Optimal Agents

Jan Starmans

November 5, 2017

Abstract

This paper studies a principal’s hiring decision when different agents generate different probability distributions of output under effort, subject to moral hazard. Contracting is subject to canonical frictions: Agents enjoy limited liability and can manipulate output ex post. The main insight is that the contracting problem determines not only optimal contract design but also the type of agent the principal hires. Different but equally productive agents require different optimal contracts, implying different agency rents. This generates a pecking order among agents with the same productivity. Moreover, the contracting problem can bias the principal towards hiring less productive agents. The results suggest a novel link between incentives and hiring, with implications for firms’ hiring decisions, the level, shape, and dispersion of incentive pay, human capital formation, the choice of corporate strategy, delegation, and firms’ production technologies.

Keywords: Principal-agent theory, contract theory, contractual frictions, hiring.

JEL Classifications: C72, D82, D86, J31, J33, J41, M55.

*I thank Tyler Abbot, Charles Angelucci, Taylor Begley, Bo Bian, Patrick Bolton, Olivier Darmouni, James Dow, Alex Edmans, Maryam Farboodi, Francisco Gomes, Denis Gromb, Alexander Guembel, Deeksha Gupta, Marina Halac, Christopher Hennessy, Ralph Kojen, Lukas Kremens, Howard Kung, Augustin Landier, Ye Li, Jean-Marie Meier, Anna Pavlova, Tomasz Piskorski, Andrea Prat, Kunal Sachdeva, José Scheinkman, David Schoenherr, Janis Skrastins, Vikrant Vig, seminar participants at London Business School and Princeton University, and participants at the 2nd HEC Finance Ph.D. Workshop for their many helpful comments and suggestions. All errors are my responsibility.

†London Business School, Regent’s Park, London, NW1 4SA, United Kingdom. Email: jstarmans@london.edu. Phone: +44 (0)75 6555 7242.
1 Introduction

The contracting literature studies optimal incentive compensation for agents such as CEOs.\(^1\) However, it takes as given the characteristics of the agent and largely bypasses the question of which agent a firm should hire in the first place, taking into account the implications of agent characteristics for the contracting problem. For example, which type of CEO, in terms of characteristics such as experience and education (see, e.g., Bertrand and Schoar, 2003), should a firm hire, anticipating the incentive and compensation problem?

This paper develops a joint theory of incentive compensation and hiring decisions. It studies a principal’s hiring decision when different agents generate different probability distributions of output under effort, subject to moral hazard. Contracting is subject to canonical frictions: Agents enjoy limited liability and can manipulate output ex post. The principal must thus decide which type of agent to hire and design the agent’s incentive contract.

Agents can generate different probability distributions of output, which reflect differences in employee characteristics such as experience and education. For example, a bank considers two CEO candidates. One candidate with a law background can implement a risk management strategy reducing litigation risk, which increases the probability of medium performance. The other candidate with a business administration background can implement an innovation strategy investing in fintech, which increases the probability of high performance.\(^2\) The bank must decide which candidate to hire and design his/her incentive contract.

The main insight of this paper is that the contracting problem determines not only optimal contract design, but also the type of agent the principal hires. Even if agents have the same productivity, they would receive different optimal contracts, implying different agency rents. This generates a pecking order among agents with the same productivity, and the principal hires the agent with the technology that requires the lowest rent. Moreover, the contracting problem can bias the principal towards hiring less productive agents. The results suggest a novel link between

\(^1\)For CEO compensation, see Frydman and Jenter (2010), Murphy (2013), and Edmans and Gabaix (2016).

\(^2\)For example, Brian Moynihan, CEO of Bank of America, majored in history at Brown University and earned a Juris Doctor from the University of Notre Dame Law School. He dealt with several large litigation cases at Bank of America. José Antonio Álvarez, CEO of Banco Santander, holds a degree in Business Economics from the University of Santiago de Compostela in Spain and the University of Chicago, and an MBA from the University of Chicago’s Graduate School of Business. Santander is one of the most active investors in fintech in the European banking industry.
incentives and hiring, with implications for firms’ hiring decisions, the level, shape, and dispersion of incentive pay, human capital formation, the choice of corporate strategy, delegation, and firms’ production technologies.

I consider the following setting: A risk-neutral principal owns a project that generates a random cash flow from a finite set of possible cash flows. Without an agent’s effort, each cash flow has a positive probability. Since the principal obtains this cash flow distribution without an agent’s effort, I call it the principal’s technology. The principal can hire a risk-neutral agent, whose effort increases the project’s expected cash flow. Different agents generate different probability distributions of cash flows under effort, referred to as agents’ technologies. Specifically, I consider agents with technologies that first-order stochastically dominate the principal’s technology with a single-peaked likelihood ratio\(^3\), a generalization of a monotone likelihood ratio.

To isolate the paper’s novel contribution, I assume that all agents generate the same increase in the project’s expected cash flow through effort (expected value of effort), incur the same disutility of effort (cost of effort), and have the same reservation utility equal to zero. In particular, agents generate the same expected surplus through effort. In an extension, I introduce differences in expected surplus across agents, in addition to differences in agents’ technologies.

Effort is subject to moral hazard and the principal can offer an agent a contract designed to induce effort. In the absence of contractual frictions, the principal would capture the full expected surplus and thus be indifferent between equally productive agents. However, I assume that contracting is subject to frictions: Agents are protected by limited liability, and they have the ability to “secretly destroy” cash flows and “secretly borrow” to inflate cash flows ex post (see, e.g., Innes, 1990). As a result, contractual payments to an agent must satisfy monotonicity constraints, in that they have to be nondecreasing in cash flows and cannot increase more than one-to-one with cash flows.

As a first step, I characterize the principal’s optimal contract for an arbitrary agent. As a benchmark case, I first focus on limited liability, ignoring the monotonicity constraints. As is standard, the agent earns a rent, and in the optimal contract, the agent’s compensation is non-zero only in the state with the maximum likelihood ratio. Indeed, each state having a positive probability

\(^3\)The likelihood ratio is the change in probability due to effort divided by the no-effort probability in each state.
under the principal’s technology, a payment in any state gives the agent a positive expected payoff if he shirks, which determines the agent’s equilibrium rent. The state with the maximum likelihood ratio has the highest incentive effect per unit of rent. Paying the agent in this state thus minimizes the agent’s rent. In particular, a higher maximum likelihood ratio implies a lower agency rent.

Next, I consider the full contracting problem with limited liability and monotonicity constraints. I show that the optimal contract is a capped bonus contract (or junior debt), that is, the agent’s compensation is zero if the cash flow falls below a first threshold, increases one-to-one with cash flows between the first and a second threshold, and remains flat beyond the second threshold. The thresholds depend on the agent’s technology, and different agents receive different optimal contracts if hired.

I then characterize the principal’s hiring decision. Despite being equally productive, different agents require different optimal contracts leading to different rents, and the principal hires the agent with the technology that requires the lowest rent. I show that the pecking order among agents is determined by the productivity of effort—the ratio of the expected value and the cost of effort—, which I have assumed is common to all agents. Consider an example with three states, cash flows 0, 1, and 2, and two agents. Agent 1’s effort increases the probability of cash flow 1, agent 2’s increases that of cash flow 2. The principal must pay agent 1 in state 1 to induce effort. Since contractual payoffs have to be nondecreasing in cash flows, the principal must pay agent 1 the same level in state 2 as in state 1, implying an increase in the agent’s rent. If the productivity of effort is high, the principal pays agent 2 a small amount only in state 2, with no further binding constraint. If both agents earn identical rents in the benchmark without monotonicity constraints, agent 1 requires a higher rent and the principal hires agent 2. In contrast, if the productivity of effort is low, the principal must pay agent 2 a high share of the cash flow in state 2 to induce effort. Since contractual payoffs cannot increase more than one-to-one with cash flows, the principal must also pay the agent a high share of the cash flow in state 1, implying an increase in the agent’s rent. Since the principal must pay agent 1 only the same level in state 2 as in state 1, but a lower share, agent 2 requires a higher rent if the productivity of effort is sufficiently low and the principal hires agent 1.

I solve the principal’s hiring decision explicitly in two classes of agents’ technologies. Specif-
ically, I consider the two subsets of agents for which the optimal contract is debt and (levered) equity, respectively. Intuitively, these are the agents who increase the probability of states mainly in the low and the high tail of the cash flow distribution, respectively. I show that in each of these two subsets of agents, a unique optimal agent exists. If the optimal agents have the same maximum likelihood ratio, there exists a threshold such that the principal hires the optimal equity agent if productivity is above the threshold but the optimal debt agent if productivity is below the threshold. The result reflects the fact that debt contracts are more constrained by the frictions compared to (levered) equity contracts if productivity is high, but less constrained if productivity is low.

The insights extend to the general case in which agents receive optimal contracts other than pure debt and pure (levered) equity. I show that an agent’s optimal contract can be characterized as more debt-like if the agent affects the probability of cash flow states only up to a certain threshold state but does not affect the probability of higher states, and as more equity-like if this threshold state is higher. Next, I show that if the principal’s hiring decision is between a more debt- and a more equity-like agent, she hires the more debt-like agent if the productivity of effort is low.

The results have a number of implications, which are summarized here and detailed in Section 6. The results show that employees’ technologies affect firms’ hiring decisions, even if they have the same productivity, implying a novel link between incentives and hiring. Further, the legal and institutional environment affects the contracting problem and thus firms’ hiring decisions, for example by affecting a firm’s ability to prevent employees manipulating output. In an extension to agents with different productivities, I show that the contracting problem can bias the principal towards hiring less productive agents, which reduces welfare. Due to the differences in technologies, a less productive agent can require a significantly lower agency rent than a more productive agent, implying an overall higher expected utility for the principal.

A further implication is that employees with the same productivity can receive different contracts and expected compensation (i.e., rents). This implies a novel theory of contract and wage dispersion (Starmans, 2017a,b)\(^4\) that captures the broad prevalence and dispersion of incentive pay in employment relationships (Lemieux et al., 2009) and the heterogeneity in employees’ technologies due to the substantial specialization of labor (Becker and Murphy, 1992). In particular, the

\(^4\)The two papers develop a theory of contract and rent dispersion in a frictional labor market and study the implications for allocations, unemployment, and welfare.
model implies that compensation should be more (levered) equity-like (e.g., an option) if the productivity of effort is high, and more debt-like (e.g., a capped bonus) if the productivity of effort is low. Interpreted as managers compared to rank and file employees, or as managers at firms with more valuable investment opportunities compared to firms with less valuable investment opportunities, the prediction is consistent with evidence on compensation practices in firms.\(^5\)

The analysis also has implications for human capital formation. Different technologies can reflect differences in education, experience, and other characteristics. These differences affect firms’ hiring decisions even in the absence of differences in productivity, in turn affecting decisions about investment in human capital and the design of educational institutions.

Different technologies can reflect different corporate strategies, which can be implemented by different types of managers, for example a risk management strategy compared to an innovation strategy. The general insight is that some corporate strategies are more costly to implement than others, even if they generate the same expected value. This has broader implications for delegation. Contracting frictions can make it more costly to delegate some tasks than others. In particular, it may be possible to delegate some tasks, while delegating others with the same expected value might be too costly.

Moreover, since different technologies lead to different probability distributions of output, the contracting problem determines the firms’ equilibrium stochastic production technology.

**Related Literature.** There is a significant literature on optimal incentive and financial contracting, which takes the nature of the production technology as given. As such, this literature does not address the choice regarding agents. For example, Innes (1990) studies a single agent with a fixed technology and shows that levered equity is the optimal incentive contract. Poblete and Spulber (2012) show that this result is not robust to different technologies. In contrast, I develop a joint theory of incentive compensation and hiring decisions in the presence of contractual frictions when agents have different technologies. Specifically, I fully characterize the principal’s and agents’ expected utilities under optimal contracts, characterize the principal’s hiring decision

\(^5\)Frydman and Jenter (2010) and Murphy (2013) document the importance of options in executive compensation. Rank and file employees often receive more debt-like incentive pay such as capped bonuses (see, e.g., Lemieux et al., 2009). Convex incentives such as options are more prevalent in new economy firms and firms with more valuable growth opportunities (Guay, 1999; Ittner et al., 2003; Murphy, 2003).
and how it varies across productivity levels, determine the hiring decision in two classes of technologies, and discuss the implications for firms’ hiring decisions, the level, shape, and dispersion of incentive pay, human capital formation, the choice of corporate strategy, delegation, and firms’ production technologies. Further, Hébert (2016) studies an agent affecting the output distribution through effort and risk shifting. In my model, the principal chooses the output distribution.

This paper is related to the literature on information systems, in particular Blackwell (1951, 1953), Holmstrom (1979), Grossman and Hart (1983), and Kim (1995). While the authors study the principal’s response to changes in the agent’s information system, they study the risk/incentive trade-off, which is absent in my model. In contrast, I study the principal’s hiring decision in the presence of canonical contractual frictions: limited liability and ex post moral hazard.

The literature on firms’ hiring decisions focuses largely on firm-employee matching in terms of productivity (Oyer and Schaefer, 2011). For example, in Lazear (2009), workers differ in two dimensions of skills, and a worker’s productivity in a firm depends on how the firm weights these skills. Several papers study CEO-firm matching theories with different CEO and firm characteristics, for example, Gabaix and Landier (2008), Terviö (2008), Edmans and Gabaix (2011), and Eisfeldt and Kuhnen (2013). In the context of principal-agent models, a number of papers study agent selection along dimensions other than productivity. For example, Legros and Newman (1996), Thiele and Wambach (1999), Newman (2007), and Chade and Vera de Serio (2014) study agent selection based on agents’ wealth, and Lewis and Sappington (1991), Sobel (1993), Silvers (2012), and von Thadden and Zhao (2012) consider the agent’s information set. Other papers study the principal’s choice between agents where the main friction is adverse selection (see, e.g., Faynzilberg and Kumar, 1997; Lewis and Sappington, 2000, 2001), which is absent in my setting. I contribute to the literature by studying agent selection based on differences in agents’ technologies, in the presence of canonical contractual frictions.

The paper proceeds as follows. Section 2 introduces the theoretical framework. Section 3 studies optimal contracts. The main contribution of the paper is to study the resulting agency rents in Section 4 and the principal’s hiring decision in Section 5. Section 6 discusses the empirical and theoretical implications of the model. All proofs can be found in Appendix A.
2 Principal-Agent Framework

2.1 Model

There are three dates $t \in \{0, 1, 2\}$ and no time discounting. The risk-neutral principal (referred to as \textit{she}) owns a project, which generates a cash flow $x_i \in \mathbb{R}_+ := [0, \infty)$ in state $i \in \Omega := \{0, \ldots, n\}$ at $t = 2$, where $0 = x_{0} < x_{1} < \cdots < x_{n}$ and $n \geq 2$.\footnote{The finite state space simplifies the analysis, but is not necessary to derive optimal contracts and agency rents.} Without an agent’s effort ($e = 0$), the cash flow $x$ is drawn according to the probability distribution $q_i = \mathbb{P}(x = x_i | e = 0) > 0$, $i \in \Omega$. Denote the probability measure by $q$. Since the principal can generate this cash flow distribution without an agent’s effort, I refer to $q$ as the principal’s technology.

The principal can hire a single agent (referred to as \textit{he}) from a set of risk-neutral agents at $t = 0$.\footnote{Appendix C discusses the case of risk-averse agents.} The agent hired at $t = 0$ chooses whether or not to exert effort $e \in \{0, 1\}$ at $t = 1$, which is not verifiable.\footnote{The main insight that the contracting problem determines not only optimal contract design but also the type of agent the principal hires, even if agents have the same productivity, extends to a model with continuous effort. As shown with binary effort, differences in agents’ technologies lead to differences in endogenous agency rents. With continuous effort, there is an additional effect of agents’ technologies on the endogenous effort level, which generally differs across agents with different technologies.} Agents have the same disutility of effort $c \geq 0$ (\textit{cost of effort}), which is noncontractible, and the same reservation utility equal to zero. Agents differ in that their effort leads to different cash flow distributions, the agents’ technologies. Specifically, denote the set of agents’ technologies by $\mathcal{P} \subset \left\{ p \in [0, 1]^{n+1} \big| \sum_{i=0}^{n} p_i = 1 \right\}$. If agent $p \in \mathcal{P}$ exerts effort ($e = 1$), the cash flow is drawn according to the agent’s technology $p$, that is, $p_i = \mathbb{P}(x = x_i | e = 1)$, $i \in \Omega$. If he does not exert effort ($e = 0$), it is drawn according to the principal’s technology $q$. I next describe the set of agents’ technologies $\mathcal{P}$.

\textbf{Definition 1.} Consider a probability measure $p$. The likelihood ratio $l(p) = (l_i(p))_{i \in \Omega} \in \mathbb{R}^{n+1}$ is defined as follows:

$$l_i(p) := \frac{p_i - q_i}{q_i}, \quad i \in \Omega.$$  

Denote the maximum of the likelihood ratio by $l^*(p) := \max_{i \in \Omega} l_i(p)$.

\textbf{Definition 2.} Consider a probability measure $p$. The likelihood ratio $l(p)$ is called single-peaked if there exists a state $m \in \Omega$, such that $m \in \arg\max_{i \in \Omega} l_i(p)$, for all $i \leq m$, $l(p)$ is nondecreasing in
Assumption 1. For all \( p \in \mathcal{P} \), \( p \) first-order stochastically dominates \( q \).

Assumption 2. For all \( p \in \mathcal{P} \), the likelihood ratio \( l(p) \) is single-peaked.

Assumptions 1 and 2 are a generalization of a monotone likelihood ratio. A monotone likelihood ratio peaks in state \( n \), which implies first-order stochastic dominance. I allow the likelihood ratio to peak in any state while maintaining first-order stochastic dominance. Differences in agents’ technologies reflect exogenous or endogenous differences in abilities, talents, education, experience, and other characteristics,\(^9\) as illustrated in the bank manager example in the introduction.

Assumption 3. Each agent \( p \in \mathcal{P} \) has the same expected value of effort \( \pi \geq c \), that is, for all \( p \in \mathcal{P} \), \( \mathbb{E}_p[x] - \mathbb{E}_q[x] = \pi \).

To isolate the paper’s novel contribution, I assume that all agents generate the same increase in the project’s expected cash flow through effort, referred to as the expected value of effort \( \pi \). Section 5.4 relaxes this assumption. Given agents’ identical cost of effort \( c \), agents generate the same first-best expected surplus of effort \( \pi - c \geq 0 \).

At \( t = 0 \), the principal can hire an agent \( p \in \mathcal{P} \) by offering the agent a contract \( s = (s_i)_{i \in \Omega} \in \mathbb{R}^{n+1}. \)\(^{10}\) Since the agent’s effort is not verifiable, the contract can only depend on cash flows, that is, on the state \( i \in \Omega \). The set of feasible contracts is restricted by three canonical contractual frictions.\(^{11}\)

Assumption 4. For all \( i \in \Omega \), the contract \( s \) satisfies \( s_i \geq 0 \).

Limited liability arises if agents have limited wealth and/or if their wealth cannot be monitored.

---

\(^9\)For example, managers differ in their investment, financial, and organizational strategies. This heterogeneity is related to differences in their education, professional and personal experience, and personal characteristics (see, e.g., Bertrand and Schoar, 2003; Kaplan et al., 2012; Huang and Kisgen, 2013; Custódio and Metzger, 2014; Benmelech and Frydman, 2015). In general, different corporate strategies should lead to different output distributions.

\(^{10}\)By a slight abuse of notation, \( s \) denotes both the vector and the random variable.

\(^{11}\)The constraints are important frictions in theories of agency (see, e.g., Sappington, 1983; Singh, 1985; Matthews, 2001; Jewitt et al., 2008; Bond and Gomes, 2009; Poblete and Spulber, 2012), financial intermediation (see, e.g., Holmstrom and Tirole, 1997), security design (see, e.g., Harris and Raviv, 1989; Innes, 1990; Nachman and Noe, 1994; DeMarzo and Duffie, 1999; Dewatripont et al., 2003; Biases and Mariotti, 2005; DeMarzo, 2005; Axelson, 2007), and executive compensation (see, e.g., Kadan and Swinkels, 2008).
and seized.\footnote{Note that the agents’ limited liability constraint and the monotonicity constraints imply the principal’s limited liability constraint: $\forall i \in \Omega : s_i \leq x_i$.} Without agents’ limited liability, the principal could sell the project to an arbitrary agent and capture the full expected surplus $\pi - c \geq 0$, in which case the principal would be indifferent between agents by construction.

**Assumption 5.** For all $i \in \{1, \ldots, n\}$, the contract $s$ satisfies $s_i \geq s_{i-1}$.

**Assumption 6.** For all $i \in \{1, \ldots, n\}$, the contract $s$ satisfies $s_i - s_{i-1} \leq x_i - x_{i-1}$.

Assumptions 5 and 6 preclude contracts that have regions in which payoffs for the agent are decreasing or increasing more than one-to-one with cash flows. The two monotonicity constraints arise from two fundamental frictions (see, e.g., Innes, 1990; Hermalin and Wallace, 2001; Dewatripont et al., 2003). First, the agent can “secretly destroy” cash flows ex post, which he would do in decreasing regions of the contract. Second, the agent can “secretly borrow” at zero cost and inflate cash flows, which he would do in regions where the contract is increasing more than one-to-one with cash flows. In reality, this corresponds to the manipulation of performance measures, which is an important concern in the design of employee incentive pay (see, e.g., Frydman and Jenter, 2010; Murphy, 2013).

### 2.2 Example

Table 1 shows an example with three states and two agents, $p$ and $\tilde{p}$, from the set of agents $\mathcal{P}$, which I refer back to throughout the paper. Agent $p$ shifts probability mass from state $i = 0$ to state $i = 1$, but does not affect the probability of state $i = 2$. The likelihood ratio therefore peaks in state $i = 1$. Agent $\tilde{p}$ shifts probability mass from state $i = 0$ to state $i = 2$, but does not affect the probability of state $i = 1$. The likelihood ratio is therefore monotone and peaks in state $i = 2$. Both agents have the same expected value of effort, given by $\pi = \mathbb{E}_p[x] - \mathbb{E}_q[x] = \mathbb{E}_p[x] - \mathbb{E}_q[x] = 0.2$. The question is whether, given the contracting problem, the principal hires agent $p$ or agent $\tilde{p}$.

Following the bank manager example from the introduction, we can interpret agent $p$ as the candidate who can implement the risk management strategy, and agent $\tilde{p}$ as the candidate who can implement the innovation strategy.
Table 1: Example with two agents. The table summarizes an example with \( n + 1 = 3 \) states, cash flows \( x_i = i, i \in \{0, 1, 2\} \), the principal’s technology \( q \), agents’ technologies \( p, \tilde{p} \in \mathcal{P} \), and the resulting likelihood ratios, \( l(p) \) and \( l(\tilde{p}) \), respectively.

<table>
<thead>
<tr>
<th>( i )</th>
<th>( x_i )</th>
<th>( q )</th>
<th>( p )</th>
<th>( \tilde{p} )</th>
<th>( l_i(p) )</th>
<th>( l_i(\tilde{p}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0.4</td>
<td>0.2</td>
<td>0.3</td>
<td>(-0.5)</td>
<td>(-0.25)</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0.4</td>
<td>0.6</td>
<td>0.4</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.3</td>
<td>0</td>
<td>0.5</td>
</tr>
</tbody>
</table>

3 Optimal Contracts

As a first step, I characterize the principal’s optimal contract for an arbitrary agent. As a benchmark case, I first focus on limited liability, ignoring the monotonicity constraints in Section 3.1. I study the full contracting problem in Section 3.2. I assume that the principal chooses to induce effort but endogenize this decision in Section 5.5. While similar contracting problems have been studied before, the contribution of my paper is to study the agency rents resulting from optimal contracts in Sections 4, and the principal’s hiring decision in Section 5.

3.1 Limited Liability Benchmark

In this section, I consider an environment where the only contractual friction is agents’ limited liability (Assumption 4). Consider an arbitrary agent \( p \in \mathcal{P} \). An optimal incentive compatible contract, denoted by \( s^*(p) \), satisfies

\[
s^*(p) \in \arg \max_s \mathbb{E}_p [x - s]
\]  \hspace{1cm} (1a)

subject to

\[
\mathbb{E}_p[s] - c \geq \mathbb{E}_q[s], \hspace{1cm} (1b)
\]

\[
\mathbb{E}_p[s] - c \geq 0, \hspace{1cm} (1c)
\]

\[\forall i \in \Omega : s_i \geq 0. \hspace{1cm} (1d)\]

\(^{13}\text{See, e.g., Poblete and Spulber (2012).}\)
Due to limited liability, the agent earns a rent. To induce effort, the principal has to pay the agent at least 0 in all states and more than 0 in some states. Since all states have a positive probability under the principal’s technology $q$, the agent gets a positive expected utility from shirking, given by $E_q [s^*(p)] > 0$. As a result, the principal has to offer the agent at least the same positive expected utility in equilibrium. In particular, the incentive constraint (1b) implies the participation constraint (1c). As shown, the incentive constraint (1b) binds, and the agent’s rent is given by his expected utility from shirking, that is,

$$E_p [s^*(p)] - c = E_q [s^*(p)] > 0. \ (2)$$

Given Assumption 3 ($E_p [x] = E_q [x] + \pi$), the principal’s expected utility is given by

$$E_p [x - s^*(p)] = E_q [x] + \pi - c - E_q [s^*(p)]. \ (3)$$

The first term, $E_q [x]$, is the expected value of the principal’s technology $q$. The second term, $\pi - c$, is the expected surplus from the agent’s effort, which is identical across agents by construction. The third term, $E_q [s^*(p)]$, is the agent’s rent. In particular, the principal designs the contract to minimize the agent’s rent.

**Lemma 1.** Consider an agent $p \in \mathcal{P}$. Let $j \in \arg \max_{i \in \Omega} l_i (p)$. An optimal contract $s^*(p)$ solving (1) satisfies, for all $i \neq j$, $s^*_i (p) = 0$ and $s^*_j (p) = \frac{c}{p_j - q_j}$.

Under optimal contracts, the principal pays the agent only in the state with the highest likelihood ratio, since it has the highest incentive effect per unit of rent. Intuitively, the likelihood ratio can be interpreted as the informativeness of cash flows in each state. Higher informativeness makes it easier to detect the agent’s effort, which reduces the agent’s rent.

The agent’s rent from an optimal contract in Lemma 1 is given by

$$E_q [s^*(p)] = \sum_{i=0}^{n} q_i s^*_i (p) = \frac{c}{p_j - q_j} = \frac{c}{l^*(p)}.$$ 

A higher cost of effort $c$ requires higher payments to the agent, which increases the expected value from shirking and hence the agent’s rent. A higher maximum likelihood ratio $l^*(p)$ implies a
higher informativeness of the agent’s technology, which reduces the rent. This benchmark allows me to distinguish between the roles of the limited liability and the monotonicity constraints in the principal’s hiring decision.

Consider the leading example from Section 2.2. If the principal hires agent $p$, she pays the agent only in state $i = 1$. If she hires agent $\bar{p}$, she pays the agent only in state $i = 2$. For example, for $c = 0.16$, we get the optimal contracts $s^*(p) = (0, 0.8, 0)$ and $s^*(\bar{p}) = (0, 0, 1.6)$. Both agents have the same maximum likelihood ratio, such that they would earn identical rents if hired by the principal. The principal is therefore indifferent between the two agents in the benchmark. Clearly, the two contracts violate the monotonicity constraints. Contract $s^*(p)$ has a decreasing region, and contract $s^*(\bar{p})$ has a region in which it increases more than one-to-one with cash flows.

3.2 Full Contracting Problem

In this section, I study the full contracting problem with the limited liability and the monotonicity constraints. Consider again an arbitrary agent $p \in \mathcal{P}$. An optimal incentive compatible contract, denoted by $s^*(p)$, satisfies

$$s^*(p) \in \arg \max_s \mathbb{E}_p [x - s]$$

subject to

$$\mathbb{E}_p[s] - c \geq \mathbb{E}_q[s],$$

$$\mathbb{E}_p[s] - c \geq 0,$$

$$\forall i \in \Omega : s_i \geq 0,$$

$$\forall i \in \{1, \ldots, n\} : s_i \geq s_i - 1,$$

$$\forall i \in \{1, \ldots, n\} : s_i - s_i - 1 \leq x_i - x_i - 1.$$

As in the limited liability benchmark in Section 3.1, the incentive constraint binds such that the agent’s rent is given by (2), and the principal’s expected payoff is given by (3).
**Definition 3.** For each agent \( p \in \mathcal{P} \), the cumulative likelihood ratio \( L(p) = (L_i(p))_{i \in \Omega} \in \mathbb{R}^{n+1} \) is defined as follows:

\[
L_i(p) := \frac{\sum_{j=i}^{n}(p_j - q_j)}{\sum_{j=i}^{n}q_j}, \ i \in \Omega.
\]

**Assumption 7.** For all \( i, j \in \Omega \) with \( i \neq j \) and \( L_i(p) > 0 \) and \( L_j(p) > 0 \), we have \( L_i(p) \neq L_j(p) \).

**Remark 1.** Assumption 7 guarantees the uniqueness of the optimal contract. Uniqueness is not relevant in my setting, since I focus on the agent’s rent, and all optimal contracts imply the same rent. I can therefore discard Assumption 7 and pick an optimal contract if the optimal contract is not unique.

The contract that minimizes the agent’s rent in the presence of limited liability and monotonicity constraints is a capped bonus contract (or junior debt).

**Proposition 1.** Consider an agent \( p \in \mathcal{P} \). Let \( L(p) \) satisfy Assumption 7. There exist two thresholds \( \bar{x}_1(p), \bar{x}_2(p) \in [0,x_n] \), such that the unique optimal contract \( s^*(p) \) is given by

\[
s^*_i(p) = \min \left\{ \max \{0,x_i - \bar{x}_1(p)\}, \bar{x}_2(p) \right\}, \ i \in \Omega.
\]

To understand the role of the monotonicity constraints for optimal contracts, consider first two classes of agents’ technologies.

**Corollary 1.** Consider an agent \( p \in \mathcal{P} \). Let \( L(p) \) satisfy Assumption 7. If \( \max_{i \in \Omega} L_i(p) = L_1(p) \), there exists a threshold \( \bar{x}(p) \in [0,x_n] \), such that the unique optimal contract \( s^*(p) \) is given by

\[
s^*_i(p) = \min \{ x_i, \bar{x}(p) \}, \ i \in \Omega.
\]

If \( \max_{i \in \Omega} L_i(p) = L_n(p) \), there exists a threshold \( \bar{x}(p) \in [0,x_n] \), such that the unique optimal contract \( s^*(p) \) is given by

\[
s^*_i(p) = \max \{0,x_i - \bar{x}(p)\}, \ i \in \Omega.
\]

Following the conventions of the literature, I call the first type of contract in Corollary 1 a debt contract and the second type of contract a (levered) equity contract. The (levered) equity contract
with an increasing cumulative likelihood ratio corresponds to the optimal contract in Innes (1990). Similar to Poblete and Spulber (2012), Corollary 1 shows that the optimality of (levered) equity is not robust to more general technologies.

Consider again the leading example from Section 2.2 with $c = 0.16$. In the limited liability benchmark, the optimal contract for agent $p$ is given by $(0, 0.8, 0)$, violating Assumption 5. As a result, the principal also has to pay the agent in state $i = 2$, such that the optimal contract is given by $s^*(p) = (0, 0.8, 0.8)$. In particular, $s^*(p)$ is a debt contract, and I refer to agent $p$ as a debt agent. Further, in the benchmark, the optimal contract for agent $\tilde{p}$ is given by $(0, 0, 1.6)$, which violates Assumption 6. As a result, the principal also has to pay the agent in state $i = 1$, such that the optimal contract is given by $s^*(\tilde{p}) = (0, 0.6, 1.6)$. In particular, $s^*(\tilde{p})$ is a (levered) equity contract, and I refer to agent $\tilde{p}$ as an equity agent.

In the full contracting problem, the cumulative likelihood ratio determines the design of the optimal contract. This is because if the principal decides to pay the agent in state $i \in \Omega$, she also has to pay the agent in all higher states $j \geq i$, since contracts have to be nondecreasing. Paying the agent in a region corresponding to a high cumulative likelihood ratio implies a low rent for the agent. Intuitively, the cumulative likelihood ratio can be interpreted as the “average informativeness” of cash flows in this region. A higher informativeness makes it easier to detect the agent’s effort, which reduces the agent’s rent.

Further, contracts cannot increase more than one-to-one with cash flows, which generally prevents the principal simply paying the agent in the region corresponding to the highest cumulative likelihood ratio. The optimal contract is therefore composed of “tranches” in the sense that the principal pays the agent first in the region corresponding to the highest cumulative likelihood ratio. When the principal reaches the additional constraint that prevents the contract from increasing more than one-to-one with cash flows, the principal pays the agent further in the region corresponding to the second highest cumulative likelihood ratio, followed by the third highest cumulative likelihood ratio, and so on. The cumulative likelihood ratio is derived from the likelihood ratio as a “weighted average” and also single-peaked. The optimal contract therefore always takes the form of a capped bonus contract.

Consider the example in Table 2. Figure 1 plots the cumulative likelihood ratio $L(p)$ and the
optimal contracts $s^*(p)$ for different levels of the cost of effort $c \in \{\bar{c}_1, \bar{c}_2, \pi\}$. The arrows show the increase in contractual payoffs between the thresholds. If the cost of effort is low ($0 \leq c \leq \bar{c}_1$), the principal pays the agent a first contract tranche in the region exceeding the state with the highest cumulative likelihood ratio (denoted by the blue dots). For medium costs of effort ($\bar{c}_1 < c \leq \bar{c}_2$), the principal adds a second contract tranche in the region exceeding the state with the second highest cumulative likelihood ratio (denoted by the green triangles). If the cost of effort is high ($\bar{c}_2 < c \leq \pi$), the principal adds a third contract tranche in the region exceeding the state with the lowest positive cumulative likelihood ratio (denoted by the red squares).

<table>
<thead>
<tr>
<th>$i$</th>
<th>$x_i$</th>
<th>$q_i$</th>
<th>$p_i$</th>
<th>$l_i(p)$</th>
<th>$L_i(p)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0.25</td>
<td>0.05</td>
<td>-0.8</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0.25</td>
<td>0.05</td>
<td>-0.8</td>
<td>0.27</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>0.25</td>
<td>0.6</td>
<td>1.4</td>
<td>0.8</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>0.25</td>
<td>0.3</td>
<td>0.2</td>
<td>0.2</td>
</tr>
</tbody>
</table>

Table 2: Example with one agent. The table summarizes an example with $n+1 = 4$ states, cash flows $x_i = i$, $i \in \{0, \ldots, 3\}$, the principal’s technology $q$, the agent’s technology $p$, and the resulting likelihood and cumulative likelihood ratios, $l_i(p)$ and $L_i(p)$, respectively.

4 Agency Rents

In this section, I consider an arbitrary agent $p \in \mathcal{P}$ and determine the agent’s rent. The principal’s expected utility, given by (3), depends on the agent’s type $p \in \mathcal{P}$ only through the agency rent, $\mathbb{E}_q[s^*(p)]$, which therefore determines the principal’s hiring decision.

4.1 Agency Rent Function

I call the mapping $[0, \pi] \ni c \mapsto \mathbb{E}_q[s^*(p)] \in \mathbb{R}_+$ the agency rent function of agent $p$, where $s^*(p)$ is the optimal contract from Proposition 1.

Proposition 2. Consider an agent $p \in \mathcal{P}$. Let $L(p)$ satisfy Assumption 7. Denote the ranking of
Consider the setting summarized in Table 2. Figure 1a plots the cumulative likelihood ratio $L(p)$. Figure 1b plots the optimal contracts for different values of the cost of effort $c \in \{ \tilde{c}_1, \tilde{c}_2, \pi \}$, where $\tilde{c}_1 = 0.4$, $\tilde{c}_2 = 0.6$, and $\pi = 0.65$, stated to the right of the respective contracts.

states according to $L(p)$ by $i_1, \ldots, i_n$, where $i_j \in \{1, \ldots, n\}$, such that

$$L_{i_1}(p) > \cdots > L_{i_k}(p) > L_{i_{k+1}}(p) = \cdots = L_{i_n}(p) = 0,$$

where $k = n$ means that, for all $i \in \{1, \ldots, n\}$, $L_i(p) > 0$.

(i) There exists a partition $(\tilde{c}_j)_{j \in \{0, \ldots, k\}}$ of the interval $[0, \pi]$, with $0 = \tilde{c}_0 < \cdots < \tilde{c}_k = \pi$, defined recursively by $\tilde{c}_0 = 0$ and, for all $j \in \{1, \ldots, k\}$, $\tilde{c}_j = \tilde{c}_{j-1} + (x_{i_j} - x_{i_{j-1}}) \sum_{i=i_j}^{n} (p_i - q_i)$, such that, for all $c \in (\tilde{c}_{j-1}, \tilde{c}_j)$, we have

$$\frac{\partial \mathbb{E}_q [s^*(p)]}{\partial c} = \frac{\sum_{i=i_j}^{n} q_i}{\sum_{i=i_j}^{n} (p_i - q_i)} = \frac{1}{L_{i_j}(p)} > 0.$$

(ii) The agency rent function is continuous, increasing, piecewise linear, (weakly) convex, and equal to 0 at $c = 0$.

Remark 2. Note that Proposition 2 also generalizes for technologies that do not satisfy Assumption 7 as discussed in Remark 1.

Proposition 2 shows that the agency rent function is a piecewise linear function. The slopes are
given by the inverse of the cumulative likelihood ratios of the states, ordered from the highest cumulative likelihood ratio (the lowest slope) to the smallest positive cumulative likelihood ratio (the highest slope). The intuition is that, if the principal pays the agent a contract tranche corresponding to state $i \in \Omega$, the marginal agency rent is given by the inverse of the cumulative likelihood ratio in the state, $\frac{1}{L_i(p)}$. Agents with different technologies $p \in \mathcal{P}$ have different cumulative likelihood ratios $L(p)$, which determine the shape of the agency rent function, implying that different agents generally earn different rents.

Figure 2 plots the optimal contracts for different values of the cost of effort and the resulting agency rent function, for the example in Table 2. If the cost of effort is low ($0 < c < \bar{c}_1$), the principal pays the agent a first contract tranche in the region exceeding the state with the highest cumulative likelihood ratio, and the marginal agency rent is equal to the inverse of the cumulative likelihood ratio $\frac{1}{L_2(p)}$ (denoted by the blue dots). For intermediate costs of effort ($\bar{c}_1 < c < \bar{c}_2$), the principal adds a second contract tranche in the region exceeding the state with the second highest cumulative likelihood ratio, and the marginal agency rent is given by $\frac{1}{L_1(p)}$ (denoted by the green triangles). If the cost of effort is high ($\bar{c}_2 < c < \pi$), the principal adds a third contract tranche in the region exceeding the state with the lowest positive cumulative likelihood ratio, and the marginal agency rent is given by $\frac{1}{L_3(p)}$ (denoted by the red squares).

The marginal agency rent depends both on the agent’s technology $p$ and the principal’s technology $q$. Agency rents therefore depend on the match between the principal’s and the agent’s technology. The marginal agency rent is low if the probability under the principal’s technology $\sum_{i=j}^{n} q_i$ is low and the improvement from the agent’s effort $\sum_{i=j}^{n} (p_i - q_i)$ is high. In this sense, the technologies exhibit a complementarity. For example, if a firm has a low probability of realizing high cash flows and hires a manager who increases the probability of high cash flows, then the cost of incentivizing the agent in high states is low. See Appendix B for further discussion of this.

4.2 Productivity of Effort and Equivalent Models

In this section, I show that the agency rent function also allows me to derive the comparative statics with respect to the expected value of effort $\pi$. I show that only the ratio of the expected value of
Figure 2: Optimal contracts and agency rent function. Consider the setting summarized in Table 2. Figure 2a plots the optimal contracts for different values of the cost of effort $c \in \{\bar{c}_1, \bar{c}_2, \pi\}$, where $\bar{c}_1 = 0.4$, $\bar{c}_2 = 0.6$, and $\pi = 0.65$, stated to the right of the respective contracts. Figure 2b plots the agency rent function and the slopes of the linear regions.

4.2.1 Parameterization of Technologies

In this section, I first construct technologies to be a direct function of the expected value of effort. Consider the set of agents $\mathcal{P}$ satisfying Assumptions 1, 2, and 3 with $\pi > 0$. The expected value of effort $\pi$ imposes a constraint on the mean of an agent’s technology $p \in \mathcal{P}$, that is, I require

$$\mathbb{E}_p[x] - \mathbb{E}_q[x] = \pi.$$ 

In particular, a technology $p \in \mathcal{P}$ is not an explicit function of the parameter $\pi$. I therefore construct agents’ technologies to be a direct function of $\pi$. For each technology $p$, define the following basic technology:

$$\hat{p} := q + \frac{p - q}{\pi},$$

---

14 A set of technologies $\mathcal{P}$ satisfying Assumptions 1, 2, and 3 amounts to a set $\mathcal{P} \subset [0, 1]^{n+1}$ such that all $p \in \mathcal{P}$ satisfy $\forall j \in \Omega: \sum_{i=0}^{j} (p_i - q_i) \leq 0, \sum_{i=0}^{n} (p_i - q_i) = 0$, $l(p)$ is single-peaked, $\mathbb{E}_p[x] - \mathbb{E}_q[x] = \pi$, where $\pi$ is low enough such that the “probability constraints”, $\forall i \in \Omega: 0 \leq p_i \leq 1$, are not binding (which I assume for all $\pi$ considered below).
which satisfies $\mathbb{E}_{\hat{p}}[x] - \mathbb{E}_q[x] = 1$ by construction. In other words, a basic technology preserves the shape and is scaled to a unit expected value of effort. Define the set of basic technologies as follows:

$$\hat{P} := \left\{ q + \frac{p - q}{\pi} \mid p \in P \right\}.$$

I can write the original set $P$ by rescaling the basic technologies as follows:

$$P = \left\{ q + \pi (\hat{p} - q) \mid \hat{p} \in \hat{P} \right\}.$$

Every technology $p \in P$ can therefore be written as $p = q + \pi (\hat{p} - q)$, where $\hat{p} \in \hat{P}$ is a basic technology. The basic technology determines the shape of the technology, and the parameter $\pi$ determines the expected value of effort.

### 4.2.2 Equivalent Models

Using the parameterization of technologies from Section 4.2.1, this section studies how changes in the cost of effort and the expected value of effort jointly affect optimal contracts and agency rents.

**Proposition 3.** Consider an agent $p \in \mathcal{P}$. Let $L(p)$ satisfy Assumption 7. Consider two sets of parameters $(c_1, \pi_1)$ and $(c_2, \pi_2)$, where $c_1, c_2 > 0$. The following three statements are equivalent.

(i) $\frac{\pi_1}{c_1} = \frac{\pi_2}{c_2}$.

(ii) $s^*(p)\big|_{(c, \pi) = (c_1, \pi_1)} = s^*(p)\big|_{(c, \pi) = (c_2, \pi_2)}$.

(iii) $\mathbb{E}_q[s^*(p)]\big|_{(c, \pi) = (c_1, \pi_1)} = \mathbb{E}_q[s^*(p)]\big|_{(c, \pi) = (c_2, \pi_2)}$.

**Remark 3.** Note that Proposition 3 also generalizes for technologies that do not satisfy Assumption 7, as discussed in Remark 1.

Proposition 3 shows that optimal contracts and agency rents are identical if the ratio of the expected value of effort and cost of effort remains constant, which I refer to as the **productivity of effort**. I can therefore interpret comparative statics with respect to the cost of effort as changes in agents’ productivity of effort.

---

15 A basic technology might not be a probability distribution, since for some $i \in \Omega$, we might have $\hat{p}_i < 0$ or $\hat{p}_i > 1$. 

---

20
Corollary 2. Consider two agents \( p, \tilde{p} \in \mathcal{P} \) and two sets of parameters \((c_1, \pi_1)\) and \((c_2, \pi_2)\), where \( c_1, c_2 > 0 \) and \( \frac{\pi_1}{c_1} = \frac{\pi_2}{c_2} \). We then have the following equivalence.

\[
\mathbb{E}_q [s^* (p)] \big|_{(c, \pi) = (c_1, \pi_1)} > \mathbb{E}_q [s^* (\tilde{p})] \big|_{(c, \pi) = (c_1, \pi_1)} \iff \mathbb{E}_q [s^* (p)] \big|_{(c, \pi) = (c_2, \pi_2)} > \mathbb{E}_q [s^* (\tilde{p})] \big|_{(c, \pi) = (c_2, \pi_2)}.
\]

Corollary 2 shows that all models with different costs and expected values of effort but the same productivity of effort generate the same ranking of agents in terms of agency rents.

5 Optimal Agents

In this section, I study the principal’s hiring decision given the solution to the contracting problem. I discuss the leading example in Section 5.1. In Section 5.2, I study the two classes of agents’ technologies from Corollary 1 with agents who receive debt and (levered) equity as optimal contracts. I extend the analysis to general agents in Section 5.3.

5.1 Example

I illustrate the intuition for the main result of the section using the leading example from Section 2.2. First, consider a low cost of effort \( c = 0.08 \), that is, a high productivity of effort. In the limited liability benchmark without monotonicity constraints from Section 3.1, the optimal contract for agent \( p \) is \((0, 0.4, 0)\). The optimal contract for agent \( \tilde{p} \) is \((0, 0, 0.8)\). Since both agents have the same maximum likelihood ratio, they receive identical rents in the benchmark. Under the full contracting problem, contracts have to be nondecreasing. This forces the principal to pay agent \( p \) in state \( i = 2 \), and the optimal contract is given by \( s^* (p) = (0, 0.4, 0.4) \), which increases the agent’s rent relative to the benchmark. In contrast, no further constraint binds for agent \( \tilde{p} \), and the optimal contract is the same as in the benchmark: \( s^* (\tilde{p}) = (0, 0, 0.8) \). If the productivity of effort is high, the rent of the debt agent \( p \) is higher, and the principal hires the equity agent \( \tilde{p} \).

Next, consider a high cost of effort \( c = 0.16 \), that is, a low productivity of effort. In the limited liability benchmark, the optimal contract for agent \( p \) is \((0, 0.8, 0)\). The optimal contract for agent
\( \tilde{p} \) is \((0,0,1.6)\). As in the first case, the first monotonicity constraint binds for the debt agent \( p \), and the optimal contract is given by \( s^*(p) = (0,0.8,0.8) \). In contrast to the first case, the second monotonicity constraint, which prevents contracts from increasing more than one-to-one with cash flows, forces the principal to pay the equity agent \( \tilde{p} \) in state \( i = 1 \) as well, and the optimal contract is given by \( s^*(\tilde{p}) = (0,0.6,1.6) \). In this case, hiring the equity agent \( \tilde{p} \) implies a higher rent, and the principal hires the debt agent \( p \).

Figure 3 plots the agency rent functions for agents \( p \) (blue solid line) and \( \tilde{p} \) (red dashed line). If the cost of effort is low (high productivity of effort), hiring the debt agent \( p \) implies a higher rent, because the principal is forced to pay the agent in the high state. If the cost of effort is high (low productivity of effort), the rent of the equity agent \( \tilde{p} \) increases, since the principal is forced to pay the agent in the intermediate state. The intuition for the switch in the pecking order is as follows. Paying the equity agent 100\% of the cash flow in the high state forces the principal to pay 100\% of the cash flow in the intermediate state as well. In contrast, paying the debt agent 100\% of the cash flow in the intermediate state means that the principal still receives 50\% of the cash flow in the high state. Hence, if the cost of effort exceeds a threshold, incentivizing the equity agent is more costly for the principal.

The main point is that the contracting problem determines not only contract design but also the type of agent the principal hires, even if agents have the same productivity. The contract and the type of agent are jointly determined. The general intuition is that the cost of the frictions depends on the optimal contract. When the principal faces the debt agent \( p \), she is only affected by the first monotonicity constraint. In contrast, if she faces the equity agent \( \tilde{p} \), she is only affected by the second monotonicity constraint.

### 5.2 Debt and Equity Agents

In this section, I explicitly solve the principal’s hiring decision in two classes of agents’ technologies. First, consider the set of debt agents, who if hired by the principal receive debt contracts akin to agent \( p \) from the leading example discussed in Section 5.1. Specifically, define the following
Figure 3: Agency rents for different costs of effort. The figure plots the agency rent functions for agents $p$ and $\tilde{p}$ from the example in Section 2.2.

Corollary 1 shows that each agent $p \in P_D$ receives a debt contract, that is, for all $p \in P_D$, there exists a threshold $\bar{x}(p) \in [0, x_n]$, such that

$$s_i^*(p) = \min \{x_i, \bar{x}(p)\}, \ i \in \Omega.$$

**Proposition 4.** There exists a unique agent $p^D \in P^D$, such that, for all $c \in (0, \pi]$, the rent of agent $p^D$ is lower than the rent of all other agents in $P^D$.

The optimal debt agent $p^D$ minimizes agency rents within the set of debt agents. In particular, the agent determines the lower bound for agency rents in the set.

Next, consider the set of equity agents, who if hired by the principal receive (levered) equity contracts akin to agent $\tilde{p}$ from the leading example discussed in Section 5.1. Specifically, define the following subset of agents:

$$P^E := \{p \in P|L_1(p) \leq \cdots \leq L_n(p)\}.$$

Corollary 1 shows that each agent $p \in P^E$ receives a (levered) equity contract, that is, for all
\( p \in \mathcal{P}^{E} \), there exists a threshold \( \bar{x}(p) \in [0,x_n] \), such that

\[
\tilde{s}_i^+(p) = \max \{0, x_i - \bar{x}(p)\} , \ i \in \Omega.
\]

**Proposition 5.** There exists a unique agent \( p^E \in \mathcal{P}^{E} \), such that, for all \( c \in (0,\pi) \), the rent of agent \( p^E \) is lower than the rent of all other agents in \( \mathcal{P}^{E} \).

The *optimal equity agent* \( p^E \) minimizes agency rents within the set of equity agents. In particular, the agent determines the lower bound for agency rents in the set.

I next determine whether the principal hires the optimal debt agent \( p^D \in \mathcal{P}^{D} \) or the optimal equity agent \( p^E \in \mathcal{P}^{E} \).

**Proposition 6.** Assume that the principal hires an agent from the set of debt agents \( \mathcal{P}^{D} \) or the set of equity agents \( \mathcal{P}^{E} \). Consider the optimal debt agent \( p^D \in \mathcal{P}^{D} \) and the optimal equity agent \( p^E \in \mathcal{P}^{E} \) from Propositions 4 and 5, respectively. There are two cases.

(i) If \( 1 - q_0 > \frac{q_n - x_n - E_q[x]}{x_1} \), then there exists a threshold \( \bar{c} \in (0,\pi) \) such that, for all \( c \in (0,\bar{c}) \), the principal hires the optimal equity agent \( p^E \), and, for all \( c \in (\bar{c},\pi] \), the principal hires the optimal debt agent \( p^D \).

(ii) If \( 1 - q_0 < \frac{q_n - x_n - E_q[x]}{x_1} \), then, for all \( c \in (0,\pi] \), the principal hires the optimal debt agent \( p^D \).

If the optimal agents \( p^D \) and \( p^E \) have the same maximum likelihood ratio and would therefore earn identical rents, if hired by the principal in the limited liability benchmark as discussed in Section 3.1, we obtain the first case of Proposition 6.\(^{16}\) In this case, the principal hires the optimal equity agent if the cost of effort is below the threshold \( \bar{c} \) (that is, if the productivity of effort is high), and hires the optimal debt agent if the cost of effort is above the threshold \( \bar{c} \) (that is, if the productivity of effort is low).

The intuition for the first result is identical to the intuition from the leading example discussed in Section 5.1. If the productivity of effort is high, the rent of the optimal debt agent implied by

\(^{16}\)See the proof of Proposition 6 for details.
the contracting problem is higher. This is because even if the principal pays the agent a small share of cash flows in state 1, she is forced to pay the agent in all higher states as well. In contrast, the principal pays the optimal equity agent a small share of the cash flow in state \( n \), and no further constraint binds. As productivity declines, the principal has to increase the optimal equity agent’s share of the cash flow in state \( n \) and is therefore forced to pay the agent in lower states too. Below a certain threshold, the principal needs to pay the equity agent a high share of the cash flow in state \( n \), which forces the principal to also pay high shares of cash flows in lower states. In contrast, paying the optimal debt agent a high share of the cash flows in state 1 forces the principal to pay the agent the same level but lower shares in higher states. In this case, the rent of the equity agent implied by the contracting problem is higher.

The result reflects the fact that the frictions constrain debt and equity agents’ optimal contracts to different degrees. The constraint that contracts have to be nondecreasing in cash flows binds for debt agents, even if the productivity of effort is high. The constraint that contracts cannot increase more than one-to-one with cash flows binds for equity agents only if productivity of effort is low but becomes very costly in this case.

The second case of Proposition 6 shows that there are cases in which the rent of the optimal debt agent \( p^D \) is always lower. In this case, the principal hires the optimal debt agent regardless of productivity. This is the case when the optimal equity agent has a significantly lower maximum likelihood ratio than the optimal debt agent and would therefore be considerably more costly to incentivize, even in the limited liability benchmark in Section 3.1.

### 5.3 General Agents

In this section, I show that the economic forces that determine the choice between debt and equity agents in Section 5.2 also apply for general agents \( p \in \mathcal{P} \), who share features of both debt and equity agents. To begin, I study the existence of an optimal agent in the general case.

**Lemma 2.** \( \mathcal{P} \) is compact.

Since \( \mathcal{P} \) is compact, and since the mapping \( \mathcal{P} \ni p \mapsto \mathbb{E}_q [s^*(p)] \in \mathbb{R}_+ \) is continuous, the extreme value theorem applies and a solution to the agent selection problem exists.
Corollary 3. There exists an optimal agent \( p^* \in \mathcal{P} \), that is, \( \min_{p \in \mathcal{P}} \mathbb{E}_q [s^*(p)] \) exists, and we have \( \min_{p \in \mathcal{P}} \mathbb{E}_q [s^*(p)] = \mathbb{E}_q [s^*(p^*)] \).

I next characterize general technologies.

Lemma 3. Let \( \pi > 0 \). Consider an agent \( p \in \mathcal{P} \). There exists a state \( m \in \{1, \ldots, n\} \) and a state \( j \in \{m, \ldots, n\} \) such that for all \( i \in \{0, \ldots, m-1\} \), \( p_i \leq q_i \) with a strict inequality in some states, for all \( i \in \{m, \ldots, j\} \), \( p_i > q_i \), and for all \( i \in \{j+1, \ldots, n\} \), \( p_i = q_i \).

Lemma 3 shows that an agent \( p \in \mathcal{P} \) reduces the probability of a region of low cash flow states \( \{0, \ldots, m-1\} \) and increases the probability of a region of high cash flow states \( \{m, \ldots, j\} \). Put differently, by exerting effort, an agent shifts probability mass from low to high states. I can therefore classify technologies as follows.

Definition 4. Let \( \pi > 0 \) and \( m \in \{1, \ldots, n\} \). Denote by \( \mathcal{P}^m \subset \mathcal{P} \) the set of agents satisfying for all \( i \in \{0, \ldots, m-1\} \), \( p_i \leq q_i \), \( p_m > q_m \), and for all \( i \in \{m+1, \ldots, n\} \), \( p_i \geq q_i \).

An agent \( p \in \mathcal{P}^m \) shifts probability mass from lower states \( i < m \) to higher states \( i \geq m \). In particular, we have \( \mathcal{P} = \bigcup_{m=1}^{n} \mathcal{P}^m \).

Lemma 4. Let \( \pi > 0 \) and \( m \in \{1, \ldots, n\} \). Consider an agent \( p \in \mathcal{P}^m \). Let \( j \in \{m, \ldots, n\} \) such that for all \( i \in \{m, \ldots, j\} \), \( p_i > q_i \), and for all \( i \in \{j+1, \ldots, n\} \), \( p_i = q_i \). For all \( c \in [0, \pi] \) and all \( i \in \Omega \), we have

\[
 s_i^*(p) \leq \min\{x_i, x_j\},
\]

which holds with equality for \( c = \pi \).

Lemma 4 determines an upper bound for an agent’s optimal contract. If an agent \( p \in \mathcal{P}^m \) affects the probability of cash flows only up to a state \( j \in \{m, \ldots, n\} \), then the bound is given by a debt contract with face value \( x_j \). I refer to an agent or agent’s technology with a lower bound, that is, a lower \( j \in \{m, \ldots, n\} \), as more debt-like. I refer to an agent or agent’s technology with a higher bound, that is, a higher \( j \in \{m, \ldots, n\} \), as more equity-like. Intuitively, consistent with the notion of debt and equity agents from Section 5.2, if an agent’s technology improves a lower region of the cash flow distribution, it is more debt-like, and if an agent’s technology improves a higher region of the cash flow distribution, it is more equity-like.
Proposition 7. Let $\pi > 0$ and $0 < m_1 < m_2 < n$. Consider two agents $p \in \mathcal{P}^{m_1}$ and $\tilde{p} \in \mathcal{P}^{m_2}$. Let $j \leq n - m_2$ such that the agents positively affect the states in the regions $\{m_1, \ldots, m_1 + j\}$ and $\{m_2, \ldots, m_2 + j\}$, respectively, that is, $p_i > q_i \iff i \in \{m_1, \ldots, m_1 + j\}$ and $\tilde{p}_i > q_i \iff i \in \{m_2, \ldots, m_2 + j\}$. There then exists a $\tilde{c} \in [0, \pi)$ such that, for all $c > \tilde{c}$, $\mathbb{E}_q [s^*(p)] < \mathbb{E}_q [s^*(\tilde{p})]$.

The result in Proposition 7 corresponds to the result in Proposition 6 in Section 5.2. It captures the fact that the frictions constrain different optimal contracts for different agents to different degrees. If the cost of effort is low (the productivity of effort is high), incentivizing a more debt-like agent $p \in \mathcal{P}^{m_1}$ can be more costly, since paying the agent a small share of cash flows in low states forces the principal to pay the agent in all higher states as well, increasing the agent’s rent. In contrast, the principal pays a more equity-like agent $\tilde{p} \in \mathcal{P}^{m_2}$ a small share of cash flows in high states and is therefore less exposed to the contractual frictions. As the cost of effort increases (the productivity declines), the principal has to pay a more equity-like agent a higher and higher share of cash flows in high states and is also forced to pay the agent higher and higher shares of cash flows in lower states. In contrast, paying a more debt-like agent a higher share of cash flows in low states forces the principal to pay the agent the same level, but a lower share in higher states. There exists a cost threshold $\tilde{c}$ such that if the cost of effort exceeds the threshold, the more debt-like agent requires a lower rent.

5.4 Agents with Different Productivities

This section extends the model to agents with different expected values of effort $\pi$ and therefore different productivities. In this case, the principal is concerned about both agency rents and productivity. If the principal hires an agent with a higher expected value of effort, the expected total surplus increases.

Denote by $\mathcal{P}_\pi$ the set of agents with expected value of effort $\pi$. Further define the set of optimal agents for a given cost of effort $c \in [0, \pi]$ as follows:

$$\mathcal{P}^*_\pi := \arg\min_{p \in \mathcal{P}_\pi} \mathbb{E}_q [s^*(p)].$$

The set $\mathcal{P}^*_\pi$ contains all agents with the lowest agency rent, potentially including the optimal debt.
agent \( p^D \) or the optimal equity agent \( p^E \) from Section 5.2.

**Proposition 8.** Consider a level of the expected value of effort \( \pi_1 > c > 0 \). For every agent \( p \in \mathcal{P}_{\pi_1} \) with \( p \notin \mathcal{P}^*_{\pi_1} \), there exists a lower level of the expected value of effort \( \pi_2 < \pi_1 \) and an agent \( \tilde{p} \in \mathcal{P}_{\pi_2} \), such that the principal prefers agent \( \tilde{p} \) to agent \( p \).

Proposition 8 shows that there can be a bias towards hiring less productive agents. An agent with an expected value of effort \( \pi_1 \) implies an expected total surplus of \( \mathbb{E}_q[x] + \pi_1 - c \). If the principal hires a less productive agent with an expected value of effort \( \pi_2 < \pi_1 \), the expected total surplus reduces by \( \pi_1 - \pi_2 > 0 \), implying a welfare loss for the economy. The bias can arise because less productive agents can still imply lower agency rents than more productive agents. The principal can therefore be biased towards hiring less productive agents, due to the agency rents resulting from the contractual frictions. The bias can arise whenever the principal cannot hire an optimal agent.

### 5.5 Decision to Induce Effort

In this section, I endogenize the principal’s decision to induce effort, which was taken as given in previous sections, and I study the implications.

Consider an arbitrary agent \( p \in \mathcal{P} \). If the agent exerts effort, the expected total surplus is given by

\[
\mathbb{E}_p[x] - c = \mathbb{E}_q[x] + \pi - c,
\]

where I use \( \mathbb{E}_p[x] = \mathbb{E}_q[x] + \pi \) (Assumption 3). This is larger or equal to the expected total surplus without effort, which is given by \( \mathbb{E}_q[x] \), since \( \pi \geq c \) by Assumption 3. In particular, effort is first-best efficient by construction.

The principal compares her expected utility from inducing effort to her expected utility from not inducing effort. She determines the optimal contract \( s^*(p) \) (Proposition 1) and the agency rent \( \mathbb{E}_q[s^*(p)] \) (Proposition 2). She induces effort if and only if

\[
\mathbb{E}_p[x - s^*(p)] \geq \mathbb{E}_q[x].
\]
Using $E_p[x] = E_q[x] + \pi$ and $E_p[s^*(p)] - c = E_q[s^*(p)]$ (equation (2)), we can rewrite (5) as

$$E_q[s^*(p)] \leq \pi - c. \quad (6)$$

The **effort condition** (6) is intuitive. The principal induces effort whenever the agency rent she incurs, $E_q[s^*(p)]$, is lower than the expected surplus she obtains from the agent’s effort, $\pi - c$.

Proposition 2 shows that the agency rent function, $[0, \pi] \ni c \mapsto E_q[s^*(p)] \in \mathbb{R}_+$, is equal to zero at $c = 0$ and increasing. The right-hand-side of the effort condition (6) is linearly decreasing in $c$ and equal to zero for $c = \pi$. Hence, we get the following result.

**Corollary 4.** Consider an agent $p \in \mathcal{P}$. There exists a $\bar{c}(p) \in (0, \pi)$, such that the principal, deciding between inducing effort for agent $p$ and not inducing effort for agent $p$, induces effort if and only if the cost of effort falls into the effort region, that is, if and only if $c \leq \bar{c}(p)$.

In particular, the effort region depends on the agent’s technology. To illustrate the result, consider two agency rent functions for two agents, $p$ and $\tilde{p}$, in Figure 4. For agent $p$, the effort region $[0, \bar{c}(p)]$ is larger than the effort region for agent $\tilde{p}$, $[0, \bar{c}(\tilde{p})]$. If we interpret the effort regions as hiring regions, the region in which the principal is willing to hire agent $p$ is larger than that for agent $\tilde{p}$.

Note that the differences in hiring regions arise even though agents have the same productivity. The reason is that agents’ rents, the cost of incentivizing agents, differs between agents with different technologies, even if they have the same productivity. For example, in Figure 4, if $c \in (\bar{c}(\tilde{p}), \bar{c}(p))$, the cost of incentivizing agent $\tilde{p}$ exceeds the expected surplus from the agent’s effort, and the principal would not hire agent $\tilde{p}$. In contrast, the cost of incentivizing agent $p$ is lower, and the principal would be willing to hire agent $p$.

### 6 Implications

In this section, I discuss several empirical and theoretical implications of the model.
Figure 4: Agency rent functions and agents’ first-best expected surplus of effort. The figure illustrates the agency rent functions for two agents $p$ and $\tilde{p}$, and the agents’ first-best expected surplus of effort $\pi - c$. The threshold $\tilde{c}$ refers to the threshold from Proposition 7 and $\bar{c}(p)$ and $\bar{c}(\tilde{p})$ refer to the thresholds from Corollary 4.

6.1 Firms’ Hiring Decisions

Implication 1. Contractual frictions bias hiring decisions.

The model implies that firms’ hiring decisions will be affected by employees’ technologies, even if they have the same productivity, since employees with different technologies require different rents. For example, a firm might hire a candidate with a law background rather than one with a business administration background, even if both candidates have the same productivity, because it can be less costly to incentivize the law candidate. This suggests a novel link between incentives and hiring. In the model, a social planner would be indifferent regarding which type of employee the firm hires, since all agents generate the same expected surplus. The rent is a pure transfer from the principal to an agent that does not affect the expected surplus. In companion papers, I show that in a labor market with canonical search frictions, the contracting problem can lead to inefficient search and hiring decisions (Starmans, 2017a) and ultimately to inefficient equilibrium search unemployment (Starmans, 2017b).

The model also shows that in an environment where employees differ in two dimensions, tech-
nology and productivity, a firm can be biased towards hiring less productive employees. The reason is that a firm might strictly prefer a less productive employee because the loss in expected surplus can be offset by a lower cost of incentivizing the agent, that is, a lower rent. This can lead to departures from social efficiency. Edmans and Gabaix (2011) also feature a bias towards agents with lower productivity, which in their setting comes from risk. In my setting, the bias is driven by contractual frictions and differences in agents’ technologies.

**Implication 2. The legal environment shapes hiring.**

The model shows that the specific nature of the contracting problem affects hiring decisions. As shown, hiring decisions in environments in which contracts are subject only to limited liability differ from hiring in environments where contracts are subject to both limited liability and monotonicity constraints. The set of frictions that a principal faces depends on the legal and institutional environment. For example, the legal environment in one jurisdiction might allow a firm to prevent or reduce the extent of output manipulation by employees, which might be impossible in another jurisdiction due to, for example, differences in legal costs or court congestion. While hiring decisions might differ for other reasons, the results show that firms’ hiring decisions should differ in the two jurisdictions purely because of differences in contractual frictions. In particular, it highlights a channel through which the legal and institutional environment can affect labor market allocations.

### 6.2 Level, Shape, and Dispersion of Incentive Pay

**Implication 3.** _Agents with the same productivity but different technologies can receive different optimal contracts and earn different expected compensation if hired._

The model has implications for the level of employees’ incentive pay. In the model, two agents with the same productivity but different technologies can earn different expected compensation if hired by the principal. Simply put, agents with the same productivity may be compensated differently. The intuition is that it can be more costly to incentivize some agents, because agents require different contracts, which are differentially affected by the contractual frictions. Evidence in line with this interpretation can be found in Graham et al. (2012), who document that manager fixed effects explain a large fraction of the variation in executive pay, including the variable component.
While manager fixed effects incorporate many differences between managers, such as differences in productivity, they can also reflect differences in managers’ technologies.

The literature on wage dispersion tries to understand how workers with the same productivity can be paid different wages.\textsuperscript{18} My model implies a novel theory of wage dispersion for employees with the same productivity resulting from contractual frictions and differences in agents’ technologies. In addition, it also implies a theory of contract dispersion that captures the broad prevalence and dispersion of incentive pay in employment relationships (see, e.g., Murphy, 1999; Oyer, 2000; Lemieux et al., 2009; Frydman and Jenter, 2010; Murphy, 2013). In my companion papers, I formally develop a theory of contract and wage dispersion in the context of a frictional labor market (Starmans, 2017a,b).

**Implication 4.** Incentive compensation should be more (levered) equity-like (e.g., an option contract) if the productivity of effort is high, and more debt-like (e.g., a capped bonus contract) if the productivity of effort is low.

In the model, when the productivity of effort is high, the principal tends to hire an agent who improves a higher region of the cash flow distribution, and the equilibrium contract is more (levered) equity-like, such as an option contract. When the productivity of effort is low, the principal tends to hire an agent who improves a lower region of the cash flow distribution, and the equilibrium contract is more debt-like, such as a capped bonus contract (see Propositions 6 and 7).\textsuperscript{19} While other papers show that different forms of compensation can be optimal, my results offer a prediction about when we should expect certain forms of compensation.

This has implications for the shape of employees’ incentive pay, both across and within different classes of employees. First, productivity can be agent specific. For example, managers are high productivity agents, whereas rank and file employees are low productivity agents. The model predicts that we should expect managers to receive more (levered) equity-like incentive pay such as options, and rank and file employees to receive more debt-like incentive pay such as capped

\textsuperscript{18}See Mortensen (2005) and Rogerson et al. (2005) for surveys of the literature.

\textsuperscript{19}Note that this insight is different from models with linear incentive schemes such as Holmstrom and Milgrom (1987). In this type of model, productivity determines the “incentive power” of the linear contract, whereas the type of contract remains the same. See Edmans and Gabaix (2016) for a summary of the literature. In my model, productivity determines the type of contract due to the principal’s hiring decision. In particular, the principal compares agents with contracts that generate the same level of incentives, that is, the same incentive power.
bonuses. Alternatively, productivity can be job specific. For example, firms with very valuable investment opportunities imply a high productivity of the manager job. In contrast, an otherwise similar firm with less valuable investment opportunities implies a low productivity of the manager job. We should therefore expect firms with more valuable investment opportunities to incentivize managers with more (levered) equity-like incentive pay such as options, and firms with less valuable investment opportunities to incentivize managers with more debt-like incentive pay such as capped bonuses.

The first prediction is consistent with the use of convex incentives such as options in executive compensation (Frydman and Jenter, 2010; Murphy, 2013), whereas incentive contracts for rank and file employees often include more debt-like incentive pay such as capped bonuses (see, e.g., Lemieux et al., 2009). The second prediction is consistent with the higher prevalence of convex incentives such as options in new economy firms and firms with more valuable growth opportunities (Guay, 1999; Ittner et al., 2003; Murphy, 2003).

6.3 Human Capital Formation

Implication 5. The contracting problem between firms and employees can affect human capital formation, in particular decisions about investment in human capital and the design of educational institutions.

Agents can generate different probability distributions of output, which reflect differences in employees’ abilities, talents, education, experience, and other characteristics, for example a law degree compared to a business administration degree. Since the heterogeneity affects firms’ hiring decisions in the presence of contractual frictions, even if there are no differences in productivity, the results have implications for human capital formation (Becker, 1994). For example, having a law or business administration degree affects the “propensity” of being hired even without differences in productivity, which in turn affects employees’ investment decisions in human capital. As such, it also affects the design of degree programs such as MBA programs, since it shapes the technologies of the graduates.
6.4 Choice of Corporate Strategy

**Implication 6.** If a firm can implement different corporate strategies with the same expected value, the contracting problem between the firm and different agents such as its managers affects the choice of corporate strategy.

I discuss two interpretations of a technology \( p \in \mathcal{P} \). First, we can interpret a technology \( p \) as manager specific. The interpretation is that manager \( p \) can implement a particular strategy that generates cash flow distribution \( p \), which can reflect the manager’s specific experience, education, and other characteristics, for example a risk management strategy compared to an innovation strategy, as discussed in the bank manager example in the introduction. Alternatively, we can interpret a technology as a strategy chosen by the firm, to be implemented by an arbitrary manager. Under this interpretation, the choice of \( p \) corresponds to a task design problem of the firm. For example, the bank decides whether the manager should pursue a risk management strategy or an innovation strategy. Since the contracting problem determines the choice of the technology, it determines the bank’s strategy. The general insight is that, even if there are different strategies that can be implemented by a firm and if all generate the same expected value, contractual frictions make some strategies more costly to implement than others.

6.5 Delegation

**Implication 7.** Some tasks are more costly to delegate than others, even if they generate the same expected value. In extreme cases, some tasks are too costly to delegate, while other tasks with the same expected value can be delegated.

More generally, the model has implications for which tasks can be delegated. A general implication of the model is that some tasks can be more expensive to delegate than others, even if they have the same expected value. Consider a task \( p \) that generates expected surplus \( \pi - c \) but implies a high agency rent \( \mathbb{E}_q [s^*(p)] \). In cases when \( \mathbb{E}_q [s^*(p)] > \pi - c > 0 \), the task is too costly to delegate, since the principal would make an expected loss \( \pi - c - \mathbb{E}_q [s^*(p)] < 0 \). However, if the agent owned the production technology, he would capture the full expected surplus and implement
the task, since \( \pi - c > 0 \). In contrast, there can exist a task \( \tilde{p} \) with the same expected surplus \( \pi - c \) but a lower agency rent \( E_{q} [s^{*} (\tilde{p})] < \pi - c \) (see Section 5.5). Although both tasks create the same expected surplus and would be implemented under the agent’s ownership, task \( p \) is too costly to delegate, whereas task \( \tilde{p} \) can be delegated.

### 6.6 Firms’ Stochastic Production Technologies

**Implication 8.** The contracting problem between the firm and different agents such as its managers affects the firm’s equilibrium stochastic production technology.

The model has implications for the equilibrium stochastic production technology. In the model, the choice of an agent \( p \in \mathcal{P} \) determines the probability distribution of output. For example, the choice of a manager \( p \) determines the probability distribution of the firm’s cash flows. The choice of the stochastic production technology is driven by the contracting problem between the principal and the agent. Cochrane (1993) argues: “It seems natural to start with the presumption that the firm has at least some control over the distribution of outputs conditional on inputs, and ask for compelling evidence that it has absolutely none.” He studies the choice of the stochastic production technology in an asset-pricing context without agency problems. My model offers a theory of firms’ equilibrium stochastic production technologies based on the agency problem between the firm and its employees.

### 7 Conclusion

This paper studies a canonical risk-neutral principal-agent model with contractual frictions in which the principal can choose between agents with different effort technologies. In the baseline model, all agents have the same productivity but differ in their technologies, that is, they generate different probability distributions of output under effort, subject to moral hazard.

The main insight is that the contracting problem determines not only optimal contract design but also the type of agent the principal hires, even if agents have the same productivity. The reason is that the contracting problem leads to an endogenous dispersion in agency rents. Due
to the different technologies, agents receive different optimal contracts, which are constrained to
different degrees by the frictions. This leads to a pecking order among agents, which determines
the principal’s hiring decision. When agents also differ in their productivity, the dispersion in
endogenous agency rents can bias the principal towards hiring less productive agents.

The analysis highlights the close link between agents’ technologies, optimal contracts, agency
rents, and the principal’s hiring decision, with implications for firms’ hiring decisions, the level,
shape, and dispersion of incentive pay, human capital formation, the choice of corporate strategy,
delegation, and firms’ production technologies. From a conceptual perspective, the paper suggests
a novel way of modeling agent heterogeneity that can capture differences in employee characteris-
tics such as abilities, talents, education, and experience.
References


37


Murphy, Kevin J., 2013, Executive compensation: Where we are, and how we got there, in George M. Constantinides, Milton Harris, and Rene M. Stulz, eds., *Handbook of the Economics of Finance*, volume 2A, 211–356 (Elsevier, Amsterdam).


A Proofs

A.1 Proof of Lemma 1

First, the agent’s incentive constraint implies the agent’s participation constraint. Incentive compatibility requires $\mathbb{E}_p [s^*(p)] - c \geq \mathbb{E}_q [s^*(p)]$. Due to the agent’s limited liability (Assumption 4), we have $\mathbb{E}_q [s^*(p)] \geq 0$, which implies $\mathbb{E}_p [s^*(p)] - c \geq 0$.

Payments to the agent have to be larger or equal to 0 in each state. Assume that there exists an $h \notin \text{arg}\max_{i \in \Omega} l_i(p)$ such that an optimal contract satisfies $s^*_h(p) > 0$. Consider a variation of the contract such that $s^*_h(p)$ is reduced by $v_h > 0$, where $v_h < s^*_h(p)$, and $s^*_j(p)$ is increased by $v_j > 0$. Denote by $s'(p)$ the resulting contract. I choose $v_h$ and $v_j$ such that the incentive constraint binds. We have

$$
\mathbb{E}_p [s'(p)] - \mathbb{E}_q [s'(p)] = \mathbb{E}_p [s^*(p)] - \mathbb{E}_q [s^*(p)] + v_j (p_j - q_j) - v_h (p_h - q_h)
$$

where I use the fact that the optimal contract satisfies the incentive constraint with equality, that is, $\mathbb{E}_p [s^*(p)] - \mathbb{E}_q [s^*(p)] = c$. If the incentive constraint was not binding, the principal could simply reduce payments to the agent by still satisfying all constraints. Hence, the contract $s'(p)$ satisfies the incentive constraint with equality if

$$
v_j (p_j - q_j) - v_h (p_h - q_h) = 0 \iff v_j = \frac{p_h - q_h}{p_j - q_j} v_h.
$$

(7)

Note that $p_j > q_j$, since $j \in \text{arg}\max_{i \in \Omega} l_i(p)$.

The principal’s expected utility under $s'(p)$ is given by

$$
\mathbb{E}_p [x - s'(p)] = \mathbb{E}_q [x] + \pi - \mathbb{E}_p [s'(p)] + \mathbb{E}_q [s'(p)] - \mathbb{E}_q [s'(p)]
$$

$$
= \mathbb{E}_q [x] + \pi - \mathbb{E}_q [s'(p)] = \mathbb{E}_q [x] + \pi - \mathbb{E}_q [s^*(p)] + v_h q_h - v_j q_j
$$

$$
= \mathbb{E}_q [x] + \pi - \mathbb{E}_q [s^*(p)] + v_h q_h - \frac{p_h - q_h}{p_j - q_j} v_h q_j,
$$

41
where I use \( E_p[x] = E_q[x] + \pi \) (Assumption 3) and equation (7). Hence, the contract \( s'(p) \) generates a higher expected utility for the principal if

\[
E_p[x - s'(p)] > E_p[x - s^*(p)]
\]

\[
\iff E_q[x] + \pi - c - E_q[s^*(p)] + \nu_h q_h - \frac{p_h - q_h}{p_j - q_j} \nu_h q_j > E_q[x] + \pi - E_q[s^*(p)]
\]

\[
\iff q_h > \frac{p_h - q_h}{p_j - q_j} q_j \iff \frac{p_j - q_j}{q_h} > \frac{p_h - q_h}{q_h} \iff l_j(p) > l_h(p),
\]

which holds, since \( j \in \arg\max_{i \in \Omega} l_i(p) \) and \( h \notin \arg\max_{i \in \Omega} l_i(p) \). Hence, \( s^*(p) \) cannot be an optimal contract.

It follows that an optimal contract \( s^*(p) \) satisfies: \( s^*_j(p) > 0 \Rightarrow j \in \arg\max_{i \in \Omega} l_i(p) \). The optimal contract \( s^*(p) \) proposed in the Lemma with, for all \( i \neq j, s^*_i(p) = 0 \), and \( s^*_j(p) = \frac{c}{p_j - q_j} \), satisfies the condition and the constraints of the problem. Note that the optimal contract is not unique if \( |\arg\max_{i \in \Omega} l_i(p)| > 1 \).

\[\Box\]

### A.2 Proof of Proposition 1

I begin by proving two technical lemmas, which I use in the proof of the proposition.

**Lemma A.1.** Each agent’s technology \( p \in \mathcal{P} \) has the following properties.

(i) \( p_0 \leq q_0 \) and \( p_n \geq q_n \).

(ii) \( p_0 < q_0 \) if \( \pi > 0 \).

(iii) \( \forall j \in \Omega : \sum_{i=j}^n (p_j - q_j) \geq 0 \).

**Proof.** First-order stochastic dominance of \( p \) over \( q \) is equivalent to \( \forall j \in \Omega : \sum_{i=j}^n (p_i - q_i) \leq 0 \), in particular for \( j = 0 \), we require \( p_0 \leq q_0 \). \( p_0 = q_0 \) implies \( p_i = q_i \) for all \( i \in \Omega \), since the likelihood ratio would not be single-peaked otherwise. In particular, \( p_0 < q_0 \) if \( \pi > 0 \).

Further, given the restriction \( \sum_{i=0}^n (p_i - q_i) = 0 \), we have for every \( j \in \{0, \ldots, n - 1\} \),

\[
\sum_{i=0}^j (p_i - q_i) \leq 0 \iff 0 \leq \sum_{i=j+1}^n (p_i - q_i).
\]
In particular, we have $p_n \geq q_n$. ■

**Lemma A.2.** For all $p \in P$, the cumulative likelihood ratio $L(p)$ is single-peaked.

*Proof.* For $i \in \{0, \ldots, n - 1\}$, we have $L_i(p) \geq L_{i+1}(p)$ if and only if

$$\frac{(p_i - q_i) + (p_{i+1} - q_{i+1}) + \cdots + (p_n - q_n)}{q_i + q_{i+1} + \cdots + q_n} \geq \frac{(p_{i+1} - q_{i+1}) + \cdots + (p_n - q_n)}{q_{i+1} + \cdots + q_n},$$

which is equivalent to

$$\frac{p_i - q_i}{q_i} \geq \frac{(p_{i+1} - q_{i+1}) + \cdots + (p_n - q_n)}{q_{i+1} + \cdots + q_n},$$

and also equivalent to

$$\frac{p_i - q_i}{q_i} \geq \frac{(p_i - q_i) + \cdots + (p_n - q_n)}{q_i + \cdots + q_n}.$$ 

Hence, we have the following equivalence

$$L_i(p) \geq L_{i+1}(p) \iff l_i(p) \geq L_{i+1}(p) \iff l_i(p) \geq L_i(p). \quad (8)$$

Note that $L_n(p) = l_n(p)$ by definition. Hence, $L_i(p)$ is nonincreasing as long as $l_i(p)$ is nonincreasing. Due to Lemma A.1, we have $L_i(p) \geq 0$ for all $i \in \Omega$. If $\mathbb{E}_p[s^*(p)] - c \geq \mathbb{E}_q[s^*(p)]$, then $p_0 < q_0$, and there exists a highest $j < \min \{\arg \max_{i \in \Omega} l_i(p)\}$, such that $l_j(p) < L_{j+1}(p)$. This implies that $L_j(p) < L_{j+1}(p)$ and $l_j(p) < L_j(p)$ by (8). Since $l(p)$ is single-peaked, $L_i(p)$ is nondecreasing for all $i \leq j$. ■

I first discuss three precedent statements. First, the agent’s incentive constraint implies the agent’s participation constraint. Incentive compatibility requires $\mathbb{E}_p[s^*(p)] - c \geq \mathbb{E}_q[s^*(p)]$. Due to agent limited liability (Assumption 4), we have $\mathbb{E}_q[s^*(p)] \geq 0$, which implies $\mathbb{E}_p[s^*(p)] - c \geq 0$.

Second, the agent’s incentive constraint binds. If this was not the case, there would exist a region in which the principal could reduce the contractual payments to the agent without violating any constraints but increasing her expected payoff.

Third, we have $s^*_0(p) = 0$. Assume that this is not the case and $s^*_0(p) > 0$. Consider a variation of the contract that reduces all contractual payments by $s^*_0(p)$. Since $s^*(p)$ is nondecreasing, and
since I have reduced all payments by the same amount $s^*_0(p)$, the variation satisfies the limited liability and monotonicity constraints. Further, the incentive constraint is invariant to scalar additions to the contract. This variation increases the principal’s expected payoff, which is a contradiction to the assumption that $s^*(p)$ is an optimal contract.

I next show that each feasible contract with $s_0 = 0$ can be written as a linear combination of $n$ contract tranches. For $j \in \{1, \ldots, n\}$, define contract tranche $s^j \in \mathbb{R}^{n+1}$ by $s^j_i = \mathbbm{1}_{\{i \geq j\}}(x_j - x_{j-1})$, $i \in \Omega$, where $\mathbbm{1}$ denotes the indicator function. Let $s$ be a feasible contract with $s_0 = 0$. There then exist $\lambda_1, \ldots, \lambda_n \in [0, 1]$ such that

$$s = \lambda_1 s^1 + \cdots + \lambda_n s^n.$$

I show this by direct construction. For $j \in \{1, \ldots, n\}$, set

$$\lambda_j = \frac{s_j - s_{j-1}}{x_j - x_{j-1}}.$$

Due to the first monotonicity constraint (Assumption 5), we have $s_j \geq s_{j-1}$, which implies $\lambda_j \geq 0$. Due to the second monotonicity constraint (Assumption 6), we have $s_j - s_{j-1} \leq x_j - x_{j-1}$, which implies $\lambda_j \leq 1$.

Consider the cumulative likelihood ratio $L(p)$ satisfying Assumption 7. Denote the agent specific ordering of $L(p)$ by $i_1, \ldots, i_n$, that is, $\{i_1, \ldots, i_n\}$ is a permutation of $\{1, \ldots, n\}$, such that

$$L_{i_1}(p) > \cdots > L_{i_k}(p) > L_{i_{k+1}}(p) = \cdots = L_{i_n}(p) = 0,$$

where $k = n$ means that, for all $i \in \{1, \ldots, n\}$, $L_i(p) > 0$. Let $s^*(p)$ be an optimal contract with the linear combination of the contract tranches given by

$$s^*(p) = \lambda_1 s^1 + \cdots + \lambda_n s^n.$$

Lemma 3 implies that, if it exists, there is an upper region of states $\{g, g+1, \ldots, n\} \subset \Omega$ such that $L_i(p) = 0 \iff i \in \{0\} \cup \{g, g+1, \ldots, n\}$. First, note that, for $i \in \{g, g+1, \ldots, n\}$, we
have $\lambda_i = 0$. If this was not the case, the principal could simply set, for all $i \in \{g, g+1, \ldots, n\}$, $\lambda_i = 0$. Since $L_i(p) = 0 \iff i \in \{0\} \cup \{g, g+1, \ldots, n\}$ implies, for all $i \geq g$, $p_i = q_i$, the variation leaves the incentive constraint unchanged. The variation also satisfies all other constraints of the program. The variation increases the expected payoff of the principal, a contradiction to the initial assumption of optimality.

I next show that we can never have two different states $j, h \in \{1, \ldots, n\}$ with $L_j(p), L_h(p) > 0$ such that $\lambda_j \in (0, 1)$ and $\lambda_h \in (0, 1)$. The proof follows a similar strategy to the proof for Lemma 1. Assume that such a case exists. Without loss of generality, let $L_j(p) > L_h(p)$. The other case is symmetric.

Consider a variation of the contract such that all contractual payments $s_h^*(p), \ldots, s_n^*(p)$ are reduced by $v_h > 0$, where $v_h < s_h^*(p) - s_{h-1}^*(p)$, and all contractual payments $s_j^*(p), \ldots, s_n^*(p)$ are increased by $v_j > 0$. Denote the resulting contract by $s^v(p)$. Note that the resulting contract is a feasible contract if $v_j$ is small enough. I choose $v_h$ and $v_j$ ($v_j$ sufficiently small), such that the incentive constraint binds. We have

$$\mathbb{E}_p[s^v(p)] - \mathbb{E}_q[s^v(p)] = \mathbb{E}_p[s^*(p)] - \mathbb{E}_q[s^*(p)] + v_j \sum_{i=j}^n (p_i - q_i) - v_h \sum_{i=h}^n (p_i - q_i).$$

Using the fact that the optimal contract satisfies the incentive constraint with equality, that is, $\mathbb{E}_p[s^*(p)] - \mathbb{E}_q[s^*(p)] = c$, the contract $s^v(p)$ satisfies the incentive constraint with equality if

$$v_j \sum_{i=j}^n (p_i - q_i) - v_h \sum_{i=h}^n (p_i - q_i) = 0 \iff v_j = \frac{\sum_{i=h}^n (p_i - q_i)}{\sum_{i=j}^n (p_i - q_i)} v_h. \quad (9)$$

Note that $\sum_{i=j}^n (p_i - q_i) > 0$, since $L_j(p) > 0$.

The principal’s expected utility under $s^v(p)$ is given by

$$\mathbb{E}_p[x - s^v(p)] = \mathbb{E}_q[x] + \pi - \mathbb{E}_p[s^v(p)] + \mathbb{E}_q[s^v(p)] - \mathbb{E}_q[s^v(p)]$$

$$= \mathbb{E}_q[x] + \pi - c - \mathbb{E}_q[s^v(p)] = \mathbb{E}_q[x] + \pi - c - \mathbb{E}_q[s^*(p)] + v_h \sum_{i=h}^n q_i - v_j \sum_{i=j}^n q_i$$

$$= \mathbb{E}_q[x] + \pi - c - \mathbb{E}_q[s^*(p)] + v_h \sum_{i=h}^n q_i - \frac{\sum_{i=h}^n (p_i - q_i)}{\sum_{i=j}^n (p_i - q_i)} v_h \sum_{i=j}^n q_i;$$

45
where I use $\mathbb{E}_p[x] = \mathbb{E}_q[x] + \pi$ (Assumption 3) and equation (9). Hence, the contract $s^*(p)$ generates a higher expected utility for the principal if

$$\mathbb{E}_p[x - s^*(p)] > \mathbb{E}_p[x - s^*(p)]$$

$$\iff \mathbb{E}_q[x] + \pi - c - \mathbb{E}_q[s^*(p)] + \nu_h \sum_{i=h}^n q_i - \sum_{i=h}^n (p_i - q_i) \nu_h \sum_{i=j}^n q_i > \mathbb{E}_q[x] + \pi - \mathbb{E}_q[s^*(p)]$$

$$\iff \sum_{i=h}^n q_i > \sum_{i=j}^n (p_i - q_i) \sum_{i=j}^n q_i \iff \sum_{i=h}^n (p_i - q_i) \sum_{i=j}^n q_i > \sum_{i=h}^n (p_i - q_i) \sum_{i=j}^n q_i \iff L_j(p) > L_h(p),$$

which holds by assumption. Hence, $s^*(p)$ cannot be an optimal contract.

The proof also reveals the construction of the optimal contract. The principal first varies $\lambda_{i_1}$, the tranche corresponding to the state with the highest cumulative likelihood ratio, between 0 and 1. If this single tranche does not generate sufficient incentives, that is, if

$$\mathbb{E}_p[s^{i_1}] - \mathbb{E}_q[s^{i_1}] = (x_{i_1} - x_{i_1 - 1}) \sum_{i = i_1}^n (p_i - q_i) =: \bar{c}_1 < c,$$

then the principal sets $\lambda_{i_1} = 1$ and next varies $\lambda_{i_2}$ between 0 and 1, and so on. Define recursively $\bar{c}_0 = 0$ and, for all $j \in \{1, \ldots, k\}$, $\bar{c}_j = \bar{c}_{j-1} + (x_{i_j} - x_{i_{j-1}}) \sum_{i = i_{j-1}}^n (p_i - q_i)$. The principal then varies $\lambda_{i_j}$ when $c \in (\bar{c}_{j-1}, \bar{c}_j)$, $j \in \{1, \ldots, k\}$.

Lemma A.2 shows that the cumulative likelihood ratio is single-peaked. This implies that we have either $i_2 = i_1 + 1$ or $i_2 = i_1 - 1$, that is, $L_{i_2}(p)$ is a “direct neighbor” of $L_{i_1}(p)$. In general, $L_{i_j}(p)$ is a “direct neighbor” of $L_{i_l}(p)$ for an $l < j$, that is, we either have $i_j = i_l + 1$ or $i_j = i_l - 1$. This implies that the tranches are always adjacent to each other, and the resulting optimal contract has the proposed form of a capped bonus contract.

### A.3 Proof of Corollary 1

The results follow directly from the proof of Proposition 1. In the first case, we have $L_1(p) \geq \cdots \geq L_n(p)$, with strict inequalities whenever $L_i(p) \neq 0$. In the second case, we have the opposite ordering $L_n(p) > \cdots > L_1(p)$, which arises in the case of a monotone likelihood ratio $l(p)$. Note that in this second case, an increasing cumulative likelihood ratio implies a technology with $p_n >
A.4 Proof of Proposition 2

From Proposition 1 we know that, for \( j \in \{1, \ldots, k\}, c \in (\bar{c}_{j-1}, \bar{c}_j) \), the principal varies \( \lambda_{ij} \), which amounts to varying payments in the region \( i \geq i_j \). Note that

\[
\sum_{j=1}^{k} (x_{ij} - x_{ij-1}) \sum_{i=i_j}^{n} (p_i - q_i) = \sum_{j=1}^{n} (x_j - x_{j-1}) \sum_{i=j}^{n} (p_i - q_i) = \sum_{i=0}^{n} (p_i - q_i)x_i = \pi,
\]

and hence \( \bar{c}_k = \pi \).

I first determine the comparative statics of the optimal contract \( s^*(p) \) with respect to the cost of effort \( c \), which then allows me to determine the comparative statics of the agency rent \( \mathbb{E}_q[s^*(p)] \) with respect to the cost of effort \( c \). Let \( \hat{c} \in (\bar{c}_{j-1}, \bar{c}_j) \) and let \( s^*(p) \) be the optimal contract. Consider a small increase of \( \hat{c} \) by \( \varepsilon > 0 \) such that \( \hat{c} + \varepsilon < \bar{c}_j \). The principal then increases each \( s^*_i(p), i \geq i_j \), by the same \( \sigma > 0 \) such that

\[
\sum_{i=i_j}^{n} \sigma(p_i - q_i) = \varepsilon \iff \sigma = \frac{\varepsilon}{\sum_{i=i_j}^{n} (p_i - q_i)},
\]

which requires \( \sum_{i=i_j}^{n} (p_i - q_i) > 0 \), which holds, since \( L_{ij}(p) > 0 \). Letting \( \varepsilon \searrow 0 \), we get, for all \( i \geq i_j \),

\[
\left. \frac{\partial s^*_i(p)}{\partial c} \right|_{c=\hat{c}} = \lim_{\varepsilon \searrow 0} \frac{\sigma}{\varepsilon} = \frac{1}{\sum_{i=i_j}^{n} (p_i - q_i)},
\]

where \( \left. \frac{\partial s^*_i(p)}{\partial c} \right|_{c=\hat{c}} \) denotes the right derivative. Hence

\[
\frac{\partial \mathbb{E}_{p}[s^*(p)]}{\partial c} \bigg|_{c=\hat{c}} = \frac{\sum_{i=i_j}^{n} p_i}{\sum_{i=i_j}^{n} (p_i - q_i)} = \frac{\sum_{i=i_j}^{n} (q_i + p_i - q_i)}{\sum_{i=i_j}^{n} (p_i - q_i)} = 1 + \frac{\sum_{i=i_j}^{n} q_i}{\sum_{i=i_j}^{n} (p_i - q_i)} = 1 + \frac{1}{L_{ij}(p)}.
\]

In particular, the agent’s expected compensation is linear in \( c \) in each subinterval \( (\bar{c}_{j-1}, \bar{c}_j) \), \( j \in \{1, \ldots, k\} \). We therefore have, for \( j \in \{1, \ldots, k\}, c \in (\bar{c}_{j-1}, \bar{c}_j) \),

\[
\frac{\partial \mathbb{E}_{p}[s^*(p)]}{\partial c} = 1 + \frac{1}{L_{ij}(p)}.
\]
Since the agent’s incentive constraint binds, the agent’s rent is given by \( E_p [s^*(p)] - c = E_q [s^*(p)] \).

Using

\[
\frac{\partial E_q [s^*(p)]}{\partial c} = \frac{\partial E_p [s^*(p)]}{\partial c} - 1,
\]

we have for \( j \in \{1, \ldots, k\}, c \in (\bar{c}_{j-1}, \bar{c}_j) \),

\[
\frac{\partial E_q [s^*(p)]}{\partial c} = \frac{1}{L_{ij}(p)}.
\]

From Proposition 1, it is clear that the agency rent function \([0, \pi] \ni c \mapsto E_q [s^*(p)] \in \mathbb{R}_+\) is continuous and equal to 0 for \( c = 0 \). The first statement of the proof implies that the agency rent function is increasing and piecewise linear. Due to the ordering of \( L(p) \), the slope of the agency rent function is increasing across the adjacent subintervals \((\bar{c}_{j-1}, \bar{c}_j)\) of \([0, \pi]\). Hence, the agency rent function is (weakly) convex.

\[\blacksquare\]

A.5 Proof of Proposition 3

I first show the equivalence of the first and the third statements. To do this, I study the agency rent functions specified in Proposition 2. Define \( D_j(p) := (x_j - x_{j-1}) \sum_{i=j}^n (p_i - q_i) \) and \( Q_j := (x_j - x_{j-1}) \sum_{i=j}^n q_i \). Let \( i_j, j \in \{1, \ldots, n\} \), denote the ordering of the cumulative likelihood ratio specified in Proposition 2. The agency rent functions are linear on intervals \((\bar{c}_{j-1}, \bar{c}_j)\) with slope \( D_{ij}(p) := \sum_{i=1}^j q_i \sum_{i=j}^n (p_i - q_i) \), where \( \bar{c}_0 = 0 \) and \( \bar{c}_j = \sum_{i=1}^j D_{ij}(p) \). Hence, each interval has length

\[
\bar{c}_j - \bar{c}_{j-1} = D_{ij}(p) = (x_{i_j} - x_{i_j-1}) \sum_{i=i_j}^n (p_i - q_i).
\]

On the interval \([\bar{c}_{j-1}, \bar{c}_j]\), the agency rent function increases by

\[
(\bar{c}_j - \bar{c}_{j-1}) \frac{\sum_{i=i_j}^n q_i}{\sum_{i=i_j}^n (p_i - q_i)} = \left( (x_{i_j} - x_{i_j-1}) \sum_{i=i_j}^n (p_i - q_i) \right) \frac{\sum_{i=i_j}^n q_i}{\sum_{i=i_j}^n (p_i - q_i)}
\]

\[
= (x_{i_j} - x_{i_j-1}) \sum_{i=i_j}^n q_i = Q_{ij}.
\]
Hence, the agency rent function has coordinates

\[(0, 0), (D_{i_1} (p), Q_{i_1}), (D_{i_1} (p) + D_{i_2} (p), Q_{i_1} + Q_{i_2}), \ldots, (D_{i_1} (p) + \cdots + D_{i_k} (p), Q_{i_1} + \cdots Q_{i_k}), \ldots].\]

From Section 4.2.1, we know that we can write \(p = q + \pi (\hat{p} - q)\), where \(\hat{p} \in \hat{\mathcal{P}}\) is a basic technology. We have \(c \in [\bar{c}_{j-1}, \bar{c}_j]\) is equivalent to

\[
\sum_{l=1}^{j-1} D_{i_l} (p) \leq c \leq \sum_{l=1}^{j} D_{i_l} (p) \iff \sum_{l=1}^{j-1} D_{i_l} (q + \pi (\hat{p} - q)) \leq c \leq \sum_{l=1}^{j} D_{i_l} (q + \pi (\hat{p} - q))
\]

\[
\iff \pi \sum_{l=1}^{j-1} D_{i_l} (\hat{p}) \leq c \leq \pi \sum_{l=1}^{j} D_{i_l} (\hat{p}) \iff \sum_{l=1}^{j-1} D_{i_l} (\hat{p}) \leq \frac{c}{\pi} \leq \sum_{l=1}^{j} D_{i_l} (\hat{p}).
\]

The agency rent for \(c \in [\bar{c}_{j-1}, \bar{c}_j]\), denoted by \(A(p, \pi, c)\), is given by

\[
A(p, \pi, c) = Q_{i_1} + \cdots + Q_{i_{j-1}} + \frac{Q_{i_j}}{D_{i_j} (p)} \left( c - \sum_{l=1}^{j-1} D_{i_l} (p) \right)
= Q_{i_1} + \cdots + Q_{i_{j-1}} + \frac{Q_{i_j}}{D_{i_j} (q + \pi (\hat{p} - q))} \left( c - \sum_{l=1}^{j-1} D_{i_l} (q + \pi (\hat{p} - q)) \right)
= Q_{i_1} + \cdots + Q_{i_{j-1}} + \frac{Q_{i_j}}{\pi D_{i_j} (\hat{p})} \left( c - \pi \sum_{l=1}^{j-1} D_{i_l} (\hat{p}) \right)
= Q_{i_1} + \cdots + Q_{i_{j-1}} + \frac{Q_{i_j}}{\pi D_{i_j} (\hat{p})} \left( \frac{c}{\pi} - \sum_{l=1}^{j-1} D_{i_l} (\hat{p}) \right) = A \left( \hat{p}, 1, \frac{c}{\pi} \right).
\]

Hence, the agency rent function can be written as a function of \(\frac{c}{\pi}\). In particular we can simply evaluate the agency rent function of the basic technology \(\hat{p} \in \hat{\mathcal{P}}\) at \(\frac{c}{\pi}\). Note that the agency rent function of the basic technology is increasing, and hence the equivalence between the first and the third statements holds.

From the proof of Proposition 1, it is clear that there is a bijective relationship between the optimal contracts and the agency rent for an arbitrary agent, and hence, the second and the third statements are equivalent, which completes the proof.
A.6 Proof of Proposition 4

The proof proceeds by guessing a technology in $\mathcal{P}^D$ and verifying that it is the unique minimum rent technology in $\mathcal{P}^D$. Define the technology $p^D \in \mathcal{P}^D$ as follows: $p^D_0 = q_0 - \frac{\pi}{x_1}$, $p^D_1 = q_1 + \frac{\pi}{x_1}$, and, for $i \in \{2, \ldots, n\}$, $p^D_i = q_i$, and denote the optimal contract by $s^*(p^D)$. By Proposition 2, the slope of the agency rent function for $p^D$ for all costs of effort $c \in (0, \pi)$ is given by

$$\frac{\partial \mathbb{E}_q[s^*(p^D)]}{\partial c} = \frac{\sum_{i=1}^{n} q_i}{\sum_{i=1}^{n} (p^D_i - q_i)} = \frac{(\sum_{i=1}^{n} q_i)x_1}{\pi}.$$ 

In particular, the agency rent function is linear with slope $\frac{(\sum_{i=1}^{n} q_i)x_1}{\pi}$.

Consider now a different agent $p \in \mathcal{P}^D$, and denote the optimal contract by $s^*(p)$. We know that $p_n \geq q_n$ by Lemma A.1 (see Appendix A.2). Since $1 \in \arg \max_{i \in \Omega} L_i(p)$, we have $p_1 > q_1$. Hence, since the likelihood ratio $l(p)$ is single-peaked by Assumption 2, we need $p_i \geq q_i$ for all $i \geq 2$. Since $p \neq p^D$, there exists a $j \geq 2$ such that $p_j > q_j$. This implies that $\sum_{i=1}^{n} (p_i - q_i) < \frac{\pi}{x_1}$, since we would have $\mathbb{E}_p[x] - \mathbb{E}_q[x] > \pi$ otherwise. Hence, the initial slope of the agency rent function for $c \leq x_1 \sum_{i=1}^{n} (p_i - q_i) < \pi$ is given by

$$\frac{\partial \mathbb{E}_q[s^*(p)]}{\partial c} = \frac{\sum_{i=1}^{n} q_i}{\sum_{i=1}^{n} (p_i - q_i)} > \frac{(\sum_{i=1}^{n} q_i)x_1}{\pi} = \frac{\partial \mathbb{E}_q[s^*(p^D)]}{\partial c}.$$ 

Since the agency rent function is (weakly) convex, we have that for all $c \in (0, \pi]$, $\mathbb{E}_q[s^*(p)] > \mathbb{E}_q[s^*(p^D)]$, which completes the proof.

A.7 Proof of Proposition 5

I make use of the notation and characterization of the agency rent function from the proof for Proposition 3 in Appendix A.5. Since for all $p \in \mathcal{P}^E$ the cumulative likelihood ratio $L(p)$ is nondecreasing, we get the following coordinates of the agency rent function, that is, the agency
rent function is piecewise linear and linear between two coordinates.

\[(0, 0), (\bar{c}_1 = D_n(p), Q_n), (\bar{c}_2 = D_n(p) + D_{n-1}(p), Q_n + Q_{n-1}), \ldots,\]
\[(\bar{c}_n = D_n(p) + D_{n-1}(p) + \cdots + D_1(p), Q_n + Q_{n-1} + \cdots + Q_1).\]

In particular, each equity technology \(p \in \mathcal{P}_E\) takes on the same values at the cutoffs \(\bar{c}_j, j \in \Omega\).

The proof proceeds by guessing a technology in \(\mathcal{P}_E\) and verifying that it is the unique minimum rent technology in \(\mathcal{P}_E\). The optimal equity technology \(p^E \in \mathcal{P}_E\) is the technology with a constant likelihood ratio for all \(i < n\), that is, it satisfies, for all \(i, j \in \{0, \ldots, n-1\}\), \(l_i(p^E) = l_j(p^E) < 0\), and \(l_n(p^E) > 0\). To show this, assume that this is not the case, and there exists an agent \(p \in \mathcal{P}_E, p \neq p^E\) and a \(c \in (0, \pi)\), such that the agency rent for technology \(p\) is lower than for \(p^E\). Since \(p \neq p^E\), there exists a smallest \(k \in \{0, \ldots, n-2\}\) such that \(l_k(p) < l_{k+1}(p)\). Consider two variations \(\nu^k \in \mathbb{R}\) and \(\delta^k \in \mathbb{R}\) that leave the ordering of \(L\) unchanged.

1. Variation \(\nu^k\): Reduce \(p_{k+1}\) by a small \(\nu^k > 0\) and increase \(p_k\) by \(\nu^k\). This variation reduces the value of the technology by \((x_{k+1} - x_k) \nu^k\).

2. Variation \(\delta^k\): Reduce \(p_{k+1}\) by a small \(\delta^k > 0\) and increase \(p_n\) by \(\delta^k\). I choose \(\delta^k\) such that both variations together leave the value of the technology unchanged, that is,

\[(x_{k+1} - x_k) \nu^k = (x_n - x_{k+1}) \delta^k \iff \nu^k = \frac{x_n - x_{k+1}}{x_{k+1} - x_k} \delta^k. \tag{10}\]

Denote the technology with the two variations applied by \(p^k\). We can now calculate and compare the cost thresholds \(\bar{c}_j\) for technology \(p\) and \(\bar{c}^k_j\) for technology \(p^k, j \in \Omega\). Consider two cases. First, let \(k = 0\).

1. For \(j = 1\), we have

\[\bar{c}^k_1 = (p^k_n - q_n) (x_n - x_{n-1}) = \left((p_n - q_n) + \delta^k\right) (x_n - x_{n-1}) > (p_n - q_n)(x_n - x_{n-1}) = \bar{c}_1.\]
2. We can proceed by iteration. Assuming \( c_{j-1}^k > \bar{c}_{j-1} \), for all \( j < n \), we then get

\[
\bar{c}_j^k = c_{j-1}^k + (x_{n-j+1} - x_{n-j}) \left( \sum_{i=n-j+1}^{n} (p_i - q_i) + \delta^k \right)
\]

\[
\bar{c}_j = \bar{c}_{j-1}^k + (x_{n-j+1} - x_{n-j}) \sum_{i=n-j+1}^{n} (p_i - q_i) = \bar{c}_j.
\]

3. Finally, using (10), we get

\[
\bar{c}_n^k = \mathbb{E}_p[x] - \mathbb{E}_q[x] + x_n \delta^k - x_1 \delta^k - \nu^k(x_1 - x_0)
\]

\[
= \mathbb{E}_p[x] - \mathbb{E}_q[x] + (x_n - x_1) \delta^k - (x_1 - x_0) \frac{x_n - x_1}{x_1 - x_0} \delta^k = \mathbb{E}_p[x] - \mathbb{E}_q[x] = \bar{c}_n.
\]

Hence, we have \( \bar{c}_j^k > \bar{c}_j \) for all \( j < n \), and therefore the variation has lower agency rents for all costs of effort \( c \in (0, \pi) \), since the agency rent functions take on the same values at the cutoffs as shown above.

Next, consider the case where \( k > 0 \). Following the same argument as above, I can show that for all \( j < n-k \), \( \bar{c}_j^k > \bar{c}_j \), and for all \( j \geq n-k \), \( \bar{c}_j^k = \bar{c}_j \). We can then proceed by iteration. We can apply further variations \( \nu^{k-1} \) and \( \delta^{k-1} \) to \( p_k \) as before (note that \( l_{k-1}(p_k) < l_k(p_k) \) due to the increase in \( p_k \)).

1. Variation \( \nu^{k-1} \): Reduce \( p_k \) by a small \( \nu^{k-1} > 0 \) and increase \( p_{k-1} \) by \( \nu^{k-1} \). This variation reduces the expected value of the technology by \( \nu^{k-1}(x_k - x_{k-1}) \).

2. Variation \( \delta^{k-1} \): Reduce \( p_k \) by a small \( \delta^{k-1} > 0 \) and increase \( p_n \) by \( \delta^{k-1} \). Choose \( \delta^{k-1} \) such that

\[
(x_k - x_{k-1}) \nu^{k-1} = (x_n - x_k) \delta^{k-1} \iff \nu^{k-1} = \frac{x_n - x_k}{x_k - x_{k-1}} \delta^{k-1}.
\]

This variation weakly decreases the agency rent function (strictly for some \( c \)). If \( k = 1 \), we are finished. If \( k > 1 \), we can then apply the same step again, applying variations \( \nu^{k-2} \) and \( \delta^{k-2} \), and continue to get to the last variation of \( p_0 \) and \( p_1 \). This reduces the agency rent function, as shown above. Hence, we have a contradiction, and we must have \( p^F = \arg \min_{p \in \mathcal{P}} \mathbb{E}_{q}[s^*(p)] \).

We can explicitly construct the technology \( p^F \). Note that it satisfies, for \( i \in \{0, \ldots, n-1 \}, \)
\[ p_i^E - q_i = \frac{q_i}{q_0} (p_0^E - q_0), \] that is, for all \( i, j \in \{0, \ldots, n-1\} \), \( l_i (p^E) = l_j (p^E) \), and we then have \( p_n^E = 1 - \sum_{i=0}^{n-1} p_i^E \), such that only \( p_0^E \) needs to be determined by \( \mathbb{E}_{p^E}[x] - \mathbb{E}_{q}[x] = \pi \). We have

\[
\begin{aligned}
\mathbb{E}_{p^E}[x] - \mathbb{E}_{q}[x] &= \pi \iff \sum_{i=0}^{n} (p_i^E - q_i) x_i = \pi \iff \sum_{i=0}^{n-1} \frac{q_i}{q_0} (p_0^E - q_0) x_i - x_n \sum_{i=0}^{n-1} \frac{q_i}{q_0} (p_0^E - q_0) = \pi \\
\iff \frac{p_0^E - q_0}{q_0} \left( \sum_{i=0}^{n-1} q_i x_i - x_n (1 - q_n) \right) = \pi \iff \frac{p_0^E - q_0}{q_0} \left( \mathbb{E}_{q}[x] - x_n \right) = \pi \\
\iff p_0^E - q_0 = -\frac{\pi q_0}{x_n - \mathbb{E}_{q}[x]} < 0.
\end{aligned}
\]

As a result, we have \( p_0^E - q_0 = -\frac{\pi q_0}{x_n - \mathbb{E}_{q}[x]} \), for all \( i \in \{1, \ldots, n-1\} \), \( p_i^E - q_i = -\frac{\pi q_i}{x_n - \mathbb{E}_{q}[x]} \), and

\[
\begin{aligned}
p_n^E - q_n &= -\sum_{i=0}^{n-1} (p_i^E - q_i) = \sum_{i=0}^{n-1} \frac{\pi q_i}{x_n - \mathbb{E}_{q}[x]} = \frac{\pi \sum_{i=0}^{n-1} q_i}{x_n - \mathbb{E}_{q}[x]} = \pi \left( 1 - q_n \right)
\end{aligned}
\]

which completes the proof. \( \blacksquare \)

### A.8 Proof of Proposition 6

From the proof of Proposition 4 (see Appendix A.6) we know that the agency rent function of agent \( p^D \) is linear. Given the (weak) convexity of the agency rent functions, the principal can only prefer \( p^E \) to \( p^D \) for some region of costs of effort \( e \) if the initial slope of the agency rent function of the agent \( p^E \) is lower than the slope of the agency rent function of agent \( p^D \). The initial slope of the agency rent function for \( p^E \) is given by \( \frac{q_0 (x_n - \mathbb{E}_{q}[x])}{\pi (1 - q_n)} \) (see Proposition 2 and use \( p_n^E \)). The slope for the agency rent function of \( p^D \) is given by \( \frac{(q_1 + \cdots + q_n) x_1}{\pi} \) (see Proposition 2 and use \( p_1^D \)). Hence, we have

\[
\frac{(q_1 + \cdots + q_n) x_1}{\pi} > \frac{q_n (x_n - \mathbb{E}_{q}[x])}{\pi (1 - q_n)} \iff 1 - q_0 > \frac{q_n}{1 - q_n} \left( \frac{x_n - \mathbb{E}_{q}[x]}{x_1} \right).
\]

Further, we know that

\[
\mathbb{E}_{q} \left[ s^* (p^D) \right]_{c=\pi} = \sum_{i=1}^{n} q_i x_1 < \mathbb{E}_{q}[x] = \mathbb{E}_{q} \left[ s^* (p^E) \right]_{c=\pi}. \]

53
Due to the linearity of $E_q[s^*(p^D)]$ in $c$ and the (weak) convexity of $E_q[s^*(p^E)]$ in $c$, there exists a unique crossing point if (11) holds and $E_q[s^*(p^E)] > E_q[s^*(p^D)]$ for all $c \in (0, \pi)$ otherwise.

Simple algebra shows that $l^*(p^D) = l^*(p^E)$ implies the inequality condition for the first case of the proposition. ■

A.9 Proof of Lemma 2

Footnote 14 in Section 4.2.1 specifies the constraints on the set of technologies $\mathcal{P}$. Ignoring the single-peaked assumption, all remaining constraints are weak inequality and equality constraints, given by $\forall j \in \Omega : \sum_{i=0}^{j}(p_i - q_i) \leq 0$, $\sum_{i=0}^{n}(p_i - q_i) = 0$, and $\mathbb{E}_p[x] - \mathbb{E}_q[x] = \pi$. The single-peaked assumption can also be written as a set of inequality constraints. Hence, the set $\mathcal{P}$ is defined by a set of weak inequality and equality constraints and is therefore closed. We also have $\mathcal{P} \subset [0, 1]^{n+1}$, and the set is clearly bounded. Hence, $\mathcal{P}$ is compact. ■

A.10 Proof of Lemma 3

Lemma A.1 in Appendix A.2 shows that $p_0 < q_0$ if $\pi > 0$. Let $m \in \{1, \ldots, n\}$ such that, for all $i \in \{0, \ldots, m-1\}$, $p_i \leq q_i$, and $p_m > q_m$.

I next show that, for all $i \in \{m+1, \ldots, n\}$, $p_i \geq q_i$. Assume that this is not the case, and there exists an $h \in \{m+1, \ldots, n\}$, such that $p_h < q_h$. Assume that $h = n$. Lemma A.1 shows that we must have $p_n \geq q_n$, and there is a contradiction. Assume that $h < n$. This implies that $l_0(p) < 0$, $l_m(p) > 0$, $l_h(p) < 0$, and $l_n(p) \geq 0$, which violates Assumption 2. Hence, we must have, for all $i \in \{m+1, \ldots, n\}$, $p_i \geq q_i$.

If it exists, let $j \in \{m, \ldots, n\}$ such that for all $i \in \{m, \ldots, j\}$, $p_i > q_i$, and $p_j = q_j$. If $j < n$, it remains to show that, for all $i \in \{j+1, \ldots, n\}$, $p_i = q_i$. Assume that this is not the case, and there exists an $h \in \{j+1, \ldots, n\}$, such that $p_h > q_h$. This implies that $l_0(p) < 0$, $l_m(p) > 0$, $l_j(p) = 0$, and $l_h(p) > 0$, which violates Assumption 2. ■
A.11 Proof of Lemma 4

This result follows directly from the construction of contracts in Proposition 1 and the assumption about \( p \) in the Lemma. If \( j \in \{m, \ldots, n\} \) such that for all \( i \in \{m, \ldots, j\} \), \( p_i > q_i \), and for all \( i \in \{j + 1, \ldots, n\} \), \( p_i = q_i \), then we have, for \( i \in \{1, \ldots, j\} \), \( L_i(p) > 0 \), and for all \( i > j \), \( L_i(p) = 0 \). The principal therefore never includes contract tranches corresponding to states exceeding state \( j \), since these states generate no incentives for the agent. The principal might pay tranches in all states \( i \leq j \). If \( c = \pi \), the optimal contract exhausts the payoff space with positive cumulative likelihood ratios, that is, for all \( i \in \{1, \ldots, j\} \), \( L_i(p) > 0 \), and for all \( i > j \), \( L_i(p) = 0 \). ■

A.12 Proof of Proposition 7

Lemma 4 shows that, for agent \( p \), we have, for all \( i \in \Omega \) and for all \( c \in [0, \pi] \), \( s^*_i(p) \leq \min \{x_i, x_{m_1+j}\} \), and \( s^*_i(p) = \min \{x_i, x_{m_1+j}\} \iff c = \pi \). For agent \( \tilde{p} \), we have, for all \( i \in \Omega \) and for all \( c \in [0, \pi] \), \( s^*_i(\tilde{p}) \leq \min \{x_i, x_{m_2+j}\} \), and \( s^*_i(\tilde{p}) = \min \{x_i, x_{m_2+j}\} \iff c = \pi \).

The agency rent functions therefore satisfy

\[
\mathbb{E}_q[s^*(p)]|_{c=\pi} = \mathbb{E}_q[\min \{x, x_{m_1+j}\}] < \mathbb{E}_q[\min \{x, x_{m_2+j}\}] = \mathbb{E}_q[s^*(\tilde{p})]|_{c=\pi},
\]

since \( m_1 + j < m_2 + j \iff m_1 < m_2 \). Since the agency rent functions are continuous and increasing, there exists a \( \tilde{c} \in [0, \pi] \) such that for all \( c > \tilde{c} \), \( \mathbb{E}_q[s^*(p)] < \mathbb{E}_q[s^*(\tilde{p})] \). ■

A.13 Proof of Proposition 8

Consider an agent \( p \in \mathcal{P}_{\pi_1} \) and \( p \notin \mathcal{P}^*_\pi_1 \). Hence, there exists an agent \( p^* \in \mathcal{P}^*_\pi_1 \), such that \( \mathbb{E}_q[s^*(p^*)] < \mathbb{E}_q[s^*(p)] \).

We can then scale the technology \( p^* \) by \( \lambda \in \left(\frac{\pi_1}{\pi}, 1\right] \), that is, define a new technology \( \tilde{p}(\lambda) \) by \( \tilde{p}(\lambda) := q + \lambda (p^* - q) \), such that

\[
\mathbb{E}_{\tilde{p}(\lambda)}[x] - \mathbb{E}_q[x] = \lambda (\mathbb{E}_{p^*}[x] - \mathbb{E}_q[x]) = \lambda \pi_1 \leq \pi_1.
\]
Consider the principal’s expected payoff with agent $\tilde{p}(\lambda)$, given by

$$P(\lambda) := \mathbb{E}_q[x] + \lambda \pi_1 - c - \mathbb{E}_q[s^*(\tilde{p}(\lambda))].$$

We know that

$$P(\lambda = 1) = \mathbb{E}_q[x] + \pi_1 - c - \mathbb{E}_q[s^*(p^*)] > \mathbb{E}_q[x] + \pi_1 - c - \mathbb{E}_q[s^*(p)].$$

Since $\left(\frac{c}{\pi_1}, 1\right) \ni \lambda \mapsto P(\lambda)$ is continuous (see Proposition 3 with the proof in Appendix A.5 to see that $\mathbb{E}_q[s^*(\tilde{p}(\lambda))]$ is continuous in $\lambda$), there exists a $\hat{\lambda} < 1$ such that

$$P(\hat{\lambda}) = \mathbb{E}_q[x] + \hat{\lambda} \pi_1 - c - \mathbb{E}_q[s^*(\tilde{p}(\hat{\lambda}))] > \mathbb{E}_q[x] + \pi_1 - c - \mathbb{E}_q[s^*(p)].$$

Set $\tilde{p} := \tilde{p}(\hat{\lambda})$, which has an expected value of effort of $\pi_2 := \hat{\lambda} \pi_1 < \pi_1$, and the principal prefers agent $\tilde{p}$ to agent $p$. ■

**B Principal’s Technology**

In the analysis of the paper, I study the principal’s optimal choice of agent, taking the principal’s technology $q$ as given. In this section, I discuss the role of the principal’s technology $q$. In this setting, I fix the impact of an agent’s effort by describing an agent’s technology using a change in the probability distribution $\Delta \in \mathbb{R}^{n+1}$. Specifically, the probability distribution of cash flows without effort is given by $q$, and the probability distribution with effort is given by $q + \Delta$. When $q$ is fixed, it is equivalent to describe an agent by either $p$ or $\Delta = p - q$.

Consider a principal’s technology $q$ and an agent’s technology $\Delta$, such that Assumptions 1, 2, and 7 are satisfied. I consider small changes in $q$ that leave the ordering of the cumulative likelihood ratio unchanged. Let $c \in (0, \pi)$ and $s^*$ be the optimal contract from Proposition 1, where I omit the dependence of the optimal contract on the agent’s type. The proof of Proposition 1 shows that, keeping $\Delta$ and the ordering of the cumulative likelihood ratio fixed, the optimal contract is independent of $q$. As discussed in Section 3.2, the expected utility of the principal under the
optimal contract is given by (3), that is, \( E_q[x] + \pi - c - E_q[s^*] \).

A change in the principal’s technology \( q \) has two effects. First, it changes the expected value of the principal’s technology, \( E_q[x] \). Second, it changes the agency rent, \( E_q[s^*] \). As in the case of agents’ technologies, I can fix the expected value of the principal’s technology, \( E_q[x] \), that is, I consider changes in \( q \) that keep \( E_q[x] \) unchanged. In this case, while a small change in the principal’s technology \( q \) does not affect the optimal contract \( s^* \), it affects the agent’s rent \( E_q[s^*] \), the cost of the optimal contract to the principal.

Consider an agent who is paid in a region \( \{j, \ldots, n\} \subset \Omega \), where \( j > 1 \), that is, \( s^*_i > 0 \iff i \in \{j, \ldots, n\} \). Reducing the probability under \( q \) in a state \( i \in \{j, \ldots, n\} \) and increasing the probability under \( q \) of a state \( h \in \{0, \ldots, j-1\} \) (accompanied by a reduction in the probability of state 0 and an increase in the probability of state 1 to keep \( E_q[x] \) constant) reduces the agency rent. The intuition is simple. The change in the probabilities from the agent’s effort determines how much the agent is paid in each state. The cost to the principal depends on the probability of these states when the agent shirks, since this determines the agent’s rent. Hence, if the region in which the principal pays the agent has a lower probability under her own technology, the agency rent is lower, since the agent can gain less from shirking. From the principal’s perspective, technologies therefore exhibit a complementarity. The principal prefers agents who improve and get paid in regions of her cash flow distribution that are unlikely under her own technology.

C Risk-Averse Agents

In this section, I discuss the case of risk-averse agents. I assume that agents’ utility from a contractual payoff \( s_i \) is measured by a utility function \( u : \mathbb{R} \mapsto \mathbb{R} \), were \( u \) is increasing, (strictly) concave, and differentiable with \( u(0) = 0, u'(0) \in (0, \infty) \), and \( \lim_{x \to \infty} u'(x) = 0 \). In particular, if the principal offers a contract \( s \) to an agent \( p \in \mathcal{P} \), and the agent exerts effort, the agent’s expected utility is given by \( E_p[u(s)] - c \).

Consider an arbitrary agent \( p \in \mathcal{P} \). An optimal incentive compatible contract that satisfies the
agent’s limited liability, denoted by $s^*(p)$, satisfies

$$s^*(p) \in \arg\max_s \mathbb{E}_p [x - s]$$

subject to

$$\mathbb{E}_p[u(s)] - c \geq \mathbb{E}_q[u(s)],$$

$$\mathbb{E}_p[u(s)] - c \geq 0,$$

$$\forall i \in \Omega : s_i \geq 0.$$  

As in the limited liability benchmark in Section 3.1, the incentive constraint implies the participation constraint, and the incentive constraint binds. Since the agent has to be paid at least 0 in all states but a positive amount in some states to satisfy the incentive constraint, we have

$$\mathbb{E}_p[u(s^*(p))] - c = \mathbb{E}_q[u(s^*(p))] > 0.$$ 

In particular, the agent earns a rent equal to $\mathbb{E}_q[u(s^*(p))] > 0$. I rewrite the incentive constraint as

$$\mathbb{E}_p[u(s)] - c \geq \mathbb{E}_q[u(s)] \Leftrightarrow \sum_{i=0}^n (p_i u(s_i) - c) \geq \sum_{i=0}^n (p_i - q_i) u(s_i) \geq c.$$ 

The principal’s optimization problem can then be written as

$$\max_s - \sum_{i=0}^n p_i s_i$$

subject to

$$\sum_{i=0}^n (p_i - q_i) u(s_i) \geq c,$$

$$\forall i \in \Omega : s_i \geq 0.$$  

The necessary and sufficient conditions for an optimal contract $s^*$ are as follows (see Léonard and Van Long, 1992):

1. $\sum_{i=0}^n(p_i - q_i)u(s_i^*) - c \geq 0$, $\mu \geq 0$, and $\mu (\sum_{i=0}^n(p_i - q_i)u(s_i^*) - c) = 0$.  

58
2. For all $i \in \Omega$: $-p_i + \mu(p_i - q_i)u'(s_i^*) \leq 0$ and $s_i^* \geq 0$.

3. For all $i \in \Omega$: $s_i^* (-p_i + \mu(p_i - q_i)u'(s_i^*)) = 0$.

I first show that $\mu > 0$ and $\sum_{i=0}^n (p_i - q_i)u(s_i^*) = c$. Assume that this is not the case, then $\mu = 0$, and we have for all $i \in \Omega$: $s_i^* p_i = 0$. In particular, for all $i \in \Omega$ with $p_i \geq q_i$, we have $p_i > 0$, which implies $s_i^* = 0$, a contradiction, since the contract would not satisfy the incentive constraint otherwise.

I next show that $p_i \leq q_i$ implies $s_i^* = 0$. Assume that this is not the case, then there exists an $i \in \Omega$ with $p_i \leq q_i$ and $s_i^* > 0$. In particular, this implies $-p_i + \mu(p_i - q_i)u'(s_i^*) = 0 \Leftrightarrow p_i = \mu(p_i - q_i)u'(s_i^*)$. If $p_i < q_i$, then this implies $p_i < 0$, since $u' > 0$, a contradiction. If $p_i = q_i$, this implies that $p_i = q_i = 0$, a contradiction.

Hence, consider states $i \in \Omega$ with $p_i > q_i$. We get the following result.

**Lemma C.1.** Let $i \in \Omega$ with $p_i > q_i$. If $u'(0) \leq \frac{p_i}{\mu(p_i - q_i)}$, then $s_i^* = 0$. If $u'(0) > \frac{p_i}{\mu(p_i - q_i)}$, then $s_i^* = (u')^{-1}\left(\frac{p_i}{\mu(p_i - q_i)}\right) > 0$.

**Proof.** To show this, let $s_i^* > 0$. We then have

$$-p_i + \mu(p_i - q_i)u'(s_i^*) = 0 \Leftrightarrow u'(s_i^*) = \frac{p_i}{\mu(p_i - q_i)}.$$

If $\frac{p_i}{\mu(p_i - q_i)} < u'(0)$, then we have

$$s_i^* = (u')^{-1}\left(\frac{p_i}{\mu(p_i - q_i)}\right) > 0.$$

If $\frac{p_i}{\mu(p_i - q_i)} \geq u'(0)$ we must therefore have $s_i^* = 0$.

It remains to show that if $\frac{p_i}{\mu(p_i - q_i)} < u'(0)$ we have $s_i^* > 0$. Assume that this is not the case, and there exists an $i \in \Omega$ with $\frac{p_i}{\mu(p_i - q_i)} < u'(0)$ and $s_i^* = 0$. We must then have

$$-p_i + \mu(p_i - q_i)u'(0) \leq 0 \Leftrightarrow u'(0) \leq \frac{p_i}{\mu(p_i - q_i)},$$

a contradiction. ■
We then get the following characterization of the optimal contract.

**Lemma C.2.** Let \( s^*_i > 0 \) and \( s^*_j > 0 \). We then have \( s^*_i > s^*_j \iff \frac{p_i - q_i}{q_i} > \frac{p_j - q_j}{q_j} \).

**Proof.** We have

\[
\begin{align*}
  s^*_i > s^*_j & \iff (u')^{-1}\left(\frac{p_i}{\mu(p_i - q_i)}\right) > (u')^{-1}\left(\frac{p_j}{\mu(p_j - q_j)}\right) \\
  & \iff \frac{q_i + (p_i - q_i)}{p_i - q_i} < \frac{q_j + (p_j - q_j)}{p_j - q_j} \\
  & \iff \frac{p_i - q_i}{q_i} > \frac{p_j - q_j}{q_j},
\end{align*}
\]

since we must have \( p_i > q_i \) and \( p_j > q_j \).

In particular, the principal pays the agent only in states with a positive likelihood ratio. Further, the principal pays the agent more in a state with a higher likelihood ratio. In the risk-neutral limited liability benchmark in Section 3.1, the principal pays the agent only in the state with the highest likelihood ratio.

This result shows that, if contracts also have to satisfy the monotonicity constraints from Assumptions 5 and 6, the general insight from Section 5 also applies in the case of risk aversion. If an agent has a high impact on the principal’s project in low states of the world (the agent has a more debt-like technology), the principal pays the agent most in these states in the limited liability benchmark. Introducing monotonicity constraints forces the principal to pay the agent at least the same amount in higher states as well. This is true even in cases when the productivity of effort is high. If an agent has a high impact on the principal’s project in high states of the world (the agent has a more equity-like technology), the principal pays the agent most in these states in the limited liability benchmark. If the productivity of effort is high, the payments to the agent are relatively low, and the effect of introducing the monotonicity constraints is limited. However, if the productivity of effort is low, the principal has to pay the equity-like agent a high share in high states, and the monotonicity constraints become very costly.