.*

Proof 18 *Note that*

$$\frac{\partial P}{\partial I_0} = \frac{1}{e^{-\delta\underline{I}^*} - e^{-\delta\bar{I}}}\left(\delta e^{-\delta I_0} + \delta e^{-\delta\underline{I}^*}\frac{e^{-\delta\bar{I}} - e^{-\delta I_0}}{e^{-\delta\underline{I}^*} - e^{-\delta\bar{I}}}\frac{\partial\underline{I}^*}{\partial I_0}\right).$$

Therefore,

$$\frac{\partial P}{\partial I_0} \geq 0$$

if and only if

$$\frac{\partial \underline{I}^*}{\partial I_0} \leq \frac{1 - e^{-\delta(\bar{I} - \underline{I}^*)}}{1 - e^{-\delta(\bar{I} - I_0)}}.$$

Lemma 15 *The completion probaility is a strictly decreasing function in ϕ. That is,*

$$\frac{\partial P}{\partial \phi} < 0.$$

Proof 19 *This holds since $\frac{\partial \underline{I}^*}{\partial \phi} > 0$ and*

$$\frac{\partial P}{\partial \phi} = \delta e^{-\delta \underline{I}^*} \frac{e^{-\delta \bar{I}} - e^{-\delta I_0}}{[e^{-\delta \underline{I}^*} - e^{-\delta \bar{I}}]^2} \frac{\partial \underline{I}^*}{\partial \phi} < 0.$$

B.1.9 Proof of Proposition 21

Note that the principal's problem is

$$\begin{aligned}
\max_{(\widehat{I},\underline{I}_1,\underline{I}_2,C,\mathcal{B})} \ & \mathcal{P}(I_0,\widehat{I},\underline{I}_1)\left[-C + \widetilde{\mathcal{P}}(\widehat{I},\bar{I},\underline{I}_2)\left(\frac{\kappa}{r} - \mathcal{B}\right)\right] \\
\text{subject to } \ & \frac{\partial \mathcal{P}_A(I_t,\widehat{I},\underline{I}_1)}{\partial I_t}(C + \widetilde{\mathcal{P}}_A(\widehat{I},\bar{I},\underline{I}_2)\mathcal{B}) \geq \frac{\phi}{\mu} \text{ for } I_t \in [\underline{I}_1,\widehat{I}], \\
& \frac{\partial \widetilde{\mathcal{P}}_A(I_t,\bar{I},\underline{I}_2)}{\partial I_t}\mathcal{B} \geq \frac{\phi}{\mu} \text{ for } I_t \in [\underline{I}_2,\bar{I}], \text{ and} \\
& C \geq 0,
\end{aligned}$$

where

$$\begin{aligned}
\mathcal{P}(I,\widehat{I},\underline{I}_1) &= \frac{\exp(\nu^- \underline{I}_1 + \nu^+ I) - \exp(\nu^- I + \nu^+ \underline{I}_1)}{\exp(\nu^- \underline{I}_1 + \nu^+ \widehat{I}) - \exp(\nu^- \widehat{I} + \nu^+ \underline{I}_1)}, \\
\widetilde{\mathcal{P}}(I,\bar{I},\underline{I}_2) &= \frac{\exp(\nu^- \underline{I}_2 + \nu^+ I) - \exp(\nu^- I + \nu^+ \underline{I}_2)}{\exp(\nu^- \underline{I}_2 + \nu^+ \bar{I}) - \exp(\nu^- \bar{I} + \nu^+ \underline{I}_2)}, \\
\mathcal{P}_A(I,\widehat{I},\underline{I}_1) &= \frac{\exp(\eta^- \underline{I}_1 + \eta^+ I) - \exp(\eta^- I + \eta^+ \underline{I}_1)}{\exp(\eta^- \underline{I}_1 + \eta^+ \widehat{I}) - \exp(\eta^- \widehat{I} + \eta^+ \underline{I}_1)}, \text{ and} \\
\widetilde{\mathcal{P}}_A(I,\bar{I},\underline{I}_2) &= \frac{\exp(\eta^- \underline{I}_2 + \eta^+ I) - \exp(\eta^- I + \eta^+ \underline{I}_2)}{\exp(\eta^- \underline{I}_2 + \eta^+ \bar{I}) - \exp(\eta^- \bar{I} + \eta^+ \underline{I}_2)}.
\end{aligned}$$

Claim 8

$$\frac{\partial \widetilde{\mathcal{P}}_A(I_t,\bar{I},\underline{I}_2)}{\partial I_t}\mathcal{B} = \frac{\phi}{\mu} \textit{ for some } I_t \in [\underline{I}_2,\bar{I}], \textit{ and}$$

$$\frac{\partial \mathcal{P}_A(I_t,\widehat{I},\underline{I}_1)}{\partial I_t}(C + \widetilde{\mathcal{P}}_A(\widehat{I},\bar{I},\underline{I}_2)\mathcal{B}) = \frac{\phi}{\mu} \textit{ for some } I_t \in [\underline{I}_1,\widehat{I}].$$

Proof 20 *First, suppose that both conditions do not bind. Then, the principal can obtain a higher profit by reducing $\mathcal{B}$ since*

$$\frac{\partial \widetilde{\mathcal{P}}_A(I_t, \bar{I}, \underline{I}_2)}{\partial I_t} > 0 \textit{ and } \frac{\partial \mathcal{P}_A(I_t, \widehat{I}, \underline{I}_1)}{\partial I_t} > 0.$$

Let's assume that the first condition binds but the second does not. There are two possibilities, $C > 0$ and $C = 0$. If $C > 0$, the principal increases his profit by reducing C. On the other hand, if $C = 0$, lowering $\underline{I}_1$ by $\epsilon > 0$ yields a higher profit to the principal since

$$\frac{\partial \mathcal{P}(I_0, \widehat{I}, \underline{I}_1)}{\partial \underline{I}_1} < 0 \textit{ and } \frac{\partial^2 \mathcal{P}(I_t, \bar{I}, \underline{I}_1)}{\partial I_t \partial \underline{I}_1} > 0.$$

Therefore, the second condition must bind.

Suppose that the first condition does not bind but the second does. Then, choosing

$$\widehat{\mathcal{B}} = \mathcal{B} - \Delta, \textit{ and } \widehat{C} = C + \widetilde{\mathcal{P}}_A(\widehat{I}, \bar{I}, \underline{I}_2)\Delta$$

does not affect the first constraint and still satisfies the second constraint, where

$$\min_{I_t \in [\underline{I}_2, \bar{I}]} \frac{\partial \widetilde{\mathcal{P}}_A(I_t, \bar{I}, \underline{I}_2)}{\partial I_t}\mathcal{B} - \frac{\phi}{\mu} \equiv \Delta > 0.$$

This compensation scheme gives a higher profit to the principal since

$$\begin{aligned} -\mathcal{P}(I_0, \widehat{I}, \underline{I}_1)\widetilde{\mathcal{P}}_A(\widehat{I}, \bar{I}, \underline{I}_2)\Delta + \mathcal{P}(I_0, \widehat{I}, \underline{I}_1)\widetilde{\mathcal{P}}(\widehat{I}, \bar{I}, \underline{I}_2)\Delta \\ = \mathcal{P}(I_0, \widehat{I}, \underline{I}_1)\Delta[\widetilde{\mathcal{P}}(\widehat{I}, \bar{I}, \underline{I}_2) - \widetilde{\mathcal{P}}_A(\widehat{I}, \bar{I}, \underline{I}_2)] \end{aligned}$$

is strictly greater than zero by ***Theorem 2****.*

This claim directly provides the following results.

Claim 9 *For given $(\widehat{I}, \underline{I}_1, \underline{I}_2)$, the optimal $\mathcal{B}$ is*

$$\mathcal{B}(\underline{I}_2) \equiv \frac{\exp(\eta^- \underline{I}_2 + \eta^+ \bar{I}) - \exp(\eta^- \bar{I} + \eta^+ \underline{I}_2)}{\eta^+ \exp(\eta^- \underline{I}_2 + \eta^+ I^{**}) - \eta^- \exp(\eta^- I^{**} + \eta^+ \underline{I}_2)} \frac{\phi}{\mu},$$

where

$$I^{**} = \min\left[\frac{1}{\eta^+ - \eta^-} \ln\left(\frac{(\eta^-)^2}{(\eta^+)^2}\right) + \underline{I}_2, \bar{I}\right],$$

and the optimal C is

$$C(\widehat{I},\underline{I}_1,\underline{I}_2) \equiv \max\left[\frac{\phi}{\mu}\frac{\exp(\eta^-\underline{I}_1+\eta^+\widehat{I})-\exp(\eta^-\widehat{I}+\eta^+\underline{I}_1)}{\eta^+\exp(\eta^-\underline{I}_1+\eta^+I^*)-\eta^-\exp(\eta^-I^*+\eta^+\underline{I}_1)}-\widetilde{\mathcal{P}}_A(\widehat{I},\bar{I},\underline{I}_2)\mathcal{B}(\underline{I}_2),0\right],$$

where

$$I^* = \min\left[\frac{1}{\eta^+-\eta^-}\ln\left(\frac{(\eta^-)^2}{(\eta^+)^2}\right)+\underline{I}_1,\widehat{I}\right].$$

Proof 21 *Convexity of* $\frac{\partial\widetilde{\mathcal{P}}_A(I_t,\bar{I},\underline{I}_2)}{\partial I_t}$ *and* $\frac{\partial\mathcal{P}_A(I_t,\widehat{I},\underline{I}_1)}{\partial I_t}$ *implies that*

$$\mathcal{B} \geq \frac{\exp(\eta^-\underline{I}_2+\eta^+\bar{I})-\exp(\eta^-\bar{I}+\eta^+\underline{I}_2)}{\eta^+\exp(\eta^-\underline{I}_2+\eta^+I^{**})-\eta^-\exp(\eta^-I^{**}+\eta^+\underline{I}_2)}\frac{\phi}{\mu}$$

and

$$C \geq \frac{\phi}{\mu}\frac{\exp(\eta^-\underline{I}_1+\eta^+\widehat{I})-\exp(\eta^-\widehat{I}+\eta^+\underline{I}_1)}{\eta^+\exp(\eta^-\underline{I}_1+\eta^+I^*)-\eta^-\exp(\eta^-I^*+\eta^+\underline{I}_1)}-\widetilde{\mathcal{P}}_A(\widehat{I},\bar{I},\underline{I}_2)\mathcal{B}(\underline{I}_2).$$

Since $C \geq 0$ *and the principal's profit is a strictly decreasing function in* C *and* $\mathcal{B}$*, the claim holds.*

Claim 10 $\underline{I}_2$ *is greater than* $\underline{I}_1$*.*

Proof 22 *Suppose not. That is,* $\underline{I}_1 > \underline{I}_2$*. This implies that* $C(\widehat{I},\underline{I}_1,\underline{I}_2) = 0$ *since*

$$\frac{\phi}{\mu}\frac{\exp(\eta^-\underline{I}_1+\eta^+\widehat{I})-\exp(\eta^-\widehat{I}+\eta^+\underline{I}_1)}{\eta^+\exp(\eta^-\underline{I}_1+\eta^+I^*)-\eta^-\exp(\eta^-I^*+\eta^+\underline{I}_1)} < \widetilde{\mathcal{P}}_A(\widehat{I},\bar{I},\underline{I}_2)\mathcal{B}(\underline{I}_2)$$

regardless of $\widehat{I}$*. Now, I show that the constraints can't bind if* $\underline{I}_1 > \underline{I}_2$*. Without loss of generality, assume that*

$$\mathcal{B} = \frac{\exp(\eta^-\underline{I}_2+\eta^+\bar{I})-\exp(\eta^-\bar{I}+\eta^+\underline{I}_2)}{\eta^+\exp(\eta^-\underline{I}_2+\eta^+I^{**})-\eta^-\exp(\eta^-I^{**}+\eta^+\underline{I}_2)}\frac{\phi}{\mu}.$$

Recall that the other constraint is

$$\frac{\partial\mathcal{P}_A(I_t,\widehat{I},\underline{I}_1)}{\partial I_t}\widetilde{\mathcal{P}}_A(\widehat{I},\bar{I},\underline{I}_2)\mathcal{B} \geq \frac{\phi}{\mu} \text{ for } I_t \in [\underline{I}_1,\widehat{I}].$$

This is equivalent to

$$\frac{\eta^+\exp(\eta^-\underline{I}_1+\eta^+I^*)-\eta^-\exp(\eta^-I^*+\eta^+\underline{I}_1)}{\eta^+\exp(\eta^-\underline{I}_2+\eta^+I^{**})-\eta^-\exp(\eta^-I^{**}+\eta^+\underline{I}_2)}\frac{\exp(\eta^-\underline{I}_2+\eta^+\widehat{I})-\exp(\eta^-\widehat{I}+\eta^+\underline{I}_2)}{\exp(\eta^-\underline{I}_1+\eta^+\widehat{I})-\exp(\eta^-\widehat{I}+\eta^+\underline{I}_1)} \geq 1.$$

This condition does not bind since the left hand side is a strictly increasing function in $\underline{I}_1$ *and equal to one if* $\underline{I}_1 = \underline{I} = 2$*. Therefore,* $\underline{I}_1 \leq \underline{I}_2$*.*

Claim 11 *There is* $\underline{I}^*$ *such that* $\underline{I}_1 \geq \underline{I}^*$.

Proof 23 *Since* $\frac{\partial \mathcal{B}(\underline{I}_2)}{\partial \underline{I}_2} < 0$ *and* $\lim_{\underline{I}_2 \to -\infty} \mathcal{B}(\underline{I}_2) = \infty$*, there is* $\underline{I}^{**}$ *such that* $\mathcal{B}(\underline{I}^{**}) = \frac{\kappa}{r}$*. Therefore,* $\underline{I}_2 \geq \underline{I}^{**}$*. For given* $\underline{I}^{**}$*, there exist* $\underline{I}^*$ *such that*

$$\frac{\phi}{\mu} \frac{\exp(\eta^- \underline{I}^* + \eta^+ \widehat{I}) - \exp(\eta^- \widehat{I} + \eta^+ \underline{I}^*)}{\eta^+ \exp(\eta^- \underline{I}^* + \eta^+ I^*) - \eta^- \exp(\eta^- I^* + \eta^+ \underline{I}^*)} - \widetilde{\mathcal{P}}_A(\widehat{I}, \bar{I}, \underline{I}^{**}) \mathcal{B}(\underline{I}^{**}) = \frac{\kappa}{r}$$

since

$$\frac{\exp(\eta^- \underline{I}_1 + \eta^+ \widehat{I}) - \exp(\eta^- \widehat{I} + \eta^+ \underline{I}_1)}{\eta^+ \exp(\eta^- \underline{I}_1 + \eta^+ I^*) - \eta^- \exp(\eta^- I^* + \eta^+ \underline{I}_1)}$$

is a strictly decreasing function in $\underline{I}_1$ *and goes to infinity as* $\underline{I}_1 \to -\infty$*. Hence, if* $\underline{I}_1 < \underline{I}^*$*, the principal never obtains a non-negative profit.*

Therefore, the principal's problem is

$$\max_{(\widehat{I}, \underline{I}_1, \underline{I}_2, C, \mathcal{B})} \mathcal{P}(I_0, \widehat{I}, \underline{I}_1) \left[-C(\widehat{I}, \underline{I}_1, \underline{I}_2) + \widetilde{\mathcal{P}}(\widehat{I}, \bar{I}, \underline{I}_2) \left(\frac{\kappa}{r} - \mathcal{B}(\underline{I}_2) \right) \right]$$

$$\text{subject to } C(\widehat{I}, \underline{I}_1, \underline{I}_2) = \max \left[\frac{\phi}{\mu} \frac{\exp(\eta^- \underline{I}_1 + \eta^+ \widehat{I}) - \exp(\eta^- \widehat{I} + \eta^+ \underline{I}_1)}{\eta^+ \exp(\eta^- \underline{I}_1 + \eta^+ I^*) - \eta^- \exp(\eta^- I^* + \eta^+ \underline{I}_1)} - \widetilde{\mathcal{P}}_A(\widehat{I}, \bar{I}, \underline{I}_2) \mathcal{B}(\underline{I}_2), 0 \right], \text{ and}$$

$$\mathcal{B}(\underline{I}_2) = \frac{\exp(\eta^- \underline{I}_2 + \eta^+ \bar{I}) - \exp(\eta^- \bar{I} + \eta^+ \underline{I}_2)}{\eta^+ \exp(\eta^- \underline{I}_2 + \eta^+ I^{**}) - \eta^- \exp(\eta^- I^{**} + \eta^+ \underline{I}_2)} \frac{\phi}{\mu}.$$

The compactness of $(\widehat{I}, \underline{I}_1, \underline{I}_2) \in ([I_0, \bar{I}], [\underline{I}^*, I_0], [\underline{I}_1, \widehat{I}])$ and the continuity of the objective function on the region guarantee the existence of a solution to the principal's problem.

B.2 Additional Mathematical Results

Theorem 1 [1] *Let X be a* (μ, σ^2) *Brownian motion with initial condition* $x \in [b, B]$*. Let* τ *be the stopping time* $\tau = \min[\tau_b, \tau_B]$*, and* $r > 0$*. If* $\sigma^2 > 0$*, then*

$$E_x[e^{-r\tau} | X(\tau) = b] P_x(X(\tau) = b) = \frac{e^{R_1 x} e^{R_2 B} - e^{R_2 x} e^{R_1 B}}{e^{R_1 b} e^{R_2 B} - e^{R_2 b} e^{R_1 B}},$$

$$E_x[e^{-r\tau} | X(\tau) = B] P_x(X(\tau) = B) = \frac{e^{R_1 b} e^{R_2 x} - e^{R_2 b} e^{R_1 x}}{e^{R_1 b} e^{R_2 B} - e^{R_2 b} e^{R_1 B}},$$

where

$$R_1 = \frac{-\mu - \sqrt{\mu^2 + 2r\sigma^2}}{\sigma^2} \text{ and } R_2 = \frac{-\mu + \sqrt{\mu^2 + 2r\sigma^2}}{\sigma^2}.$$

[1] Proposition 5.3. in Stokey (2008)

Theorem 2 *Let X be a (μ, σ^2) Brownian motion with initial condition $x \in [b, B]$. Let τ be the stopping time $\tau = \min[\tau_b, \tau_B]$, and $r > 0$. If $\sigma^2 > 0$, then*

$$\frac{\partial\{E_x[e^{-r\tau}|X(\tau) = B]P_x(X(\tau) = B)\}}{\partial r} < 0.$$

Proof 24 *First. notice that*

$$\frac{\partial R_1}{\partial r} = -\frac{1}{\sqrt{\mu^2 + 2r\sigma^2}} \text{ and } \frac{\partial R_2}{\partial r} = \frac{1}{\sqrt{\mu^2 + 2r\sigma^2}}.$$

After some algebra, I can obtain

$$\frac{\partial\{E_x[e^{-r\tau}|X(\tau) = B]P_x(X(\tau) = B)\}}{\partial r} = \frac{(\sqrt{\mu^2 + 2r\sigma^2})^{-1}}{[e^{R_1 b + R_2 B} - e^{R_1 B + R_2 b}]^2} e^{2R_1 b + R_2(B+x)} \cdot$$
$$\underbrace{\left[(1 - e^{-(R_2 - R_2)(B+x-2b)})(x - B) + (e^{-(R_2 - R_1)(x-b)} - e^{-(R_2 - R_1)(B-b)})(B + x - 2b)\right]}_{\equiv f(x)}.$$

Since $e^{2R_1 b + R_2(B+x)} > 0$, it is enough to show that $f(x) < 0$ for $b < x < B$. Notice that $f(B) = f(b) = 0$. Now, I show that $f(x)$ is a strictly convex function on $b < x < B$. The second derivative of $f(x)$ is

$$\begin{aligned} f''(x) = {} & a^2 e^{-a(B+x-2b)}(B - x) + a^2 e^{-a(x-b)}(B + x - 2b) + 2a(e^{-a(B+x-2b)} \\ & - e^{-a(x-b)}), \end{aligned}$$

where $a \equiv (R_2 - R_1) > 0$. Since $(B - b) > 0$ and $a > 0$,

$$\begin{aligned} f''(x) &= ae^{-a(x-b)}[ae^{-a(B-b)}(B - b) + 2e^{-a(B-b)} - ae^{-a(B-b)}(x - b) \\ &\quad + a(B + x - 2b) - 2] \\ &> ae^{-a(x-b)}[2 - a(B - b) - ae^{-a(B-b)}(x - b) + a(B + x - 2b) - 2] \\ &= ae^{-a(x-b)}a(x - b)(1 - e^{-a(B-b)}) > 0. \end{aligned}$$

Theorem 3 [2] *Let X be a (μ, σ^2) Brownian motion with initial condition $x \in [b, B]$, and let τ be the stopping time $\tau = \min[\tau_b, \tau_B]$. If $\sigma^2 > 0$, then*

$$P_x(X(\tau) = b) = \frac{e^{-\delta B} - e^{-\delta x}}{e^{-\delta B} - e^{-\delta b}},$$

$$P_x(X(\tau) = B) = \frac{e^{-\delta x} - e^{-\delta b}}{e^{-\delta B} - e^{-\delta b}}, \quad \text{if } \mu \neq 0,$$

[2] Proposition 5.4. in Stokey (2008)

where $\delta \equiv 2\mu/\sigma^2$*, and*

$$P_x(X(\tau) = b) = \frac{B - x}{B - b},$$

$$P_x(X(\tau) = B) = \frac{x - b}{B - b}, \quad \text{\textit{if} } \mu = 0.$$

Appendix C

APPENDIX TO CHAPTER 4

C.1 Derivation of the Simple Example

Consider a strategy profile ($\hat{s}$) of the following form:

- An agent with ability $\theta \geq \widehat{\theta}$ pursues the high-paying career.
- An agent with ability $\theta < \widehat{\theta}$ participates in the labor market of the low-paying career paths.

If an agent with ability $\widehat{\theta}$ pursues the high-paying career, then her expected utility would be

$$\begin{aligned} E[u(\cdot)|(s_i(\widehat{\theta}) = \mathcal{H}, \hat{s}_{-i}(\theta_{-i}))] &= \sum_{j=2}^{3} \binom{3}{j} \widehat{\theta}^j (1 - \widehat{\theta})^{3-j} u(H_{4-j}) \\ &+ \sum_{j=0}^{1} \binom{3}{j} \widehat{\theta}^j (1 - \widehat{\theta})^{3-j} u(O_H) \\ &= -2\widehat{\theta}^3 + 6\widehat{\theta}^2. \end{aligned}$$

If this agent pursues the low-paying career, her expected utility would be

$$E[u(\cdot)|(s_i(\widehat{\theta}) = \mathcal{L}, \hat{s}_{-i}(\theta_{-i}))] = u(L_1) = 3$$

since she is the best candidate in the labor market of the low-paying career paths if she chooses to pursue the low-paying career. In order for $\widehat{\theta}$ to be the equilibrium cut-off, the two expected utilities must have the same value. Hence,

$$-2\widehat{\theta}^3 + 6\widehat{\theta}^2 = 3.$$

Now, if an agent with ability $\theta \in [0, < \widehat{\theta})$ pursues the high-paying career, her expected utility would be

$$\begin{aligned} E[u(\cdot)|(s_i(\theta) = \mathcal{H}, \hat{s}_{-i}(\theta_{-i}))] &= \sum_{j=2}^{3} \binom{3}{j} (\widehat{\theta} - \mathcal{P}(\theta,\widehat{\theta}))^j (1 - \widehat{\theta} + \mathcal{P}(\theta,\widehat{\theta}))^{3-j} u(H_{4-j}) \\ &+ \sum_{j=2}^{3} \binom{3}{j} (\widehat{\theta} - \mathcal{P}(\theta,\widehat{\theta}))^j (1 - \widehat{\theta} + \mathcal{P}(\theta,\widehat{\theta}))^{3-j} u(O_H) \\ &= -2\widehat{\theta}^3 + 6\widehat{\theta}^2 + 6\widehat{\theta}^2\mathcal{P}(\theta,\widehat{\theta}) - 6\widehat{\theta}\mathcal{P}(\theta,\widehat{\theta})^2 - 12\widehat{\theta}\mathcal{P}(\theta,\widehat{\theta}) \\ &+ 6\mathcal{P}(\theta,\widehat{\theta})^2 + 2\mathcal{P}(\theta,\widehat{\theta})^3, \end{aligned}$$

where $\mathcal{P}(\theta,\widehat{\theta}) = \int_{\theta}^{\widehat{\theta}} f(x)g(x)dx = \int_{\theta}^{\widehat{\theta}} g(x)dx$. On the other hand, if this agent decides to attend the other labor market, her expected utility would be

$$\begin{aligned} E[u(\cdot)|(s_i(\theta) = \mathcal{L}, \hat{s}_{-i}(\theta_{-i}))] &= \sum_{j=0}^{1} \binom{3}{j} (\widehat{\theta} - \theta - \mathcal{P}(\theta,\widehat{\theta}))^j (1 - \widehat{\theta} + \mathcal{P}(\theta,\widehat{\theta}) + \theta)^{3-j} u(L_{j+1}) \\ &+ \sum_{j=0}^{1} \binom{3}{j} (\widehat{\theta} - \theta - \mathcal{P}(\theta,\widehat{\theta}))^j (1 - \widehat{\theta} + \mathcal{P}(\theta,\widehat{\theta}) + \theta)^{3-j} u(O_L) \\ &= 3 - 6\widehat{\theta} + 3\widehat{\theta}^2 - 6\widehat{\theta}\theta + 6\theta + 3\theta^2 + 6\mathcal{P}(\theta,\widehat{\theta}) - 6\widehat{\theta}\mathcal{P}(\theta,\widehat{\theta}) \\ &+ 6\theta\mathcal{P}(\theta,\widehat{\theta}) + 3\mathcal{P}(\theta,\widehat{\theta})^2. \end{aligned}$$

Equating these equations gives the equilibrium strategy.

C.2 Proofs

C.2.1 Proof of Proposition 23

Consider a strategy profile $(\hat{s})$ of the following form:

- An agent with ability $\theta \geq \widehat{\theta}$ pursues the high-paying career.
- An agent with ability $\theta < \widehat{\theta}$ participates in the labor market of the low-paying career paths.

I show that this strategy can not be an equilibrium. When agents follow this strategy, the expected utility of an agent with ability $\widehat{\theta}$ is

$$\begin{aligned} E[u(\cdot)|(\hat{s}_i(\widehat{\theta}) = \mathcal{H}, \hat{s}_{-i}(\theta_{-i}))] = &\sum_{j=I-n}^{I-1} \binom{I-1}{j} F(\widehat{\theta})^j (1 - F(\widehat{\theta}))^{I-1-j} u(H_{n-j}) \\ &+ \sum_{j=0}^{I-n-1} \binom{I-1}{j} F(\widehat{\theta})^j (1 - F(\widehat{\theta}))^{I-1-j} u(O_H) \end{aligned}$$

if she pursues the high-paying career.

However, if she participates in the other labor market, her expected utility would be

$$E[u(\cdot)|(\hat{s}_i(\widehat{\theta}) = \mathcal{L}, \hat{s}_{-i}(\theta_{-i}))] = u(L_1).$$

Therefore, the difference in the expected utility of an agent with ability $\widehat{\theta}$ between pursuing the high-paying career and the low-paying career is

$$\widehat{D}(\widehat{\theta}) \equiv E[u(\cdot)|(\hat{s}_i(\theta) = \mathcal{H}, \hat{s}_{-i}(\theta_{-i}))] - E[u(\cdot)|(\hat{s}_i(\theta) = \mathcal{L}, \hat{s}_{-i}(\theta_{-i}))]$$

when agents follow the pure strategy profile, $\hat{s}$, where I use the hat notation to clarify that this difference depends on the strategy profile $\hat{s}$ through $\widehat{\theta}$.

Then, $\widehat{D}(\widehat{\theta}) = u(H_1) - u(L_1) > 0$ when $\widehat{\theta} = \bar{\theta}$ and $\widehat{D}(\widehat{\theta}) = u(\emptyset) - u(L_1) < 0$ if $\widehat{\theta} = \underline{\theta}$. Now, I show that $\partial\widehat{D}/\partial\widehat{\theta} > 0$ for $\widehat{\theta} \in (\underline{\theta}, \bar{\theta})$. This implies that there is a unique $\widehat{\theta}$ such that $\widehat{D}(\widehat{\theta}) = 0$. Note that

$$\begin{aligned}\frac{\partial\widehat{D}(\widehat{\theta})}{\partial\widehat{\theta}} = f(\widehat{\theta})\Bigg\{&\sum_{j=I-n}^{I-1}\binom{I-1}{j}F(\widehat{\theta})^{j-1}(1-F(\widehat{\theta}))^{I-2-j}u(G_{I-j})[j-(I-1)F(\widehat{\theta})]\\ &+\sum_{j=0}^{I-n-1}\binom{I-1}{j}F(\widehat{\theta})^{j-1}(1-F(\widehat{\theta}))^{I-2-j}u(O_H)[j-(I-1)F(\widehat{\theta})]\Bigg\}.\end{aligned}$$

First, notice that

$$\lim_{\widehat{\theta}\downarrow\underline{\theta}}\frac{\partial\widehat{D}(\widehat{\theta})}{\partial\widehat{\theta}} = 0 \quad \text{and} \quad \lim_{\widehat{\theta}\uparrow\bar{\theta}}\frac{\partial\widehat{D}(\widehat{\theta})}{\partial\widehat{\theta}} = (n-1)f(U)(u(H_1)-u(H_2)) > 0.$$

For $\widehat{\theta}$ such that $\frac{k-1}{I-1} < F(\widehat{\theta}) \leq \frac{k}{I-1}$, $k = 1, 2, \ldots, I-2$ and $\frac{I-2}{I-1} < F(\widehat{\theta}) < 1$,

$$\begin{aligned}\frac{\partial\widehat{D}(\widehat{\theta})}{\partial\widehat{\theta}} &= f(\widehat{\theta})\Bigg\{\sum_{j=I-n}^{I-1}\binom{I-1}{j}F(\widehat{\theta})^{j-1}(1-F(\widehat{\theta}))^{I-2-j}u(H_{I-j})[j-(I-1)F(\widehat{\theta})]\\ &\quad+\sum_{j=0}^{I-n-1}\binom{I-1}{j}F(\widehat{\theta})^{j-1}(1-F(\widehat{\theta}))^{I-2-j}u(O_H)[j-(I-1)F(\widehat{\theta})]\Bigg\}\\ &> \frac{f(\widehat{\theta})}{F(\widehat{\theta})(1-F(\widehat{\theta}))}u(\widetilde{H}_k)\Bigg\{\sum_{j=I-n}^{I-1}\binom{I-1}{j}F(\widehat{\theta})^{j}(1-F(\widehat{\theta}))^{I-1-j}[j-(I-1)F(\widehat{\theta})]\\ &\quad+\sum_{j=0}^{I-n-1}\binom{I-1}{j}F(\widehat{\theta})^{j}(1-F(\widehat{\theta}))^{I-1-j}[j-(I-1)F(\widehat{\theta})]\Bigg\}\\ &= \frac{f(\widehat{\theta})}{F(\widehat{\theta})(1-F(\widehat{\theta}))}u(H_k)[(I-1)F(\widehat{\theta})-(I-1)F(\widehat{\theta})]\\ &= 0,\end{aligned}$$

where

$$u(\widetilde{H}_k) = \begin{cases} u(H_k) & \text{if } k \leq n \\ u(O_H) & \text{if } k > n \end{cases}.$$

Therefore, there is a unique $\underline{\theta} < \widehat{\theta}^* < \bar{\theta}$ such that $\widehat{D}(\widehat{\theta}^*) = 0$. However, if agents use this strategy, an agent with ability $\theta \in [\underline{\theta}, \widehat{\theta}^*)$ has an incentive to deviate. In particular, if she follows this strategy, her expected utility is strictly less than $u(L_1)$. However, if she pursues the high-paying career instead, her expected utility would be $u(L_1)$ since her expected utility when she pursue the high-paying career is exactly the same as an agent with ability $\widehat{\theta}^*$. Hence, there is no symmetric pure strategy equilibrium in the form of $\hat{s}$. Also, for any other form of pure strategy, there is no symmetric pure strategy equilibrium by the same logic.

C.2.2 Proof of Proposition 24

Consider a strategy profile ($\hat{s}$) of the following form:

- An agent with ability $\theta \geq \widehat{\theta}$ pursues the high-paying career with probability one.
- An agent with ability $\theta < \widehat{\theta}$ pursue the high-paying career with probability $g(\theta)$.

Also, recall that

$$\mathcal{P}(a,b) \equiv \left(\int_a^b f(\theta)g(\theta)d\theta\right), \quad \text{and} \quad Q(a,b) \equiv \left(\int_a^b f(\theta)(1 - g(\theta))d\theta\right).$$

If an agent with ability $\widehat{\theta}$ pursues the high-paying career, her expected utility would be

$$E[u(\cdot)|(\hat{s}_i(\widehat{\theta}) = \mathcal{H}, \hat{s}_{-i}(\theta_{-i}))] = \sum_{j=0}^{n-1} \binom{I-1}{j} (1 - F(\widehat{\theta}))^j F(\widehat{\theta})^{I-1-j} u(H_{j+1}) + \sum_{j=n}^{I-1} \binom{I-1}{j} (1 - F(\widehat{\theta}))^j F(\widehat{\theta})^{I-1-j} u(O_H).$$

Clearly, if she participates in the other labor market, her expected utility would be $u(L_1)$.

Now, if an agent with ability $\theta \in [L, \widehat{\theta})$ pursues the high-paying career, her expected utility is equal to

$$\begin{aligned} &E[u(\cdot)|(\hat{s}_i(\theta) = \mathcal{H}, \hat{s}_{-i}(\theta_{-i}))] \\ &\quad = \sum_{j=0}^{n-1} \binom{I-1}{j} (1 - F(\widehat{\theta}) + \mathcal{P}(\theta, \widehat{\theta}))^j (F(\theta) + \mathcal{Q}(\theta, \widehat{\theta}))^{I-1-j} u(H_{j+1}) \\ &\qquad + \sum_{j=n}^{I-1} \binom{I-1}{j} (1 - F(\widehat{\theta}) + \mathcal{P}(\theta, \widehat{\theta}))^j (F(\theta) + \mathcal{Q}(\theta, \widehat{\theta}))^{I-1-j} u(O_H). \end{aligned}$$

However, if she decides to participate in the other labor market, her expected utility would be

$$\begin{aligned} &E[u(\cdot)|(\hat{s}_i(\theta) = \mathcal{L}, \hat{s}_{-i}(\theta_{-i}))] \\ &\quad = \sum_{j=0}^{m-1} \binom{I-1}{j} \mathcal{Q}(\theta, \widehat{\theta})^j (1 - F(\widehat{\theta}) + \mathcal{P}(\theta, \widehat{\theta}) + F(\theta))^{I-1-j} u(L_{j+1}) \\ &\qquad + \sum_{j=m}^{I-1} \binom{I-1}{j} \mathcal{Q}(\theta, \widehat{\theta})^j (1 - F(\widehat{\theta}) + \mathcal{P}(\theta, \widehat{\theta}) + F(\theta))^{I-1-j} u(O_L). \end{aligned}$$

Recall that the function $D(\cdot)$ is

$$D(\theta) = E[u(\cdot)|(\hat{s}_i(\theta) = \mathcal{H}, \hat{s}_{-i}(\theta_{-i}))] - E[u(\cdot)|(\hat{s}_i(\theta) = \mathcal{L}, \hat{s}_{-i}(\theta_{-i}))].$$

Proof 25 (Proof of Claim 1) *If* $\widehat{\theta} = \underline{\theta}$, $D(\underline{\theta}) = u(O_H) - u(L_1) < 0$. *Hence,* $\widehat{\theta} > \underline{\theta}$.

Proof 26 (Proof of Claim 2) *For* $\underline{\theta} \leq \theta < \widehat{\theta}$, *the difference between the two actions is*

$$\begin{aligned} D(\theta) = &\sum_{j=0}^{n-1} \binom{I-1}{j} (1 - F(\widehat{\theta}) + \mathcal{P}(\theta, \widehat{\theta}))^j (F(\theta) + \mathcal{Q}(\theta, \widehat{\theta}))^{I-1-j} u(H_{j+1}) \\ &+ \sum_{j=n}^{I-1} \binom{I-1}{j} (1 - F(\widehat{\theta}) + \mathcal{P}(\theta, \widehat{\theta}))^j (F(\theta) + \mathcal{Q}(\theta, \widehat{\theta}))^{I-1-j} u(O_H) \\ &- \sum_{j=0}^{m-1} \binom{I-1}{j} \mathcal{Q}(\theta, \widehat{\theta})^j (1 - F(\widehat{\theta}) + \mathcal{P}(\theta, \widehat{\theta}) + F(\theta))^{I-1-j} u(L_{j+1}) \\ &- \sum_{j=m}^{I-1} \binom{I-1}{j} \mathcal{Q}(\theta, \widehat{\theta})^j (1 - F(\widehat{\theta}) + \mathcal{P}(\theta, \widehat{\theta}) + F(\theta))^{I-1-j} u(O_L). \end{aligned}$$

Consider the case where $\theta = \underline{\theta}$. Then,

$$
\begin{aligned}
D(\underline{\theta}) = &\sum_{j=I-n}^{I-1} \binom{I-1}{j} Q(\underline{\theta},\widehat{\theta})^j (1 - Q(\underline{\theta},\widehat{\theta}))^{I-1-j} u(H_{I-j}) \\
&+ \sum_{j=0}^{I-n-1} \binom{I-1}{j} Q(\underline{\theta},\widehat{\theta})^j (1 - Q(\underline{\theta},\widehat{\theta}))^{I-1-j} u(O_H) \\
&- \sum_{j=0}^{m-1} \binom{I-1}{j} Q(\underline{\theta},\widehat{\theta})^j (1 - Q(\underline{\theta},\widehat{\theta}))^{I-1-j} u(L_{j+1}) \\
&- \sum_{j=m}^{I-1} \binom{I-1}{j} Q(\underline{\theta},\widehat{\theta})^j (1 - Q(\underline{\theta},\widehat{\theta}))^{I-1-j} u(O_L).
\end{aligned}
$$

Note that $D(\underline{\theta}) = u(O_H) - u(L_1) < 0$ when $Q(\underline{\theta},\widehat{\theta}) = 0$, and $D(\underline{\theta}) = u(H_1) - u(O_L) > 0$ when $Q(\underline{\theta},\widehat{\theta}) = 1$. For $0 < Q(\underline{\theta},\widehat{\theta}) < 1$, the differential of this difference with respect to $Q(\underline{\theta},\widehat{\theta})$ is

$$
\begin{aligned}
\frac{\partial D(\underline{\theta})}{\partial Q(\underline{\theta},\widehat{\theta})} = &\sum_{j=I-n}^{I-1} \binom{I-1}{j} Q(\underline{\theta},\widehat{\theta})^{j-1} (1 - Q(\underline{\theta},\widehat{\theta}))^{I-2-j} u(H_{I-j})[j - (I-1)Q(\underline{\theta},\widehat{\theta})] \\
&+ \sum_{j=0}^{I-n-1} \binom{I-1}{j} Q(\underline{\theta},\widehat{\theta})^{j-1} (1 - Q(\underline{\theta},\widehat{\theta}))^{I-2-j} u(O_H)[j - (I-1)Q(\underline{\theta},\widehat{\theta})] \\
&- \sum_{j=0}^{m-1} \binom{I-1}{j} Q(\underline{\theta},\widehat{\theta})^{j-1} (1 - Q(\underline{\theta},\widehat{\theta}))^{I-2-j} u(L_{j+1})[j - (I-1)Q(\underline{\theta},\widehat{\theta})] \\
&- \sum_{j=m}^{I-1} \binom{I-1}{j} Q(\underline{\theta},\widehat{\theta})^{j-1} (1 - Q(\underline{\theta},\widehat{\theta}))^{I-2-j} u(O_L)[j - (I-1)Q(\underline{\theta},\widehat{\theta})].
\end{aligned}
$$

I define a pseudo-utility function $U(\cdot)$ by

$$
U(H_i) = \begin{cases} u(H_{I-i}) & \text{if } I - n \le i \le I - 1 \\ u(O_H) & \text{if } 0 \le i < I - n, \end{cases} \quad \text{and } U(L_i) = \begin{cases} u(O_L) & \text{if } m \le i \le I - 1 \\ u(L_{i+1}) & \text{if } 0 \le i < m. \end{cases}
$$

Using this function, I can rewrite $\frac{\partial D(\underline{\theta})}{\partial Q(\underline{\theta},\widehat{\theta})}$ *as*

$$\begin{aligned}
\frac{\partial D(\underline{\theta})}{\partial Q(\underline{\theta},\widehat{\theta})} &= \frac{1}{Q(\underline{\theta},\widehat{\theta})(1-Q(\underline{\theta},\widehat{\theta}))}\Bigg[\sum_{j=0}^{I-1}\binom{I-1}{j} jU(H_j)Q(\underline{\theta},\widehat{\theta})^j(1-Q(\underline{\theta},\widehat{\theta}))^{I-1-j} \\
&\quad -(I-1)Q(\underline{\theta},\widehat{\theta})\sum_{j=0}^{I-1}\binom{I-1}{j} U(H_j)Q(\underline{\theta},\widehat{\theta})^j(1-Q(\underline{\theta},\widehat{\theta}))^{I-1-j} \\
&\quad -\sum_{j=0}^{I-1}\binom{I-1}{j} jU(L_j)Q(\underline{\theta},\widehat{\theta})^j(1-Q(\underline{\theta},\widehat{\theta}))^{I-1-j} \\
&\quad +(I-1)Q(\underline{\theta},\widehat{\theta})\sum_{j=0}^{I-1}\binom{I-1}{j} U(L_j)Q(\underline{\theta},\widehat{\theta})^j(1-Q(\underline{\theta},\widehat{\theta}))^{I-1-j}\Bigg] \\
&> \frac{1}{Q(\underline{\theta},\widehat{\theta})(1-Q(\underline{\theta},\widehat{\theta}))}\Big\{E[J]E[U(H_J)] - (I-1)Q(\underline{\theta},\widehat{\theta})E[U(H_J)] \\
&\quad -E[J]E[U(L_J)] + (I-1)Q(\underline{\theta},\widehat{\theta})E[U(L_J)]\Big\} \\
&= \frac{1}{Q(\underline{\theta},\widehat{\theta})(1-Q(\underline{\theta},\widehat{\theta}))}\Big\{(I-1)Q(\underline{\theta},\widehat{\theta})E[U(H_J)] - (I-1)Q(\underline{\theta},\widehat{\theta})E[U(H_J)] \\
&\quad -(I-1)Q(\underline{\theta},\widehat{\theta})E[U(L_J)] + (I-1)Q(\underline{\theta},\widehat{\theta})E[U(L_J)]\Big\} \\
&= 0,
\end{aligned}$$

where

$$\begin{aligned}
E[J] &= \sum_{j=0}^{I-1}\binom{I-1}{j} jQ(\underline{\theta},\widehat{\theta})^j(1-Q(\underline{\theta},\widehat{\theta}))^{I-1-j} \\
E[U(H_J)] &= \sum_{j=0}^{I-1}\binom{I-1}{j} U(H_j)Q(\underline{\theta},\widehat{\theta})^j(1-Q(\underline{\theta},\widehat{\theta}))^{I-1-j}, \text{ and} \\
E[U(L_J)] &= \sum_{j=0}^{I-1}\binom{I-1}{j} U(L_j)Q(\underline{\theta},\widehat{\theta})^j(1-Q(\underline{\theta},\widehat{\theta}))^{I-1-j}.
\end{aligned}$$

The above inequality holds since the function $U(H_j)$ *is an increasing function in* j*,* $U(L_j)$ *is a decreasing function in* j*, and* $0 < Q(\underline{\theta},\widehat{\theta}) < 1$*. This implies that* $E[JU(H_J)] > E[J]E[U(H_J)]$ *and* $E[JU(L_J)] < E[J]E[U(L_J)]$

Therefore, there exists a unique value of $Q(\underline{\theta},\widehat{\theta})$ *such that* $D(\underline{\theta}) = 0$*. I denote this value by* $Q^* \in (0,1)$*.*

Proof 27 (Proof of Claim 3) *I have already shown that* $Q(\underline{\theta},\widehat{\theta})$ *has a unique value* Q^* *in equilibrium. Therefore, it suffices to show that there exists a unique* $\widehat{\theta}$ *such* *that* $D(\widehat{\theta}) = 0$ *when* $Q(\underline{\theta}, \widehat{\theta}) = Q^*$. *Note that*

$$D(\widehat{\theta}) = \sum_{j=0}^{n-1} \binom{I-1}{j} (1 - F(\widehat{\theta}))^j F(\widehat{\theta})^{I-1-j} u(H_{j+1}) \\ + \sum_{j=n}^{I-1} \binom{I-1}{j} (1 - F(\widehat{\theta}))^j F(\widehat{\theta})^{I-1-j} u(O_H) - u(L_1).$$

This difference has the value $u(H_1) - u(L_1) > 0$ *when* $F(\widehat{\theta}) = 1$, *and*

$$D(\widehat{\theta}) = \sum_{j=0}^{m-1} \binom{I-1}{j} (Q^*)^j (1 - Q^*)^{I-1-j} u(L_{j+1}) \\ + \sum_{j=m}^{I-1} \binom{I-1}{j} (Q^*)^j (1 - Q^*)^{I-1-j} u(O_L) - u(L_1) < 0$$

when $F(\widehat{\theta}) = Q^*$.[1]

The differential of this difference with respect to $F(\widehat{\theta})$ *is*

$$\frac{\partial D(\widehat{\theta})}{\partial F(\widehat{\theta})} = \sum_{j=0}^{n-1} \binom{I-1}{j} (1 - F(\widehat{\theta}))^{j-1} F(\widehat{\theta})^{I-2-j} u(H_{j+1}) [(I-1)(1 - F(\widehat{\theta})) - j] \\ + \sum_{j=n}^{I-1} \binom{I-1}{j} (1 - F(\widehat{\theta}))^{j-1} F(\widehat{\theta})^{I-2-j} u(O_H) [(I-1)(1 - F(\widehat{\theta})) - j].$$

Using the pseudo-utility function I defined before, the above equation can be rewritten as

$$\begin{aligned} \frac{\partial D(\widehat{\theta})}{\partial F(\widehat{\theta})} &= \frac{1}{F(\widehat{\theta})(1 - F(\widehat{\theta}))} \sum_{j=0}^{I-1} \binom{I-1}{j} F(\widehat{\theta})^j (1 - F(\widehat{\theta}))^{I-1-j} U(H_j) [j - (I-1) F(\widehat{\theta})] \\ &> \frac{1}{F(\widehat{\theta})(1 - F(\widehat{\theta}))} \left\{ E[J] E[U(H_J)] - (I-1) F(\widehat{\theta}) E[U(H_J)] \right\} \\ &= \frac{1}{F(\widehat{\theta})(1 - F(\widehat{\theta}))} \left\{ (I-1) F(\widehat{\theta}) E[U(H_J)] - (I-1) F(\widehat{\theta}) E[U(H_J)] \right\} \\ &= 0. \end{aligned}$$

Therefore, the differential of the difference is strictly greater than 0. Hence, there exist a unique $\widehat{\theta}$ *and* $Q(\underline{\theta}, \widehat{\theta})$ *such that* $D(\widehat{\theta}) = 0$ *and* $D(\underline{\theta}) = 0$. *From now on, I denote these unique* $\widehat{\theta}$, $\mathcal{P}(\underline{\theta}, \widehat{\theta})$, *and* $Q(\underline{\theta}, \widehat{\theta})$ *by* $\widehat{\theta}^*$, $\mathcal{P}(\underline{\theta}, \widehat{\theta}^*)$, *and* $Q(\underline{\theta}, \widehat{\theta}^*) \equiv Q^*$.

[1] Recall that Q^* is the solution to the equation $D(\underline{\theta}) = 0$.

Proof 28 (Proof of Claim 4) *Consider an agent with ability $\theta \in (\underline{\theta}, \widehat{\theta}^*)$. Recall that*

$$\begin{aligned}
D(\theta) = &\sum_{j=0}^{n-1} \binom{I-1}{j} (1 - F(\theta) - Q^* + Q(\underline{\theta}, \theta))^j (F(\theta) + Q^* - Q(\underline{\theta}, \theta))^{I-1-j} u(H_{j+1}) \\
&+ \sum_{j=n}^{I-1} \binom{I-1}{j} (1 - F(\theta) - Q^* + Q(\underline{\theta}, \theta))^j (F(\theta) + Q^* - Q(\underline{\theta}, \theta))^{I-1-j} u(O_H) \\
&- \sum_{j=0}^{m-1} \binom{I-1}{j} (Q^* - Q(\underline{\theta}, \theta))^j (1 - Q^* + Q(\underline{\theta}, \theta))^{I-1-j} u(L_{j+1}) \\
&- \sum_{j=m}^{I-1} \binom{I-1}{j} (Q^* - Q(\underline{\theta}, \theta))^j (1 - Q^* + Q(\underline{\theta}, \theta))^{I-1-j} u(O_L).
\end{aligned}$$

Now, I show that there is a unique value of $Q(\underline{\theta}, \theta)$ such that $D(\theta) = 0$. First, notice that $\max\{0, Q^* - F(\widehat{\theta}) + F(\theta)\} \leq Q(\underline{\theta}, \theta) \leq \min\{F(\theta), Q^*\}$ *from $\mathcal{P}(\underline{\theta}, \theta) + Q(\underline{\theta}, \theta) = F(\theta)$ and $Q(\underline{\theta}, \theta) + Q(\theta, \widehat{\theta}) = Q^*$.*

Consider the four possible boundaries. This shows that $D(\theta)$ has the value $D(\theta) > 0$ when $Q(\underline{\theta}, \theta) = \max\{0, Q^* - F(\widehat{\theta}) + F(\theta)\}$ and $D(\theta) < 0$ if $Q(\underline{\theta}, \theta) = \min\{F(\theta), Q^*\}$.

- $Q(\underline{\theta}, \theta) = 0$*: From $D(\underline{\theta}) = 0$,*

$$\begin{aligned}
D(\theta) = &\sum_{j=0}^{n-1} \binom{I-1}{j} (1 - F(\theta) - Q^*)^j (F(\theta) + Q^*)^{I-1-j} u(H_{j+1}) \\
&+ \sum_{j=n}^{I-1} \binom{I-1}{j} (1 - F(\theta) - Q^*)^j (F(\theta) + Q^*)^{I-1-j} u(O_H) \\
&- \sum_{j=0}^{n-1} \binom{I-1}{j} (1 - Q^*)^j (Q^*)^{I-1-j} u(H_{j+1}) \\
&- \sum_{j=n}^{I-1} \binom{I-1}{j} (1 - Q^*)^j (Q^*)^{I-1-j} u(O_H) \\
&> 0.
\end{aligned}$$

  The last inequality holds since $u(H_{j+1})$ is a strictly decreasing function in j and $F(\theta) > 0$.

- $Q(\underline{\theta}, \theta) = Q^* - F(\widehat{\theta}) + F(\theta)$*:*

From $D(\widehat{\theta}) = 0$,

$$
\begin{aligned}
D(\theta) = & \sum_{j=0}^{n-1} \binom{I-1}{j} (1 - F(\widehat{\theta}))^j F(\widehat{\theta})^{I-1-j} u(H_{j+1}) \\
& + \sum_{j=n}^{I-1} \binom{I-1}{j} (1 - F(\widehat{\theta}))^j F(\widehat{\theta})^{I-1-j} u(O_H) \\
& - \sum_{j=0}^{m-1} \binom{I-1}{j} (F(\widehat{\theta}) - F(\theta))^j (1 - F(\widehat{\theta}) + F(\theta))^{I-1-j} u(L_{j+1}) \\
& - \sum_{j=m}^{I-1} \binom{I-1}{j} (F(\widehat{\theta}) - F(\theta))^j (1 - F(\widehat{\theta}) + F(\theta))^{I-1-j} u(O_L) \\
= & \, u(L_1) - \sum_{j=0}^{m-1} \binom{I-1}{j} (F(\widehat{\theta}) - F(\theta))^j (1 - F(\widehat{\theta}) + F(\theta))^{I-1-j} u(L_{j+1}) \\
& - \sum_{j=m}^{I-1} \binom{I-1}{j} (F(\widehat{\theta}) - F(\theta))^j (1 - F(\widehat{\theta}) + F(\theta))^{I-1-j} u(O_L) \\
> & \, 0.
\end{aligned}
$$

- $Q(\underline{\theta}, \theta) = F(\theta)$*:*

  From $D(\underline{\theta}) = 0$,

$$
\begin{aligned}
D(\theta) = & \sum_{j=0}^{m-1} \binom{I-1}{j} (Q^*)^j (1 - Q^*)^{I-1-j} u(L_{j+1}) \\
& + \sum_{j=m}^{I-1} \binom{I-1}{j} (Q^*)^j (1 - Q^*)^{I-1-j} u(O_L) \\
& - \sum_{j=0}^{m-1} \binom{I-1}{j} (Q^* - F(\theta))^j (1 - Q^* + F(\theta))^{I-1-j} u(L_{j+1}) \\
& - \sum_{j=m}^{I-1} \binom{I-1}{j} (Q^* - F(\theta))^j (1 - Q^* + F(\theta))^{I-1-j} u(O_L) \\
< & \, 0.
\end{aligned}
$$

  The last inequality holds since $u(H_{j+1})$ *is a strictly decreasing function and* $F(\theta) > 0$.

- $Q(\underline{\theta}, \theta) = Q^*$*:*

Then,

$$D(\theta) = \sum_{j=0}^{n-1} \binom{I-1}{j} (1-F(\theta))^j F(\theta)^{I-1-j} u(H_{j+1})$$
$$+ \sum_{j=n}^{I-1} \binom{I-1}{j} (1-F(\theta))^j F(\theta)^{I-1-j} u(O_H) - u(L_1).$$

I have already shown that the differential of this equation is strictly greater than zero in ***Claim 3****. Since* $\theta < \widehat{\theta}$*,* $D(\theta) < D(\widehat{\theta}) = 0$.

Remaining proof is to show that $\partial D(\theta)/\partial Q(\underline{\theta}, \theta) < 0$*. Define* $S(\theta) := Q^* - Q(\underline{\theta}, \theta)$*. Then, it suffices to prove* $\partial D(\theta)/\partial S(\theta) > 0$*. The differential of the difference* $D(\theta)$ *with respect to* $S(\theta)$ *is*

$$\frac{\partial D(\theta)}{\partial S(\theta)} = \sum_{j=0}^{n-1} \binom{I-1}{j} (1-F(\theta)-S(\theta))^{j-1} (F(\theta)+S(\theta))^{I-2-j} u(H_{j+1}) \cdot$$
$$[(I-1)(1-F(\theta)-S(\theta)) - j]$$
$$+ \sum_{j=n}^{I-1} \binom{I-1}{j} (1-F(\theta)-S(\theta))^{j-1} (F(\theta)+S(\theta))^{I-2-j} u(O_H) \cdot$$
$$[(I-1)(1-F(\theta)-S(\theta)) - j]$$
$$- \sum_{j=0}^{m-1} \binom{I-1}{j} S(\theta)^{j-1} (1-S(\theta))^{I-2-j} u(L_{j+1}) [j - (I-1)S(\theta)]$$
$$- \sum_{j=m}^{I-1} \binom{I-1}{j} S(\theta)^{j-1} (1-S(\theta))^{I-2-j} u(O_L) [j - (I-1)S(\theta)]$$
$$= \frac{1}{(F(\theta)+S(\theta))(1-F(\theta)-S(\theta))} \cdot$$
$$\sum_{j=0}^{n-1} \binom{I-1}{j} (F(\theta)+S(\theta))^j (1-F(\theta)-S(\theta))^{I-1-j} U(H_j) \cdot$$
$$[j - (I-1)(F(\theta)+S(\theta))]$$
$$- \frac{1}{S(\theta)(1-S(\theta))} \sum_{j=m}^{I-1} \binom{I-1}{j} S(\theta)^{j-1} (1-S(\theta))^{I-2-j} U(L_j) [j - (I-1)S(\theta)]$$
$$> \frac{1}{(F(\theta)+S(\theta))(1-F(\theta)-S(\theta))} \cdot$$
$$\{E[J]E[U(H_J)] - (I-1)(F(\theta)+S(\theta))E[U(H_J)]\}$$
$$- \frac{1}{S(\theta)(1-S(\theta))} \{E[J]E[U(L_J)] - (I-1)S(\theta)E[U(L_J)]\}$$
$$= 0.$$

Hence, $\frac{\partial D(\theta)}{\partial Q(\underline{\theta},\theta)} < 0$.

Therefore, by the *Lebesgue Differentiation Theorem*, the symmetric equilibrium strategy $g(\theta)$ is unique among integrable functions.

C.2.3 Proof of Proposition 25

First, I derive $g(\theta)$. Recall that the equilibrium strategy $g(\theta)$ for $\theta \in [\underline{\theta}, \widehat{\theta})$ satisfies

$$
\begin{aligned}
0 = &\sum_{j=0}^{n-1} \binom{I-1}{j} (1 - F(\theta) - Q^* + Q(\underline{\theta},\theta))^j (F(\theta) + Q^* - Q(\underline{\theta},\theta))^{I-1-j} u(H_{j+1}) \\
&+ \sum_{j=n}^{I-1} \binom{I-1}{j} (1 - F(\theta) - Q^* + Q(\underline{\theta},\theta))^j (F(\theta) + Q^* - Q(\underline{\theta},\theta))^{I-1-j} u(O_H) \\
&- \sum_{j=0}^{m-1} \binom{I-1}{j} (Q^* - Q(\underline{\theta},\theta))^j (1 - Q^* + Q(\underline{\theta},\theta))^{I-1-j} u(L_{j+1}) \\
&- \sum_{j=m}^{I-1} \binom{I-1}{j} (Q^* - Q(\underline{\theta},\theta))^j (1 - Q^* + Q(\underline{\theta},\theta))^{I-1-j} u(O_L).
\end{aligned}
$$

Differentiating this equation with respect to θ gives the following result:

$$
\begin{aligned}
0 = {} & f(\theta)g(\theta)\cdot \\
& \left\{\sum_{j=0}^{n-1}\binom{I-1}{j}(1-F(\theta)-Q^*+Q(\underline{\theta},\theta))^{j-1}(F(\theta)+Q^*-Q(\underline{\theta},\theta))^{I-2-j}u(H_{j+1})\cdot\right. \\
& [(I-1)(1-F(\theta)-Q^*+Q(\underline{\theta},\theta))-j] \\
& +\sum_{j=n}^{I-1}\binom{I-1}{j}(1-F(\theta)-Q^*+Q(\underline{\theta},\theta))^{j-1}(F(\theta)+Q^*-Q(\underline{\theta},\theta))^{I-2-j}u(O_H)\cdot \\
& [(I-1)(1-F(\theta)-Q^*+Q(\underline{\theta},\theta))-j] \\
& -\sum_{j=0}^{m-1}\binom{I-1}{j}(Q^*-Q(\underline{\theta},\theta))^{j-1}(1-Q^*+Q(\underline{\theta},\theta))^{I-2-j}u(L_{j+1})\cdot \\
& [j-(I-1)(Q^*-Q(\underline{\theta},\theta))] \\
& -\sum_{j=m}^{I-1}\binom{I-1}{j}(Q^*-Q(\underline{\theta},\theta))^{j-1}(1-Q^*+Q(\underline{\theta},\theta))^{I-2-j}u(O_L)\cdot \\
& \left.[j-(I-1)(Q^*-Q(\underline{\theta},\theta))]\right\} \\
& + f(\theta)\left\{\sum_{j=0}^{m-1}\binom{I-1}{j}(Q^*-Q(\underline{\theta},\theta))^{j-1}(1-Q^*+Q(\underline{\theta},\theta))^{I-2-j}u(L_{j+1})\cdot\right. \\
& [j-(I-1)(Q^*-Q(\underline{\theta},\theta))] \\
& +\sum_{j=m}^{I-1}\binom{I-1}{j}(Q^*-Q(\underline{\theta},\theta))^{j-1}(1-Q^*+Q(\underline{\theta},\theta))^{I-2-j}u(O_L)\cdot \\
& \left.[j-(I-1)(Q^*-Q(\underline{\theta},\theta))]\right\}.
\end{aligned}
$$

Aligning this equation gives

$$
\begin{aligned}
g(\theta) = -\frac{1}{\frac{\partial D(\theta)}{\partial S(\theta)}} & \left\{\sum_{j=0}^{m-1}\binom{I-1}{j}(Q^*-Q(\underline{\theta},\theta))^{j-1}(1-Q^*+Q(\underline{\theta},\theta))^{I-2-j}u(L_{j+1})\cdot\right. \\
& [j-(I-1)(Q^*-Q(\underline{\theta},\theta))] \\
& +\sum_{j=m}^{I-1}\binom{I-1}{j}(Q^*-Q(\underline{\theta},\theta))^{j-1}(1-Q^*+Q(\underline{\theta},\theta))^{I-2-j}u(O_L)\cdot \\
& \left.[j-(I-1)(Q^*-Q(\underline{\theta},\theta))]\right\}.
\end{aligned}
$$

From the proof of **Claim 4**, it can be easily shown that $0 < g(\theta) < 1$. Therefore,

$$
g(\theta) = -\frac{\mathcal{N}(\theta)}{\mathcal{M}(\theta)-\mathcal{N}(\theta)},
$$

where

$$\mathcal{M}(\theta) \equiv \sum_{j=0}^{n-1} \binom{I-1}{j} (1 - F(\theta) - Q^* + Q(\underline{\theta},\theta))^{j-1} (F(\theta) + Q^* - Q(\underline{\theta},\theta))^{I-2-j} u(H_{j+1}) \cdot$$
$$[(I-1)(1 - F(\theta) - Q^* + Q(\underline{\theta},\theta)) - j]$$
$$+ \sum_{j=n}^{I-1} \binom{I-1}{j} (1 - F(\theta) - Q^* + Q(\underline{\theta},\theta))^{j-1} (F(\theta) + Q^* - Q(\underline{\theta},\theta))^{I-2-j} u(O_H) \cdot$$
$$[(I-1)(1 - F(\theta) - Q^* + Q(\underline{\theta},\theta)) - j]$$
$$\mathcal{N}(\theta) \equiv \sum_{j=0}^{m-1} \binom{I-1}{j} (Q^* - Q(\underline{\theta},\theta))^{j-1} (1 - Q^* + Q(\underline{\theta},\theta))^{I-2-j} u(L_{j+1}) \cdot$$
$$[j - (I-1)(Q^* - Q(\underline{\theta},\theta))]$$
$$+ \sum_{j=m}^{I-1} \binom{I-1}{j} (Q^* - Q(\underline{\theta},\theta))^{j-1} (1 - Q^* + Q(\underline{\theta},\theta))^{I-2-j} u(O_L) \cdot$$
$$[j - (I-1)(Q^* - Q(\underline{\theta},\theta))].$$

Now, I prove the proposition. First, notice that $g(\theta) = 1$ for $\theta \in [\widehat{\theta}, \bar{\theta}]$ by **Proposition 24**. Therefore, it suffices to show that

$$\lim_{\theta \uparrow \widehat{\theta}} g(\theta) < 1.$$

Since

$$\lim_{\theta \uparrow \widehat{\theta}} g(\theta) = \lim_{\theta \uparrow \widehat{\theta}} \left(-\frac{\mathcal{N}(\theta)}{\mathcal{M}(\theta) - \mathcal{N}(\theta)} \right) = -\frac{\mathcal{N}(\widehat{\theta})}{\mathcal{M}(\widehat{\theta}) - \mathcal{N}(\widehat{\theta})},$$
$$\mathcal{M}(\widehat{\theta}) - \mathcal{N}(\widehat{\theta}) > 0, \text{ and } \mathcal{N}(\widehat{\theta}) < 0$$

, it is enough to show that $\mathcal{M}(\widehat{\theta}) > 0$. Note that $\mathcal{M}(\widehat{\theta})$ is the same as $\frac{\partial D(\widehat{\theta})}{\partial F(\widehat{\theta})}$ in **Claim 3**. Therefore, $\mathcal{M}(\widehat{\theta}) > 0$.

C.2.4 Proof of Proposition 26

Notice that $D(\underline{\theta}) = 0$ when $Q(\underline{\theta}, \widehat{\theta}) = Q^*$, and $D(\underline{\theta}) = E[\mathcal{H}] - E[\mathcal{L}] = 0$ if $Q(\underline{\theta}, \widehat{\theta}) = Q^* = \frac{m}{I}$. Hence, $\frac{m}{I} > (<)\ Q^*$ if $E[\mathcal{H}] > (<) E[\mathcal{L}]$ since $D(\underline{\theta})$ is an increasing function in $Q(\underline{\theta}, \widehat{\theta})$. Moreover, the high-paying career path is more (less) competitive than the low-paying career path if

$$\frac{I(1 - Q^*)}{n} > (<)\ \frac{IQ^*}{m}.$$

This condition is equivalent to $\frac{m}{I} > (<)\ Q^*$ when $I = n + m$.

C.2.5 Proof of Proposition 27

Here, I only prove the case when $I > n+m$. The proof for the case when $I = n+m$ is similar to this proof. The proof for non-existence of symmetric pure strategy equilibrium is similar to the proof of **Proposition 23**. Therefore, I skip the proof.

Again, consider a strategy profile $(\hat{s})$ of the following form:

- An agent with ability $\theta \geq \widehat{\theta}$ pursues the high-paying career with probability one.
- An agent with ability $\theta < \widehat{\theta}$ pursue the high-paying career with probability $g(\theta)$.

Without loss of generality, let's assume that $u(O_L) = u(\emptyset) = 0$ for brevity. Note that if an agent with ability $\widehat{\theta}$ pursues the high-paying career, her expected utility would be :

$$E[u(\cdot)|(\hat{s}_i(\widehat{\theta}) = \mathcal{H}, \hat{s}_{-i}(\theta_{-i}))] = \sum_{j=0}^{n-1} \binom{I-1}{j} (1-F(\widehat{\theta}))^j F(\widehat{\theta})^{I-1-j} u(H_{j+1})$$
$$+ \sum_{j=n}^{n+m-1} \sum_{i=0}^{n+m-j-1} \binom{I-1}{i,\ j,\ I-1-i-j} (1-F(\widehat{\theta}))^j \mathcal{Q}(\underline{\theta},\widehat{\theta})^i \mathcal{P}(\underline{\theta},\widehat{\theta})^{I-1-i-j} u(L_{i+j-n+1})$$
$$+ \sum_{j=n}^{n+m-1} \sum_{i=n+m-j}^{I-1-j} \binom{I-1}{i,\ j,\ I-1-i-j} (1-F(\widehat{\theta}))^j \mathcal{Q}(\underline{\theta},\widehat{\theta})^i \mathcal{P}(\underline{\theta},\widehat{\theta})^{I-1-i-j} u(\emptyset)$$
$$+ \sum_{j=n+m}^{I-1} (1-F(\widehat{\theta}))^j F(\widehat{\theta})^{I-1-j} u(\emptyset).$$

Clearly, if she participates in the other labor market, her expected utility would be $u(L_1)$.

Now, if an agent with ability θ, $L \leq \theta < \widehat{\theta}$, pursues the high-paying career, her expected utility has the following value:

$$
\begin{aligned}
&E[u(\cdot)|(\hat{s}_i(\theta) = \mathcal{H}, \hat{s}_{-i}(\theta_{-i}))] \\
&\quad = \sum_{j=0}^{n-1} \binom{I-1}{j} (1 - F(\widehat{\theta}) + \mathcal{P}(\theta,\widehat{\theta}))^j (F(\theta) + \mathcal{Q}(\theta,\widehat{\theta}))^{I-1-j} u(H_{j+1}) \\
&\quad + \sum_{j=n}^{n+m-1} \sum_{i=0}^{n+m-j-1} \binom{I-1}{i,\ j,\ I-1-i-j} (1 - F(\widehat{\theta}) + \mathcal{P}(\theta,\widehat{\theta}))^j \mathcal{Q}(\underline{\theta},\widehat{\theta})^i \mathcal{P}(\underline{\theta},\theta)^{I-1-i-j} \cdot \\
&\quad u(L_{i+j-n+1}) \\
&\quad + \sum_{j=n}^{n+m-1} \sum_{i=n+m-j}^{I-1-j} \binom{I-1}{i,\ j,\ I-1-i-j} (1 - F(\widehat{\theta}) + \mathcal{P}(\theta,\widehat{\theta}))^j \mathcal{Q}(\underline{\theta},\widehat{\theta})^i \mathcal{P}(\underline{\theta},\theta)^{I-1-i-j} u(\emptyset) \\
&\quad + \sum_{j=n+m}^{I-1} (1 - F(\widehat{\theta}) + \mathcal{P}(\theta,\widehat{\theta}))^j (F(\theta) + \mathcal{Q}(\theta,\widehat{\theta}))^{I-1-j} u(\emptyset).
\end{aligned}
$$

However, if she decides to participate in the other labor market, her expected utility would be

$$
\begin{aligned}
&E[u(\cdot)|(\hat{s}_i(\theta) = \mathcal{L}, \hat{s}_{-i}(\theta_{-i}))] \\
&\quad = \sum_{j=0}^{m-1} \binom{I-1}{j} \mathcal{Q}(\theta,\widehat{\theta})^j (1 - F(\widehat{\theta}) + \mathcal{P}(\theta,\widehat{\theta}) + F(\theta))^{I-1-j} u(L_{j+1}) \\
&\quad + \sum_{j=m}^{I-1} \binom{I-1}{j} \mathcal{Q}(\theta,\widehat{\theta})^j (1 - F(\widehat{\theta}) + \mathcal{P}(\theta,\widehat{\theta}) + F(\theta))^{I-1-j} u(\emptyset).
\end{aligned}
$$

Claim 12 *The cut-off ability $\widehat{\theta}$ is strictly greater than $\underline{\theta}$.*

Proof 29 *If $\widehat{\theta} = \underline{\theta}$, $D(\underline{\theta}) = u(\emptyset) - u(L_1) < 0$. Hence, $\widehat{\theta} > \underline{\theta}$.*

Claim 13 *There exists a unique value of $\mathcal{Q}(\underline{\theta},\widehat{\theta}) \in (0,1)$ such that $D(\underline{\theta}) = 0$.*

Proof 30 *For $\theta \in [\underline{\theta},\widehat{\theta})$, the difference between two actions is*

$$
\begin{aligned}
D(\theta) &= \sum_{j=0}^{n-1} \binom{I-1}{j} (1 - F(\widehat{\theta}) + \mathcal{P}(\theta,\widehat{\theta}))^j (F(\theta) + \mathcal{Q}(\theta,\widehat{\theta}))^{I-1-j} u(H_{j+1}) \\
&\quad + \sum_{j=n}^{n+m-1} \sum_{i=0}^{n+m-j-1} \binom{I-1}{i,\ j,\ I-1-i-j} (1 - F(\widehat{\theta}) + \mathcal{P}(\theta,\widehat{\theta}))^j \mathcal{Q}(\underline{\theta},\widehat{\theta})^i \mathcal{P}(\underline{\theta},\theta)^{I-1-i-j} \cdot \\
&\quad u(L_{i+j-\ell+1}) \\
&\quad - \sum_{j=0}^{m-1} \binom{I-1}{j} \mathcal{Q}(\theta,\widehat{\theta})^j (1 - F(\widehat{\theta}) + \mathcal{P}(\theta,\widehat{\theta}) + F(\theta))^{I-1-j} u(L_{j+1}).
\end{aligned}
$$

Consider the case when $\theta = \underline{\theta}$. Then,

$$D(\underline{\theta}) = \sum_{j=I-n}^{I-1} \binom{I-1}{j} Q(\underline{\theta},\widehat{\theta})^j (1 - Q(\underline{\theta},\widehat{\theta}))^{I-1-j} u(H_{n-j})$$
$$- \sum_{j=0}^{m-1} \binom{I-1}{j} Q(\underline{\theta},\widehat{\theta})^j (1 - Q(\underline{\theta},\widehat{\theta}))^{I-1-j} u(L_{j+1}).$$

Note that $D(\underline{\theta}) = -u(L_1) < 0$ when $Q(\underline{\theta},\widehat{\theta}) = 0$ and $D(\underline{\theta}) = u(H_1) > 0$ when $Q(\underline{\theta},\widehat{\theta}) = 1$. Also, for $Q(\underline{\theta},\widehat{\theta}) \in (0,1)$, the differential of this difference with respect to $Q(\underline{\theta},\widehat{\theta})$ is

$$\frac{\partial D(\underline{\theta})}{\partial Q(\underline{\theta},\widehat{\theta})} = \sum_{j=I-n}^{I-1} \binom{I-1}{j} Q(\underline{\theta},\widehat{\theta})^{j-1} (1 - Q(\underline{\theta},\widehat{\theta}))^{I-2-j} u(H_{I-j})[j - (I-1)Q(\underline{\theta},\widehat{\theta})]$$
$$- \sum_{j=0}^{m-1} \binom{I-1}{j} Q(\underline{\theta},\widehat{\theta})^{j-1} (1 - Q(\underline{\theta},\widehat{\theta}))^{I-2-j} u(L_{j+1})[j - (I-1)Q(\underline{\theta},\widehat{\theta})].$$

Define a pseudo-utility function $U(\cdot)$ by

$$U(i) = \begin{cases} u(H_{I-i}) & \text{if } I - n \le i \le I \\ 0 & \text{if } m - 1 < i < I - n \\ -u(L_{i+1}) & \text{if } 0 \le i \le m - 1. \end{cases}$$

Using this function, I can express $\frac{\partial D(\underline{\theta})}{\partial Q(\underline{\theta},\widehat{\theta})}$ as

$$\frac{\partial D(\underline{\theta})}{\partial Q(\underline{\theta},\widehat{\theta})} = \frac{1}{Q(\underline{\theta},\widehat{\theta})(1 - Q(\underline{\theta},\widehat{\theta}))} \Bigg[\sum_{j=0}^{I-1} \binom{I-1}{j} jU(j) Q(\underline{\theta},\widehat{\theta})^j (1 - Q(\underline{\theta},\widehat{\theta}))^{I-1-j}$$
$$-(I-1) \sum_{j=0}^{I-1} \binom{I-1}{j} U(j) Q(\underline{\theta},\widehat{\theta})^j (1 - Q(\underline{\theta},\widehat{\theta}))^{I-1-j} \Bigg]$$
$$> \frac{1}{Q(\underline{\theta},\widehat{\theta})(1 - Q(\underline{\theta},\widehat{\theta}))} \left\{ E[J]E[U(J)] - (I-1)Q(\underline{\theta},\widehat{\theta})E[U(J)] \right\}$$
$$= \frac{1}{Q(\underline{\theta},\widehat{\theta})(1 - Q(\underline{\theta},\widehat{\theta}))} \left[(I-1)Q(\underline{\theta},\widehat{\theta})E[U(J)] - (I-1)Q(\underline{\theta},\widehat{\theta})E[U(J)] \right]$$
$$= 0,$$

where

$$E[J] = \sum_{j=0}^{I-1} \binom{I-1}{j} j Q(\underline{\theta},\widehat{\theta})^j (1 - Q(\underline{\theta},\widehat{\theta}))^{I-1-j}$$

$$E[U(J)] = \sum_{j=0}^{I-1} \binom{I-1}{j} U(j) Q(\underline{\theta},\widehat{\theta})^j (1 - Q(\underline{\theta},\widehat{\theta}))^{I-1-j}.$$

The inequality holds since the function $U(j)$ is increasing function in j and $0 < \mathcal{P}(L,\widehat{\theta}) < 1$. This implies $E[JU(J)] > E[J]E[U(J)]$.

Therefore, there exists a unique value of $Q(\underline{\theta},\widehat{\theta}) \in (0,1)$ such that $D(\underline{\theta}) = 0$. From now on, I denote this value by Q_e^*.

Claim 14 *There exist a unique $\widehat{\theta}$ such that $D(\widehat{\theta}) = 0$ and $D(\underline{\theta}) = 0$.*

Proof 31 *I have already shown that $Q(\underline{\theta},\widehat{\theta})$ has a unique value Q_e^*. Therefore, it suffices to show that there exists a unique $\widehat{\theta}$ such that $D(\widehat{\theta}) = 0$ when $Q(\underline{\theta},\widehat{\theta}) = Q_e^*$. Note that*

$$D(\widehat{\theta}) = \sum_{j=I-n}^{I-1} \binom{I-1}{j} F(\widehat{\theta})^j (1 - F(\widehat{\theta}))^{I-1-j} u(H_{I-j})$$
$$+ \sum_{j=0}^{I-1-n} \sum_{i=0}^{I-1-n-j} \binom{I-1}{j,\ i,\ I-1-j-i} (F(\widehat{\theta}) - Q_e^*)^j (Q_e^*)^i (1 - F(\widehat{\theta}))^{I-1-j-i} \cdot$$
$$u(L_{I-n-j}) - u(L_1),$$

where $u(L_{I-n-j}) = 0$ if $I-n-j > m$. This difference has the value $u(H_1) - u(L_1) > 0$ when $F(\widehat{\theta}) = 1$, and

$$D(\widehat{\theta}) = \sum_{j=0}^{n-1} \binom{I-1}{j} (1 - Q_e^*)^j (Q_e^*)^{I-1-j} u(H_{j+1}) - u(L_1)$$
$$= \sum_{j=0}^{m-1} \binom{I-1}{j} (Q_e^*)^j (1 - Q_e^*)^{I-1-j} u(L_{j+1}) - u(L_1) < 0$$

when $F(\widehat{\theta}) = Q_e^*$.[2]

[2]Recall that Q_e^* is the solution to the equation $D(\underline{\theta}) = 0$.

The differential of this difference with respect to $F(\widehat{\theta})$ is

$$\frac{\partial D(\widehat{\theta})}{\partial F(\widehat{\theta})} = \sum_{j=I-n}^{I-1} \binom{I-1}{j} F(\widehat{\theta})^{j-1}(1-F(\widehat{\theta}))^{I-2-j} u(H_{I-j})[j-(I-1)F(\widehat{\theta})]$$

$$+ \sum_{j=0}^{I-n-1} \sum_{i=0}^{I-1-n-j} \binom{I-1}{j,\ i,\ I-1-j-i} (F(\widehat{\theta})-Q_e^*)^{j-1}(Q_e^*)^i(1-F(\widehat{\theta}))^{I-2-j-i}.$$

$$u(L_{I-n-j})[j(1-F(\widehat{\theta}))+(j+i)(F(\widehat{\theta})-Q_e^*)-(I-1)(F(\widehat{\theta})-Q_e^*)].$$

Now, I show that $\frac{\partial D(\widehat{\theta})}{\partial F(\widehat{\theta})} > 0$ for $F(\widehat{\theta}) \in (Q_e^*, 1)$. The differentiation can be expressed as

$$\frac{\partial D(\widehat{\theta})}{\partial F(\widehat{\theta})} = \sum_{j=I-n}^{I-1} \binom{I-1}{j} F(\widehat{\theta})^{j-1}(1-F(\widehat{\theta}))^{I-2-j} u(H_{I-j})[j-(I-1)F(\widehat{\theta})]$$

$$+ \sum_{j=0}^{I-n-1} \binom{I-1}{j} \frac{F(\widehat{\theta})^j(1-F(\widehat{\theta}))^{I-1-j}}{(F(\widehat{\theta})-Q_e^*)(1-F(\widehat{\theta}))} \sum_{i=0}^{j} \binom{j}{i} \left(\frac{F(\widehat{\theta})-Q_e^*}{F(\widehat{\theta})}\right)^i \left(\frac{Q_e^*}{F(\widehat{\theta})}\right)^{j-i}.$$

$$u(L_{I-n-i},\widehat{\theta})[i(1-F(\widehat{\theta}))+j(F(\widehat{\theta})-Q_e^*)-(I-1)(F(\widehat{\theta})-Q_e^*)]$$

$$> \sum_{j=I-n}^{I-1} \binom{I-1}{j} F(\widehat{\theta})^{j-1}(1-F(\widehat{\theta}))^{I-2-j} u(H_{I-j},\widehat{\theta})[j-(I-1)F(\widehat{\theta})]$$

$$+ \sum_{j=0}^{I-1-n} \binom{I-1}{j} \frac{F(\widehat{\theta})^j(1-F(\widehat{\theta}))^{I-1-j}}{(F(\widehat{\theta})-Q_e^*)(1-F(\widehat{\theta}))} \{(1-F(\widehat{\theta}))E[A_j]E[u(A_j)]$$

$$+ j(F(\widehat{\theta})-Q_e^*)E[u(A_j)] - (I-1)(F(\widehat{\theta})-Q_e^*)E[u(A_j)]\}$$

$$= \frac{1}{F(\widehat{\theta})(1-F(\widehat{\theta}))} \left[\sum_{j=0}^{I-1} \binom{I-1}{j} jF(\widehat{\theta})^j(1-F(\widehat{\theta}))^{I-1-j} U_e(j)\right.$$

$$\left. -(I-1)F(\widehat{\theta}) \sum_{j=0}^{I-1} \binom{I-1}{j} F(\widehat{\theta})^j (1-F(\widehat{\theta}))^{I-1-j} U_e(j,\widehat{\theta})\right]$$

$$> 0,$$

where

$$E[A_j] = \binom{j}{i} j \left(\frac{F(\widehat{\theta}) - Q_e^*}{F(\widehat{\theta})}\right)^i \left(\frac{Q_e^*}{F(\widehat{\theta})}\right)^{j-i},$$

$$E[u(A_j)] = \binom{j}{i} \left(\frac{F(\widehat{\theta}) - Q_e^*}{F(\widehat{\theta})}\right)^i \left(\frac{Q_e^*}{F(\widehat{\theta})}\right)^{j-i} u(L_{I-n-i}), \text{ and}$$

$$U_e(j) = \begin{cases} u(H_{I-j}, \theta) & \text{if } j \geq I - n \\ E[u(A_j)] & \text{otherwise} \end{cases}.$$

The last inequality holds since $U_e(j)$ is a increasing function in j and $0 < F(\widehat{\theta}) < 1$.

Therefore, the differential of the difference is strictly greater than 0. Hence, there exist a unique $\widehat{\theta}$ and $Q(\underline{\theta}, \widehat{\theta})$ such that $D(\widehat{\theta}) = 0$ and $D(\underline{\theta}) = 0$.

Claim 15 *For given $\widehat{\theta}$ and $Q(\underline{\theta}, \widehat{\theta})$, there is a unique $Q(\underline{\theta}, \theta)$ satisfying $D(\theta) = 0$ for $\theta \in (\underline{\theta}, \widehat{\theta})$.*

Proof 32 *Consider an agents with ability $\theta \in (\underline{\theta}, \widehat{\theta}^*)$. Recall that*

$$\begin{aligned} D(\theta) = &\sum_{j=I-n}^{I-1} \binom{I-1}{j} (F(\theta) + Q_e^* - Q(\underline{\theta}, \theta))^j (1 - F(\theta) - Q_e^* + Q(\underline{\theta}, \theta))^{I-1-j} u(H_{I-j}) \\ &+ \sum_{j=0}^{I-1-n} \sum_{i=0}^{I-1-n-j} \binom{I-1}{j,\ i,\ I-1-j-i} (F(\theta) - Q(\underline{\theta}, \theta))^j (Q_e^*)^i \cdot \\ &(1 - F(\theta) - Q_e^* + Q(\underline{\theta}, \theta))^{I-1-j-i} u(L_{I-n-j}) \\ &- \sum_{j=0}^{m-1} \binom{I-1}{j} (Q_e^* - Q(\underline{\theta}, \theta))^j (1 - Q_e^* + Q(\underline{\theta}, \theta))^{I-1-j} u(L_{j+1}). \end{aligned}$$

Now, I show that there is a unique $Q(\underline{\theta}, \theta)$ such that $D(\theta) = 0$. First, notice that $\max\{0, Q_e^* - F(\widehat{\theta}) + F(\theta)\} \leq Q(\underline{\theta}, \theta) \leq \min\{F(\theta), Q_e^*\}$ from $\mathcal{P}(\underline{\theta}, \theta) + Q(\underline{\theta}, \theta) = F(\theta)$ and $Q(\underline{\theta}, \theta) + Q(\theta, \widehat{\theta}) = Q_e^*$.

The difference $D(\theta)$ has a positive value when $Q(\underline{\theta}, \theta) = \max\{0, Q_e^* - F(\widehat{\theta}) + F(\theta)\}$, and a negative value when $Q(\underline{\theta}, \theta) = \min\{F(\theta), Q_e^*\}$.

- *$Q(\underline{\theta}, \theta) = 0$: From $D(\underline{\theta}) = 0$,*

$$
\begin{aligned}
D(\theta) = & \sum_{j=I-n}^{I-1} \binom{I-1}{j} (F(\theta) + Q_e^*)^j (1 - F(\theta) - Q_e^*)^{I-1-j} u(H_{I-j}) \\
& + \sum_{j=0}^{I-1-n} \sum_{i=0}^{I-1-n-j} \binom{I-1}{j,\ i,\ I-1-j-i} F(\theta)^j (Q_e^*)^i \cdot \\
& (1 - F(\theta))^{I-1-j-i} u(L_{I-n-j}) \\
& - \sum_{j=I-n}^{I-1} \binom{I-1}{j} (Q_e^*)^j (1 - Q_e^*)^{I-1-j} u(H_{I-j}) \\
& > 0.
\end{aligned}
$$

The inequality holds since the second summation is greater than 0 and the first summation is strictly greater than the last summation because $F(\theta) > 0$.

- $Q(\underline{\theta}, \theta) = Q_e^* - F(\widehat{\theta}) + F(\theta)$*: From* $D(\widehat{\theta}) = 0$*,*

$$
\begin{aligned}
D(\theta) = u(L_1) - & \sum_{j=0}^{m-1} \binom{I-1}{j} (F(\widehat{\theta}) - F(\theta))^j (1 - F(\widehat{\theta}) + F(\theta))^{I-1-j} u(L_{j+1}) \\
> 0. &
\end{aligned}
$$

- $Q(\underline{\theta}, \theta) = F(\theta)$*: From* $D(\underline{\theta}) = 0$*,*

$$
\begin{aligned}
D(\theta) = & \sum_{j=0}^{m-1} \binom{I-1}{j} (Q_e^*)^j (1 - Q_e^*)^{I-1-j} u(L_{j+1}) \\
& - \sum_{j=0}^{m-1} \binom{I-1}{j} (Q_e^* - F(\theta))^j (1 - Q_e^* + F(\theta))^{I-1-j} u(L_{j+1}) \\
& < 0.
\end{aligned}
$$

The inequality holds since $F(\theta) > 0$.

- $Q(\underline{\theta}, \theta) = Q_e^*$*: From* $D(\widehat{\theta}) = 0$*,*

$$
\begin{aligned}
D(\theta) = & \sum_{j=I-n}^{I-1} \binom{I-1}{j} F(\theta)^j (1 - F(\theta))^{I-1-j} u(H_{I-j}) \\
& + \sum_{j=0}^{I-1-n} \sum_{i=0}^{I-1-n-j} \binom{I-1}{j,\ i,\ I-1-j-i} (F(\theta) - Q_e^*)^j (Q_e^*)^i \cdot \\
& (1 - F(\theta))^{I-1-j-i} u(L_{I-n-j}) - u(L_1) < 0
\end{aligned}
$$

since $F(\theta) < F(\widehat{\theta})$.

Remaining proof is to show that $\partial D(\theta)/\partial Q(\underline{\theta},\theta) < 0$. *Define* $S_e(\theta) := Q_e^* - Q(\underline{\theta},\theta)$. *Then, it suffices to show that* $\partial D(\theta)/\partial S_e(\theta) > 0$. *The differentiation of the difference* $D(\theta)$ *with respect to* $S_e(\theta)$ *is*

$$\frac{\partial D(\theta)}{\partial S_e(\theta)} = \sum_{j=I-n}^{I-1} \binom{I-1}{j} (F(\theta) + S_e(\theta))^{j-1} (1 - F(\theta) - S_e(\theta))^{I-2-j} u(H_{I-j}) \cdot$$

$$[j - (I-1)(F(\theta) + S_e(\theta))]$$

$$+ \sum_{j=0}^{I-1-n} \sum_{i=0}^{I-1-n-j} \binom{I-1}{j,\ i,\ I-1-j-i} (F(\theta) - Q_e^* + S_e(\theta))^{j-1} (Q_e^*)^i \cdot$$

$$(1 - F(\theta) - S_e(\theta))^{I-2-j-i} u(L_{I-n-j})$$

$$[j(1 - F(\theta) - S_e(\theta)) + (j+i)(F(\theta) - Q_e^* + S_e(\theta)) - (I-1)(F(\theta) - Q_e^* + S_e(\theta))]$$

$$- \sum_{j=0}^{m-1} \binom{I-1}{j} S_e(\theta)^{j-1} (1 - S_e(\theta))^{I-2-j} u(L_{j+1}) [j - (I-1) S_e(\theta)]$$

$$= \sum_{j=I-n}^{I-1} \binom{I-1}{j} (F(\theta) + S_e(\theta))^{j-1} (1 - F(\theta) - S_e(\theta))^{I-2-j} u(H_{I-j}) \cdot$$

$$[j - (I-1)(F(\theta) + S_e(\theta))]$$

$$+ \sum_{j=0}^{I-1-n} \binom{I-1}{j} \frac{(F(\theta) + S_e(\theta))^j (1 - F(\theta) - S_e(\theta))^{I-1-j}}{(F(\theta) - Q_e^* + S_e(\theta))(1 - F(\theta) - S_e(\theta))} \cdot$$

$$\sum_{i=0}^{j} \binom{j}{i} \left(\frac{F(\theta) - Q_e^* + S_e(\theta)}{F(\theta) + S_e(\theta)} \right)^i \left(\frac{Q_e^*}{F(\theta) + S_e(\theta)} \right)^{j-i} u(L_{I-n-i}) \cdot$$

$$[i(1 - F(\theta) - S_e(\theta)) + j(F(\theta) - Q_e^* + S_e(\theta)) - (I-1)(F(\theta) - Q_e^* + S_e(\theta))]$$

$$- \sum_{j=0}^{m-1} \binom{I-1}{j} S_e(\theta)^{j-1} (1 - S_e(\theta))^{I-2-j} u(L_{j+1}) [j - (I-1) S_e(\theta)]$$

$$> \frac{1}{(F(\theta) + S_e(\theta))(1 - F(\theta) - S_e(\theta))} \cdot$$

$$\left[\sum_{j=0}^{I-1} \binom{I-1}{j} j (F(\theta) + S_e(\theta))^j (1 - F(\theta) - S_e(\theta))^{I-1-j} \widetilde{U}_e(j) \right.$$

$$\left. - (I-1)(F(\theta) + S_e(\theta)) \sum_{j=0}^{I-1} \binom{I-1}{j} (F(\theta) + S_e(\theta))^j (1 - F(\theta) - S_e(\theta))^{I-1-j} \widetilde{U}_e(j) \right]$$

$$- \sum_{j=0}^{m-1} \binom{I-1}{j} S_e(\theta)^{j-1} (1 - S_e(\theta))^{I-2-j} u(L_{j+1}) [j - (I-1) S_e(\theta)],$$

where

$$\widetilde{U}_e(j) = \begin{cases} u(H_{I-j}) & \text{if } j \geq I-n \\ E[u(B_j)] & \text{otherwise} \end{cases} \quad \text{with}$$

$$E[u(B_j)] = \sum_{i=0}^{j} \binom{j}{i} \left(\frac{F(\theta) - Q_e^* + S_e(\theta)}{F(\theta) + S_e(\theta)}\right)^i \left(\frac{Q_e^*}{F(\theta) + S_e(\theta)}\right)^{j-i} u(L_{I-n-i}).$$

The summation

$$\Bigg[\sum_{j=0}^{I-1} \binom{I-1}{j} j(F(\theta) + S_e(\theta))^j (1 - F(\theta) - S_e(\theta))^{I-1-j} \widetilde{U}_e(j)$$

$$- (I-1)(F(\theta) + S_e(\theta)) \sum_{j=0}^{I-1} \binom{I-1}{j} (F(\theta) + S_e(\theta))^j (1 - F(\theta) - S_e(\theta))^{I-1-j} \widetilde{U}_e(j)\Bigg]$$

is strictly greater than zero since $\widetilde{U}_e(j)$ *is a increaing function in* j *and* $0 < F(\theta) + S_e(\theta) < 1$. *Also, the summation*

$$\sum_{j=0}^{m-1} \binom{I-1}{j} S_e(\theta)^{j-1} (1 - S_e(\theta))^{I-2-j} u(L_{j+1})[j - (I-1)S_e(\theta)]$$

is strictly less than zero since $u(L_{j+1})$ *is a strictly decreasing function in* j *and* $0 < S_e(\theta) < 1$.

Hence, $\frac{\partial D(\theta)}{\partial Q(\underline{\theta}, \theta)} < 0$.

Therefore, by the *Lebesgue Differentiation Theorem*, the symmetric equilibrium strategy $g(\theta)$ is unique among integrable functions.

C.2.6 Proof of Proposition 28

Note that $\widehat{\theta}$ satisfies :

$$D(\widehat{\theta}) = \sum_{j=0}^{n-1} \binom{I-1}{j} (1 - F(\widehat{\theta}))^j F(\widehat{\theta})^{I-1-j} u(H_{j+1})$$

$$+ \sum_{j=n}^{I-1} \binom{I-1}{j} (1 - F(\widehat{\theta}))^j F(\widehat{\theta})^{I-1-j} u(O_H) - u(L_1) = 0,$$

and $\widehat{\theta}_e$ satisfies :

$$D(\widehat{\theta}_e) = \sum_{j=I-n}^{I-1} \binom{I-1}{j} F(\widehat{\theta}_e)^j (1-F(\widehat{\theta}_e))^{I-1-j} u(H_{I-j})$$
$$+ \sum_{j=0}^{I-1-n} \sum_{i=0}^{I-1-n-j} \binom{I-1}{j,\ i,\ I-1-j-i} (F(\widehat{\theta}_e) - Q_e^*)^j (Q_e^*)^i (1-F(\widehat{\theta}_e))^{I-1-j-i} \cdot$$
$$u(L_{I-n-j}) - u(L_1)$$
$$= \sum_{j=I-n}^{I-1} \binom{I-1}{j} F(\widehat{\theta}_e)^j (1-F(\widehat{\theta}_e))^{I-1-j} u(H_{I-j})$$
$$+ \sum_{j=0}^{I-1-n} \binom{I-1}{j} F(\widehat{\theta}_e)^j (1-F(\widehat{\theta}))^{I-1-j} E[u(A_j)] - u(L_1) = 0.$$

Since

$$0 < \sum_{j=0}^{I-1-n} \binom{I-1}{j} F(\widehat{\theta}_e)^j (1-F(\widehat{\theta}_e))^{I-1-j} E[u(A_j)]$$
$$< \sum_{j=0}^{I-1-n} \binom{I-1}{j} F(\widehat{\theta}_e)^j (1-F(\widehat{\theta}_e))^{I-1-j} u(H_n),$$

there exists a unique $0 < \mathcal{K} < u(H_n)$ such that

$$\sum_{j=0}^{I-1-n} \binom{I-1}{j} F(\widehat{\theta}_e)^j (1-F(\widehat{\theta}_e))^{I-1-j} E[u(A_j)]$$
$$= \sum_{j=0}^{I-1-n} \binom{I-1}{j} F(\widehat{\theta}_e)^j (1-F(\widehat{\theta}_e))^{I-1-j} \mathcal{K}.$$

Hence, if $u(O_H) = \mathcal{K}$, $D(\widehat{\theta}) = D(\widehat{\theta}_e) = 0$ with $\widehat{\theta} = \widehat{\theta}_e$. Moreover, if $u(O_H) > (<)\ \mathcal{K}$, $D(\widehat{\theta}) > (<)\ D(\widehat{\theta}_e) = 0$ when $\widehat{\theta} = \widehat{\theta}_e$. Since $D(\widehat{\theta})$ is a strictly increasing function in $\widehat{\theta}$, the equilibrium cut-off ability $\widehat{\theta} < (>)\ \widehat{\theta}_e$ if $u(O_H) > (<)\ \mathcal{K}$.