RESEARCH INSTITUTE OF INDUSTRIAL ECONOMICS

IFN Working Paper No. 1171, 2017

Delegation of Investment Decisions, and Optimal Renumeration of Agents

Assar Lindbeck and Jörgen Weibull

Research Institute of Industrial Economics
P.O. Box 55665
SE-102 15 Stockholm, Sweden
info@ifn.se
www.ifn.se

DELEGATION OF INVESTMENT DECISIONS, AND OPTIMAL REMUNERATION OF AGENTS

Assar Lindbeck* and Jörgen Weibull†

December 5, 2019. Revised August 2, 2020.‡

Abstract. We analyze an investor who delegates information acquisition and investment decisions to an agent. The investor cannot monitor the agent's effort or information. Optimal pay schemes contain bonuses that increase with the net return rate of the investment, but, unlike conventional contracts, at a decreasing rate. Moreover, investments with low return rates are penalized, again unlike conventional contracts. Nevertheless, it may be optimal for the investor to reward the agent above the agent's reservation utility. We examine the role of the agent's risk attitude for the shape of the pay scheme, and whether firing after bad investments is a more effective threat than reduced pay. We also analyze how the nature of the contract changes if the agent is given bargaining power.

Keywords: Delegation, principal-agent, principal-expert, investment, information acquisition, rational inattention, contract, bonus, penalty.
JEL codes: D01, D82, D86, G11, G23, G30.

---

*Institute for International Economic Studies, at Stockholm University, and Research Institute of Industrial Economics, Stockholm. Assar Lindbeck thanks the Jan Wallander and Tom Hedelius Foundation, and the Marianne and Marcus Wallenberg Foundation for financial support.

†Stockholm School of Economics (SSE), and the Institute for Advanced Study in Toulouse (IAST). Jörgen Weibull thanks the Knut and Alice Wallenberg Research Foundation, and the Agence Nationale de la Recherche, Chaire IDEX ANR-11-IDEX-0002-02, for financial support.

‡We thank the editor in charge and two anonymous referees for helpful comments. We also thank Sylvestre Boittin-Duchesne for excellent research assistance, and Atahan Afsar, Mogens Fosgerau and Joacim Tåg for helpful comments to earlier versions of this paper. A precursor of parts of this study is reported in IFN WP 1171, "Investment, rational inattention, and delegation" (Lindbeck and Weibull, 2017).

1. Introduction

In standard principal-agent models, the task of the agent is to increase the success probability of an agreed-upon project. A complementary, and arguably equally important task for many agents, such as CEOs of large corporations, pension-fund managers and consultants, is to make a well-informed choice of which project, if any, to undertake. To be well-informed usually requires skill and effort to acquire, process, and assess relevant information, and such efforts, as well as the obtained information, are usually difficult to monitor. This class of moral-hazard problems, sometimes called principal-expert problems, was noted and discussed already by Demski and Sappington (1987), but has received less attention among model builders than standard principal-agent models. We here develop an operational principal-agent (or principal-expert) model in order to analyze precisely such moral-hazard problems.

In our model, an investor delegates information acquisition and investment decisions to a risk-neutral or risk-averse agent, subject to a limited-liability constraint. We take the agent to be someone who has a comparative advantage in acquiring and evaluating pertinent information about projects or investment opportunities. Canonical examples are CEOs of large corporations, division managers, and fund managers. Other examples exist in the consulting industry, where firms, institutions and wealthy individuals sometimes delegate information acquisition and, in effect also decision-making, to a consultant or expert.

Information about investment projects is often a mix of hard and soft pieces. To assess the relevance and reliability of such information for balanced decision-making requires expertise, effort and time. Moreover, the information obtained is often difficult to communicate to a non-expert. There are two obstacles to communication. First, the principal may not have the expertise or time needed. The project may, for example, concern a new medical product, technical innovation, or the entry into a foreign market. Second, it may be in the agent's self-interest not to share all information. The agent may, more generally, opportunistically mis-report, suppress, or even distort information. For this reason, not only information acquisition but also the investment decision itself is in practice often formally or informally delegated to the agent. Even if the principal makes the formal decision, she may have to rely on the information and advise given by the agent (see Aghion and Tirole, 1997). The key issue is thus how to motivate the agent to acquire relevant and reliable information and, when making or recommending the investment decision, to use that information in the principal's interest. As pointed out by Demski and Sappington (1987), there is in general a tension between these two objectives.

The topic being rich and complex, we abstract from many important real-life factors and focus only on a few key elements.[1] We assume that the investor cannot

[1]For recent contributions to the topic of management compensation, see e.g. Kaplan and Ström-

monitor the agent's efforts or information. A contract between the two parties can only be conditioned on whether investment is made or not, and if made, its realized net return. In particular, the agent's remuneration in case of non-investment cannot be conditioned on the return that would have materialized, had investment been made.[2] In order to keep back the agent's potential eagerness to invest, because of his hope to earn a bonus, the agent has to be paid also for not investing, and has also to face some sort of penalty if making a loss-creating investment. The penalty may take pecuniary form or take the form of firing. We require pecuniary penalties to meet the limited-liability constraint of never resulting in a net pay from the agent to the principal. To deter the agent from investing when prospects—about which he has private information—are not good is just as important for the principal as to motivate the agent to invest when prospects are good. Under the mentioned limited-liability constraint, the magnitude of pecuniary penalties after bad investments is bounded by the pay after non-investment, since the harshest pecuniary penalty is to withhold all payment, a loss for the agent equal in size to his pay when not investing. Firing may be a more powerful penalty, especially if this is harmful for the agent's future career, and if it is not very costly for the principal. The limited-liability constraint causes an asymmetry between bonuses and pecuniary penalties, even for risk neutral agents facing investment projects that are *ex ante* symmetric in terms of the probability distribution of potential losses and gains.

We show that optimal contract for risk-neutral agents belong to a family of well-behaved non-linear functions known in physics, but, to the best of our knowledge, not known in economics. Our model's optimal contracts, for risk-neutral as well as risk-averse agents, are non-linear and reward investments with very high return rates less than proportionally. Investments that result in low or negative return rates are penalized. The reason for the non-linearity is that contracts here serve two distinct purposes: to incentivize the agent's information acquisition, and, once information has been obtained, to align the agent's incentives with those of the principal at the investment decision (had the principal then had access to the agent's private information). It is the first objective, giving incentives for information acquisition under the limited-liability constraint that lies behind the additional non-linearity, also for risk-neutral agents. We model the agent's information acquisition within a rational-inattention approach, whereby the agent optimally reduces his uncertainty, given the cost of uncertainty reduction and given prior beliefs and contract. Uncertainty is

berg (2002,2004), Friebel and Rath (2004), Gabaix and Landier (2008), Terviö (2008), Kaplan and Rauh (2010), Murphy (2012), Kaplan (2013), Ibert et al (2018), and Ma, Tang and Gómez (2019).

[2] This restriction is natural in many situations, since it is usually not possible to know and in court verify, *ex post*, what the return to by-passed investment opportunities would have been However, in some situations, such as investment in stocks, this information may be available and, if verifiable, can then be part of a contract. This is assumed in Carroll (2019).

measured by the Shannon entropy, whereby the marginal cost of uncertainty reduction is low when there is lots of uncertainty and tends to plus infinity as uncertainty vanishes.

The threat of firing the agent after an unprofitable investment may be more effective than the threat to reduce his pay, especially for agents for whom career concerns are important (assuming that being fired acts as a negative signal in the future market for the agent). We also find that if the agent has bargaining power when the contract is set up, then this enhances his salary more than the bonuses. As for the risk attitude of the agent, one might conjecture that it is more profitable for the principal to hire a risk-neutral than a risk-averse agent, since the latter will need to be better insured against bad outcomes. But this is not the case. On the contrary, the principal benefits from hiring a risk-averse agent. The reason is that the above-mentioned asymmetry between bonuses and penalties (due to the limited-liability constraint) is mitigated, since for a risk-averse agent income reductions are more painful.

We use numerical simulations throughout in order to illustrate the nature of optimal contracts, and to show how they depend on the nature of the return distribution and other factors. We find that the presence or absence of extremely rare and high return rates ("black swans") does not much influence the shape of optimal contracts. In such rare states of nature, optimal contracts reward investment by way of only a small share of the realized net return (unlike the fixed share under standard profit-share or stock-based contracts).

The presentation of the material is organized as follows. In Section 2 we outline the base-line model and provide a necessary first-order condition for optimality when the agent's participation constraint is slack. Section 3 is devoted to the case of a risk-neutral agent, and we show that optimal contracts take a certain mathematical form that, to the best of our knowledge, has not appeared before in contract theory. Section 4 defines a class of risky projects that we use to illustrate the nature of optimal contracts. Section 5 compares optimal contracts with standard profit-share and stock-based contracts. Section 6 analyzes bargaining between a principal and a risk-neutral agent. In Section 7 we analyze optimal contracts for risk averse agents. Section 8 provides a generalization of the base-line model by allowing for firing clauses in contracts. Section 9 discusses the related literature, and Section 10 concludes. Mathematical proofs and complementary numerical results are given in an appendix at the end of the manuscript.

2. Model

A risk-neutral investor faces an indivisible investment opportunity, or *project*. The project requires a fixed lump-sum investment, $I > 0$, and gives a random return (or cash flow), Z. The project's *net return rate* is defined as

$$X = \frac{Z - I}{I} - r, \tag{1}$$

where r is the investor's unit cost of funds (a risk-free interest rate or the return to treasury bonds). This net return rate X from the project is taken to be a random variable with finite support $M = \{x_1, ..., x_m\}$, where $x_1 < x_2 < ... < x_m$. The investor's prior probability for each *state of nature* $\omega = \Omega = \{1, 2, ..., m\}$ is positive and denoted μ_k. The vector $\mu = (\mu_1, ..., \mu_m)$ is thus the investor's prior belief concerning the true state of nature. This may e.g. be based on public information, or knowledge about the economy at large, or about the industry in question, and/or on freely available information about the project at hand. We assume throughout that $x_1 < 0 < x_m$.

The investor considers the possibility of hiring an agent who has some comparative advantage in obtaining and processing relevant information about investment projects. We assume that the principal knows the agent's information cost-function and risk attitude. The agent strives to maximize the expected utility from his remuneration. His Bernoulli function of income, $u : \mathbb{R}_+ \to \mathbb{R}_+$, is taken to be twice differentiable with $u(0) = 0$, $u' > 0$ and $u'' \leq 0$.[3] The agent's outside option has expected utility $\bar{u}$.

We assume that the agent shares the principal's prior. However, if the agent is hired, then the principal will not know the agent's effort to acquire and process information, the quality of the information so obtained, or what that information is. The only verifiable information upon which a contract can be based is whether or not investment was made, and, if made, its subsequently realized return. In view of this informational asymmetry, it is immaterial for the subsequent analysis if the principal asks the agent for investment advice or if the principal instead delegates the investment decision to the agent.

In the base-line model, outlined in this section, we focus on purely pecuniary contracts. In a Section 6, we analyze contracts with a firing clause. Hence, for now, a contract is a vector $\vec{y} = (y_0, y_1, ..., y_m) \in \mathbb{R}^{m+1}$, where y_0 is the agent's pay if he does not invest, and y_k, for $k = 1, ..., m$, is his pay if he invests and the state of nature turns out to be $k \in \Omega$. We focus on contracts that meet the limited-liability constraint of never requiring the agent to make a net payment to the principal; all components of the vector $\vec{y}$ have to be non-negative. Such contracts will be called *feasible*. A contract $\vec{y}$ can be interpreted as a *salary* and a package of *bonuses* and *penalties*. In this interpretation, y_0 is the salary and, for each state $k \in \Omega$, the

[3] We will later make one exception and consider the case when u is logarithmic and thus only defined for strictly positive arguments.

difference $t_k = y_k - y_0$, if positive, is the agent's *bonus* for investing in state k, while the difference $y_0 - y_k$, if $t_k = y_k - y_0$ is negative, is the agent's *penalty* for investing in that state. Expressed in these terms the limited liability constraint thus requires no penalty to exceed the salary.

Acquisition and processing information is costly to the agent. We analyze this within a rational-inattention approach. This currently very active research field was pioneered by Sims (2003, 2006), followed up and developed further by Woodford (2008), Cabrales, Gossner and Serrano (2013), Matejka and McKay (2015), Yang (2015), Lindbeck and Weibull (2017), Caplin, Dean, and Leahy (2019), Yang and Zeng (2019), and Fosgerau et al. (2020). This approach views information acquisition as a choice of a joint probability distribution over signals and states of nature under the constraint that the marginal distribution over states equals the decision maker's — here the agent's prior — and with information costs represented in terms of entropy reduction.[4]

Using this approach, and adapting it to a principal-agent setting, we obtain investment probabilities that depend on the agent's optimal choice of how well-informed he wants to be, given the project, his risk attitude, his information costs and contract. By spending more time and effort on information acquisition, the agent can reduce the risk of investing in states of nature where he earns little, and enhance the chances for investing in states where he earns much. An information-cum-investment strategy for the agent consists of (i) a decision of how much effort to spend on acquiring and processing information, (ii) a decision rule that specifies what investment decision to make conditional upon the information obtained.

If the contract specifies the same or similar pay to the agent from investment in all states of nature, or if information acquisition is very costly, then it may be optimal for the agent not to acquire any information. He will then base his investment decision directly on his prior and his pay after investment and non-investment, respectively. Likewise, if the agent is always or never paid more when investing than when not investing, he will optimally chose not to acquire information. However, if investment in some states of nature pays below non-investment and investment in other states pays above, and information costs are not very high, then it may be optimal for the agent to make an effort to acquire and process information about which states are more likely to prevail. The agent then strives to obtain a precise binary signal that essentially says "invest" when his pay exceeds the pay from not investing, and "do not invest" in the opposite case.

By thus acquiring and processing information, the agent will replace his prior belief

[4] The information-theoretic interpretation of entropy is due to Shannon (1948), who axiomatized entropy as a measure of uncertainty. The approach was further developed by Shannon and Weaver (1949), Jaynes (1957), Kullback (1959) and Hobson (1969).

about the state of nature by a more precise (and *ex ante* random) posterior belief. Within the rational-inattention approach, this intricate information acquisition *cum* investment problem becomes analytically tractable. In this approach, the information costs associated with a more precise posterior is measured in terms of the expected entropy reduction. When choosing how much effort to make, the agent optimally trades off the (certain and immediate) cost or disutility of acquiring and processing information against the (uncertain and future) benefit from making a well-informed investment decision. At that second stage, when the agent's refined probabilistic belief has been formed, he will face a standard Bayesian decision problem and will decide for or against investment depending on his then expected utility from investing or not investing.[5]

Let q denote the probability for investment at the *ex ante* stage when the agent has accepted the contract and chosen his information-*cum*-investment strategy, but before his private information has been obtained. This *ex ante* investment probability is zero or one if his expected benefit from information acquisition cannot outweigh the cost of obtaining the information. If he thus opts for $q \in \{0,1\}$, he will choose $q = 1$ ($q = 0$) if, according to his prior belief μ, his expected utility from payment when investing exceeds (falls short of) the utility from his (certain) pay when not investing. If he instead opts for $q \in (0,1)$, then he will acquire and process information, and base his investment decision upon the so obtained information. From Theorem 1, Lemma 2 and Corollary 2 in Matejka and McKay (2015), one obtains (after some algebraic manipulation) that the agent's optimal information - *cum* - investment strategy then induces the following conditional investment probabilities in all states of nature $k \in \Omega$:

$$p_k = \Pr\left[\text{invest } \mid \omega = k\right] = \frac{qe^{u(y_k)/c}}{qe^{u(y_k)/c} + (1-q)\, e^{u(y_0)/c}}\ , \qquad (2)$$

where $c > 0$ is the agent's *unit cost of information* (to be detailed below). We note that this conditional investment probability is a continuously and strictly increasing function of the agent's utility difference $u(y_k) - u(y_0)$ between investing in that state and not investing. Moreover, this probability p_k exceeds (falls short of) the prior investment probability q if the payment y_k after investment in that state exceeds (falls short of) the "salary" y_0 obtained when not investing. We also note that the conditional investment probabilities (2) are non-linear in the payments to the agent even if the agent is risk-neutral.

By Bayes' law, the *ex ante* investment probability, q, equals the weighted sum of the conditional investment probabilities, each weighted by the likelihood of that

[5] For excellent discussions, analyses and extensions of the rational-inattention approach, see Matejka and McKay (2015), Caplin, Dean and Leahy (2019), and Fosgerau et al. (2020).

state:

$$\sum_{k \in \Omega} \mu_k p_k = q. \tag{3}$$

This equation is always met if $q = 0$ or $q = 1$. However, for values of q between these extremes, (3) is met if and only if

$$\sum_{k \in \Omega} \frac{\mu_k e^{u(y_k)/c}}{q e^{u(y_k)/c} + (1-q)\, e^{u(y_0)/c}} = 1. \tag{4}$$

This equation characterizes the agent's optimal choice of investment probability q under any given contract.[6] Once q has been solved for, the conditional investment probabilities are immediately obtained from (2). It is easily verified that the left-hand side of (4) is a continuous and strictly convex function of q that takes unit value at $q = 1$. Hence, there exists at most one $q \in (0,1)$ satisfying (3), or, equivalently, (4). In sum, under any contract $\vec{y}$, the agent will either choose $q^* = 0$, $q^* = 1$, or, if it exists, the unique $q^* \in (0,1)$ that satisfies (4).

It follows that $q^* = 1$ if there are no penalties in the contract (that is, if $y_k \geq y_0$ for all $k \in \Omega$), since then the left-hand side in (4) exceeds unity for all $q < 1$). Likewise, $q^* = 0$ if there are no bonuses in the contract (that is, if $y_k \leq y_0$ for all $k \in \Omega$), since then the left-hand side in (4) falls short of unity for all $q < 1$. In the first case, the agent invests with probability one in all states of the world, while in the second case he never invests. However, these two extreme behaviors occur also under less extreme contracts. We will say that a contract $\vec{y}$ provides *sweet* investment conditions to the agent if it is optimal for him to always invest, $q^* = 1$, and *sour* investment conditions if it is optimal for him to never invest, $q^* = 0$. In all other cases, that is when $0 < q^* < 1$, the agent's investment conditions will be called *normal*. Then the agent sometimes invests, and sometimes not, depending on the information he has acquired. Clearly, this is the only case of interest for the principal, since the principal could otherwise do just as well without the agent.

The following result characterizes the three mentioned investment conditions.[7] Writing Y for the random payment to the agent after investment (that is, $Y = y_k$ if the state of nature is $k \in \Omega$):

Lemma 1. *Investment conditions are normal for the agent if*

$$-c \ln \mathbb{E}\left[e^{-u(Y)/c}\right] < u(y_0) < c \ln \mathbb{E}\left[e^{u(Y)/c}\right], \tag{5}$$

[6] The equation also agrees with the necessary and sufficient optimality condition in Proposition 1 in Caplin, Dean and Leahy (2019).

[7] For other results that provide conditions under which a rationally inattentive decision-maker will not acquire any information, see Lemma 2 in Woodford (2008), Proposition 1 in Yang (2015), and, for a recent, general and thorough treatment of this issue, see Caplin, Dean and Leahy (2019).

sour if

$$u(y_0) \geq c \ln \mathbb{E}\left[e^{u(Y)/c}\right], \tag{6}$$

and sweet if

$$u(y_0) \leq -c \ln \mathbb{E}\left[e^{-u(Y)/c}\right]. \tag{7}$$

In other words: with probability one the agent will not invest if his salary is too high, as expressed by inequality (6). Information acquisition being costly for him, he will not acquire any information under such contracts. Likewise, with probability one he will invest if his salary is too low, as expressed by inequality (7). Also in this case it is optimal for him to acquire no information. For contracts that satisfy (5), he will acquire some information and thereafter invest if and only if the obtained information is sufficiently favorable for investment, in terms of his expected utility from investing as compared with not investing under the contract.

The agent's achieved expected utility, when acting optimally under any contract $\vec{y}$ that provides normal investment conditions for the agent, is

$$\begin{aligned} U(\vec{y}) \quad = \quad & \sum_{k \in \Omega} \mu_k p_k^* \cdot [u(y_k) - u(y_0)] \; + \; u(y_0) \\ & - c \cdot \left(H(q^*) - \sum_{k \in \Omega} \mu_k H(p_k^*) \right), \end{aligned} \tag{8}$$

where q^* is the unique solution in $(0,1)$ to (4), the probability that the agent will invest, and p_k^*, for each $k \in \Omega$, is the associated conditional probability that the agent will invest, given the state k (see (2)). $H : [0,1] \to \mathbb{R}_+$ is the (Shannon) entropy function for a binary probability distribution, defined by

$$H(p) = -\, p \ln p - (1-p) \ln (1-p), \tag{9}$$

with the convention $0 \ln 0 = 0$.[8] Consequently, the difference $H(q^*) - \sum_k \mu_k H(p_k^*)$ is the expected *entropy reduction* when moving from the prior investment probability, q^*, to the (*ex ante* random) posterior investment probability, p_k^*, in the current state of nature, $k \in \Omega$.[9] Maximal entropy reduction would be obtained if the agent were to know the true state of nature, and thus invest precisely in those states in which he obtains a bonus. However, it is prohibitively costly, and hence suboptimal for

[8] The marginal cost of reducing entropy is thus infinite at the boundaries: $H'(p) \to +\infty$ as $p \downarrow 0$ and $H'(p) \to -\infty$ as $p \uparrow 1$.

[9] We note that entropy reduction is symmetric in the sense that it is invariant under permutation of the states of nature. This is arguably a strong assumption, since in practice some states may be harder to identify than others. See Fosgerau et al (2020) for a new generalized entropy that permits asymmetry.

the agent to obtain such precise information. The agent thus has to trade off his information costs against the benefit for him of making a better informed investment decision.

What contract, if any, will a rational and risk-neutral principal propose the agent? First, the contract has to meet the agent's participation constraint that his *ex ante* expected utility under the contract does not fall short of his reservation utility:

$$U(\vec{y}) \geq \bar{u}, \tag{10}$$

where $\bar{u}$ is the agent's reservation utility.

Second, the contract must be such that it provides normal investment conditions for the agent. This is, precisely, (5). In essence, this condition requires that the salary, bonuses and penalties should in a precise sense be "well balanced". In particular, there should be enough bonuses (positive differences $y_k - y_0$) and penalties (positive differences $y_0 - y_k$) to incentivize the agent to acquire information.

Third, if there are contracts that meet the above two requirements, then an optimal contract should yield the highest possible expected profit to the principal among these. In order to state this condition, we need to step backwards in time to the moment when a contract is offered to the agent, and identify the principal's expected profit from hiring the agent under any given contract $\vec{y}$. This expected profit is

$$\Pi(\vec{y}) \quad = \quad \sum_{k\in\Omega} \mu_k p_k^* \cdot (Ix_k - y_k) \quad - \quad (1 - q^*)\, y_0. \tag{11}$$

In words, this is the sum of the net profits to the principal in all states of nature, each weighted by the probability that the agent will invest in that state, minus the pay to the agent if he does not invest, 's the probability for that event.

Under (5), the agent acquires information before making the investment decision and then $q^* \in (0,1)$ is uniquely determined by (3). If the contract is such that the agent's investment conditions under the contract are sour, then the agent, if hired, will acquire no information and will not invest. In that case $q^* = 0$, and the principal's net profit is simply $\Pi(\vec{y}) = -y_0$. Likewise, if the contract would render the agent's investment conditions sweet, then the agent would invest "blindly", $q^* = 1$, resulting in expected profit $\Pi(\vec{y}) = \sum_k \mu_k (Ix_k - y_k)$. However, neither "sweet" nor "sour" contracts are of interest to the principal. We are now in a position to state the third condition, that the contract should be optimal for the principal from among all feasible contracts that meet the agent's participation constraint. It should solve the program

$$\max_{\vec{y}\geq 0 \text{ s.t. } (10)} \Pi(\vec{y}). \tag{12}$$

This brings us to the fourth and final condition, namely, that all of this should be worthwhile for the principal. Suppose, thus, that $\vec{y}^*$ solves program (12) and results in expected profit Π^*. It is optimal for the principal to hire the agent under contract $\vec{y}^*$ if and only if $\Pi^* \geq I \cdot \max\{0, \mathbb{E}[X]\}$. The last expression is the profit obtained by the principal single-handedly.[10]

We will henceforth proceed step by step, and consider some special cases, and a variant of this model in which the agent is fired rather than given a penalty.

2.1. A necessary first-order condition. We first identify a necessary condition for a contract to be optimal in situations when the agent's participation constraint is not binding.

Proposition 1. *Suppose that $\vec{y}$ is an optimal contract and that the agent's participation constraint is not binding. Then there exist parameters $\alpha, \xi \in \mathbb{R}$ such that (a) $x_k \leq \xi \Rightarrow y_k = 0$, (b) $x_k > \xi \Rightarrow y_k > 0$, and (c) every $y_k > 0$ is uniquely determined by the equation*

$$y_k \;=\; Ix_k - \frac{c}{u'(y_k)}\left(\frac{q}{1-q}e^{[u(y_k)-u(y_0)]/c} + 1\right) + \alpha, \tag{13}$$

where $q \in (0,1)$ is uniquely determined by (4). This defines the after-investment pay y_k as a continuous function of the net return rate x_k, strictly increasing wherever positive. Moreover, $y_1 < y_0 < y_m$.

In other words, optimal contracts are characterized by a critical net return rate, ξ, such that the agent is paid nothing if he invests and the net return rate falls short of ξ. By contrast, at net return rates above ξ, the agent is paid something if investing, and that pay is continuously and strictly increasing in the realized net return rate. The after-investment pay y_k is lower than the "salary" y_0 in the worst state of nature, $k = 1$, and higher than the salary in the best state of nature, $k = m$. We also note that the shape of the optimal contract does not depend in any direct way on the shape of the prior distribution; the latter enters only in so far as it influences the investment probability q and the parameters α and ξ.

2.2. Private benefits to the agent. Our model presumes that the agent's only concern is his pay. In practice, the agent may derive private benefits from investment in certain states of nature, and potentially also from non-investment. One major example is career concerns. As is well-known such concerns may reduce the need for monetary incentives (see e.g. Ma, Tang and Gomez, 2018). Suppose that the realized net return rate after investment become publicly known *ex post*. This return rate may then influence the agent's future career. This incentive channel may be

[10]Recall that X is defined as the net return rate, that is the return rate net of the risk-free interest rate r available to the principal.

particularly important for young agents. How does such career concerns affect the nature of optimal contracts in the present model?

In order to illustrate the role of private benefits in a tractable form (and in line with Tirole and Aghion, 1997), suppose that the agent's utility from investing in state $k \in \Omega$ is $u(y_k) + \gamma_k$ and from not investing is $u(y_0) + \gamma_0$, where γ_0, γ_1,..., γ_m represent the agent's private benefits. In the case of career concerns, γ_k may be the present utility to the agent from future earnings. Arguably, the higher the realized net return rate x_k after investment, the higher is then γ_k. Proposition 1 and the subsequent analysis easily generalizes to allow for such additive private benefits.

See Appendix for a numerical example in which it is shown how the agent's career concerns enhance the principal's expected profit. The reason simply being that this intrinsic motivation for good performance makes it optimal for the principal to pay smaller or no bonuses after good investments.

3. Risk-neutral agent

Suppose that the agent is risk neutral. An application of Proposition 1 enables us to obtain a closed-form expression, in terms of a transcendental function used in physics, for optimal contracts when the participation constraint is slack, and also when it is not "binding too hard". For any feasible contract $\vec{y}$, all payments y_k after investment can be written in the form

$$y_k = \max\{0, y_0 + \phi(x_k)\} \qquad \forall k \in \Omega, \tag{14}$$

where $\phi : \mathbb{R} \to \mathbb{R}$ is a *transfer function* that specifies the net transfer $t_k = y_k - y_0 = \phi(x_k)$ to the agent if he invests in state $k \in \Omega$. When positive (negative), $\phi(x_k)$ can be interpreted as a bonus (penalty or "malus"). It turns out that if a contract $\vec{y}$ is optimal and the the agent's participation constraint does "not bind too hard" (in a precise sense), then the transfer function ϕ can be explicitly represented in terms of the so-called *Lambert W function*, $W : (-1/e, +\infty) \to \mathbb{R}_+$.[11] The optimal transfer function ϕ is strictly increasing and strictly concave. Hence, since the after-investment payments y_k according to the contract are given in the form (14), these payments are a non-increasing function of the net return rates x_k, but a function that is neither concave nor convex. Formally:

Proposition 2. *If a contract $\vec{y}$ is optimal and the Lagrangian associated with the agent's participation constraint does not exceed unity, then* $y_k = \max\{0, y_0 + \phi(x_k)\}$

[11] To be precise, this is a correspondence used in theoretical physics (when analyzing the Planck, Bose–Einstein, and Fermi–Dirac distributions), in biochemistry (when analyzing enzyme kinetics), and in combinatorics. See Appendix for its definitoin and basic properties.

for all $k \in \Omega$*, where* $\phi : \mathbb{R} \to \mathbb{R}$ *is defined by*

$$\phi(x) = c \cdot \left[\ln W\left(\frac{\sigma}{c} e^{(Ix-\tau)/c}\right) - \ln \frac{\sigma}{c}\right] \quad \forall x \in \mathbb{R}, \tag{15}$$

for parameters $\sigma, \tau \in \mathbb{R}$*, with* $\sigma > 0$*. Moreover,* ϕ *is twice differentiable, strictly increasing and strictly concave, from minus to plus infinity, with* $0 < \phi' < I$*,* $\phi'' < 0$*,* $\lim_{x \to -\infty} \phi'(x) = I$*, and* $\lim_{x \to +\infty} \phi'(x) = 0$*, and* $\phi((\sigma + \tau)/I) = 0$*. Furthermore, if the agent's participation constraint is slack, then* $\sigma = cq/(1-q)$*, where* q *is the unique solution in* $(0,1)$ *to (4).*

In other words: irrespective of how many potential outcomes the project has, what their magnitudes and prior probabilities are, the optimal payment schedule always belongs to a family of mathematically well-known functions, where each member of the family, for given I and c is completely characterized by only two parameters, σ and τ. It follows that the agent receives a bonus (penalty) if and only he invests in a state of nature with net return rate above (below) $(\sigma + \tau)/I$. In states of nature where his total pay (salary plus net transfer) is positive, the payments to the agent follows a strictly increasing and strictly concave curve (in terms of the net return rate) that gradually reduces the share of the gains of trade, if positive, that befalls the agent. Indeed, this share shrinks towards zero as the capital gains tend to plus infinity. In particular, optimal contracts are not linear, and neither concave nor convex; they pay nothing to the agent for all net return rates below a critical value, and they provide positive pay above that critical rate, pay that is a strictly increasing and concave function of the net return rate. When the agent's participation constraint is slack, the parameter σ equals the agent's unit cost of information, c, multiplied by $q/(1-q)$, the odds for investment (before information has been acquired). In particular, when investment is just as likely as non-investment, then σ is the agent's unit information cost, c.

These general qualitative features of net-transfer functions ϕ of the form (15) are illustrated in Figure 1, drawn for $I = 1$, $c = 0.05$, $\sigma = 0.05$ and $\tau = 0.15$. As expected, the graph of the net-transfer function ϕ intersects the x-axis precisely where $x = \sigma + \tau$. For net return rates above this value, a bonus is given, and below this net return rate, penalties are are imposed. The dashed straight lines in the diagram provide upper and lower bounds on the transfer $\phi(x)$. For $x \leq \tau$, the upper bound is $y = Ix - \tau$ (the left-most upward-sloping dashed straight line, for $x \leq \tau$), and the lower bound is $y = Ix - \tau - \sigma$ (the right-most parallel dashed straight line). For $\tau < x < \sigma + \tau$, the upper bound is zero (horizontal dashed line), and the lower bound is the same as for $x \leq \tau$. For $x \geq \sigma + \tau$, the upper bound is (the former lower bound) $y = Ix - \tau - \sigma$ (the right-most upward-sloping dashed straight line) and the lower bound is (the former upper bound) zero.

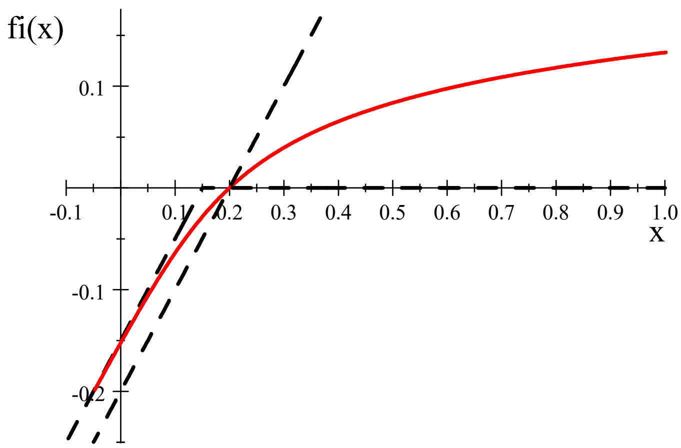


Figure 1: The net-transfer function ϕ.

The above general qualitative observations about solutions ϕ to (29) amounts to saying that the random pay, $Y = \{0, y_0 + \phi(X)\}$ to the agent, when the agent invests, is a security with the properties that $0 \leq \phi(X) < IX$, and such that, whenever the agent's pay is positive, both the agent's pay, $y_0 + \phi(X)$, and the residual that befalls the principal, $V = IX - y_0 - \phi(X)$, is increasing in X. Moreover, since this security, when taking positive values, is strictly concave in the net return rate, the principal's residual is strictly convex in the net return rate. In other words, the higher the capital gain from investment, the smaller is the agent's share. It follows that standard profit-share or stock-share contracts are suboptimal whenever there are more than two potential outcomes. The reason is that they are (piece-wise) linear and hence either pay too little at low return rates or too much at high return rates, see Section 5.

We note, however, that the lower the agent's unit information cost $c > 0$, the closer is the optimal contract to another type of piece-wise linear contracts, namely those that pay nothing if the agent invests in bad states of nature, a linearly increasing pay in intermediate states of nature, and a fixed pay in all good states of nature. In the limit as $c \to 0$, it is as if the agent can costlessly and without error observes the true state of nature. Under an optimal contract, such an agent almost surely earns his salary, y_0, because he almost never invests in states k where $y_k < y_0$, and he almost always invests when $y_k > y_0$. The principal, knowing this, has no reason to pay him "more than a penny" above his "salary" y_0 in those states k where the principal wants him to invest (that is, when $x_k > 0$). The optimal contract for the principal, in the limit case when the agent's unit information cost tends to zero, has $y_0 = \bar{u}$. In this limit case, the agent is indifferent between accepting and rejecting the contract, and the optimal contract is continuous and piecewise linear, and it induces him to invest if and only if the net return rate is positive. Formally, the limit contract (as $c \to 0$) is[12]

$$y_k = \begin{cases} 0 & x_k < -\bar{u}/I \\ \bar{u} + Ix_k & -\bar{u}/I \le x_k \le 0 \\ \bar{u} & x_k > 0 \end{cases} \qquad \forall k \in \Omega. \tag{16}$$

Going back to the main case studied here, an agent with an arbitrary (and fixed) unit information cost $c > 0$, we finally note that the payment vector $\vec{y}$ is not in general proportional to I, the size of the investment. This may be surprising since the agent is risk neutral. However, the agent's information costs are non-linear, and it is easily verified from the proof of Proposition 2 that if the agent's participation constraint is slack, then each positive payment y_k increases with I at a decreasing rate:

$$\frac{\partial y_k}{\partial I} = \left(1 + \frac{\sigma}{c} e^{y_k/c}\right)^{-1} \cdot x_k.$$

4. A class of test projects

Having so far allowed for projects X with arbitrary finite support, we now apply the model to a class of projects with three potential outcomes, in order to shed light on the nature of optimal contracts and their dependence on the riskiness of projects. Three is the smallest number of outcomes that allows distinction between linear and non-linear contracts.

We consider a class of projects with mean-preserving spreads, allowing for a high but rare net return rate. More precisely, the three net return rates are

$$x_1 = -1, \quad x_2 = \frac{1}{2} \quad \text{and} \quad x_3 = \frac{1}{2} + \frac{1}{4\varepsilon}, \tag{17}$$

for $0 < \varepsilon < 1/2$, with prior probabilities $\mu_1 = 1/2$, $\mu_2 = 1/2 - \varepsilon$, and $\mu_3 = \varepsilon$. Hence, with probability one half all money is lost, with probability $1/2 - \varepsilon$ the net return rate is 50%, and with probability ε it is $1/2 + (4\varepsilon)^{-1}$. These projects all have zero expected net return rate; $\mathbb{E}[X] = 0$.[13] They are mean-preserving spreads of the actuarially fair double-or-nothing project in which there is equal probability of loosing all money ($x = -1$) or having it doubled ($x = 1$). That project is obtained when $\varepsilon = 1/2$. As ε decreases from 1/2 towards zero, the return rate x_3 tends to plus infinity. Hence,

[12] The case of costless information, $c = 0$, is not allowed for in our model. In that case, it is easily verified that there are infinitely many optimal contracts, one of which is the mentioned limit contract when $c \to 0$.

[13] The variance of X is decreasing in ε; $Var(X) = 7/8 + 1/(16\varepsilon) \approx 0.88 + 0.06\varepsilon^{-1}$.

for very small $\mu_3 = \varepsilon > 0$, state of nature 3 is a "black swan"—an extreme and very unlikely event.[14]

We have used numerical simulation methods to find optimal contracts for the following five projects in this family:

TABLE 1: The five projects, all with $\mathbb{E}[X] = 0$.

	x_1	x_2	x_3	μ_1	$\mu_2 = 1/2 - \varepsilon$	$\mu_3 = \varepsilon$	$Var(X)$
Project 1	-1	$1/2$	1	$1/2$	0	$1/2$	1
Project 2	-1	$1/2$	$3/2$	$1/2$	$1/4$	$1/4$	1.125
Project 3	-1	$1/2$	4	$1/2$	$3/7$	$1/14 \approx 0.071$	1.750
Project 4	-1	$1/2$	12	$1/2$	$11/23$	$1/46 \approx 0.022$	3.750
Project 5	-1	$1/2$	30	$1/2$	$29/59$	$1/118 \approx 0.009$	8.250

Numerically found optimal contracts with $I = 1$, for a risk neutral agent with $c = 0.05$ and reservation utility $\bar{u} = 0.1$ (and thus reservation wage $\bar{w} = 0.1$), are reported in the next table, with parameter estimates σ and τ for the formula in (15). We also report the expected profit to the principal, expected utility to the agent, and expected mutual-entropy reduction. (See Section 8 for optimal contracts for risk-averse agents.)

TABLE 2: Optimal contracts for a risk neutral agent.

project	y_0	y_1	y_2	y_3	Π	U	σ	τ	entropy red.
1	0.1101	0	–	0.2396	0.2493	0.1446	–	–	0.4019
2	0.1071	0	0.1898	0.2645	0.2552	0.1380	0.0511	0.1494	0.3719
3	0.1032	0	0.1891	0.3209	0.2646	0.1275	0.0468	0.1542	0.3291
4	0.1018	0	0.1889	0.3768	0.2699	0.1217	0.0453	0.1545	0.3106
5	0.1015	0	0.1890	0.4268	0.2721	0.1197	0.0441	0.1586	0.3052

We note that the salary (y_0) varies only slightly across the projects. While the remuneration (y_2) after investment in the state with the intermediate net return rate ($x_2 = 1/2$) is virtually the same in Projects 2-5, the remuneration (y_3) after investment in the most favorable state of nature increases with its net return rate (x_3).[15]

[14] However, unlike "true" black swans, the principal and agent are here aware of the possibility of a black swan.

[15] The prior probability distribution μ varies across the five projects, and so does the size of the third outcome. Therefore, the invariance result in Proposition 2 in Caplin, Dean and Leahy (2019), for variations in a decision-maker's prior (only), does not apply directly. However, there may be some connection with that result, a topic for future research.

The actuarially fair double-or-nothing Project 1 is a priori symmetric. Nevertheless, the risk-neutral agent is rewarded asymmetrically; he is paid $y_0 - y_1 \approx 0.1101$ for not investing in the bad state of nature, and $y_3 - y_0 \approx 0.1295$ for investing in the good state. He is thus paid less for making the right decision in one state than in the other, although the two states are a priori equally likely. This illustrates the tension between (a) incentivizing the agent to acquire information, and (b) incentivizing the agent to make the right decision once his information has been obtained. Taken in isolation and without regard to the limited-liability constraint, the second motive requires full alignment with the principal's interest at that stage, that is, equal rewards to the agent for "doing the right thing" in both states of nature.[16] The reason why the agent is paid less for doing the right thing in the bad state of nature $(y_0 - y_1)$ is that the limited-liability constraint forces any raise of that pay to occur by way of raising y_0, which, however, simultaneously reduces the reward $(y_3 - y_0)$ for making the right thing in the good state of nature. In order to incentivize the agent to acquire information, the principal needs to reward the agent sufficiently to make information acquisition worth-while, and also reward for taking the right action. However, the limited liability constraint induces an asymmetry in the contract costs for the principal.

We also note that the principal's expected profit and the agent's expected utility vary across projects, with the profit increasing and utility decreasing as the highest return rate increases and its likelihood decreases, see Figures 2 and 3 below, drawn for a large variety of projects in this class.

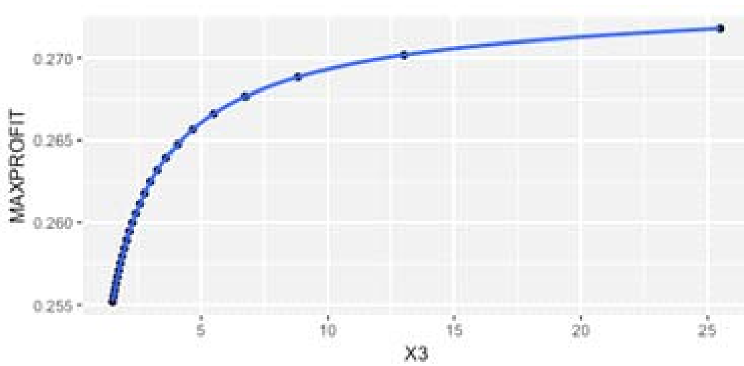


Figure 2: The principal's expected profit.

[16]Indeed, in the absence of information costs and decision costs for the agent, it would be sufficient to pay the agent "one penny" for making the right decision in each state of nature. This would fully align the agent's incentives with the interest of the principal.

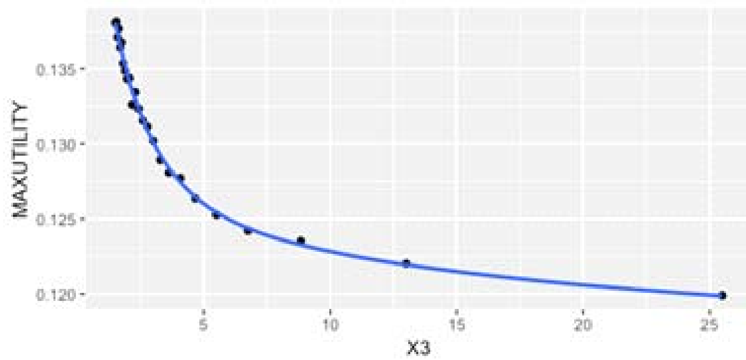


Figure 3: The agent's achieved utility under optimal contracts for different projects.

Our numerical simulations also show that the agent makes less effort to acquire information the smaller ε is; the reduction that he achieves in the entropy of his endogenous uncertainty falls from approximately 0.405 to approximately 0.305 as we move from Project 1 to Project 5. Hence, for this class of projects, a mean-preserving spread that gradually permits a more valuable but rarer black swan it is optimal for the principal to somewhat weaken the agent's incentive to acquire information about the state of nature the more valuable and rarer the black swan is. An intuitive explanation for this result is that the principal wants the agent to invest in states 2 and 3, and not in state 1. The prior probability for the latter state is, by construction, the same as the prior probability for the event that the state is either 2 or 3, in all projects in the present family. Conditional upon the state being either 2 or 3, a project's expectation is independent of ε: $\mathbb{E}[X \mid X \geq 0] = 1$. It appears that, as ε decreases and the net return x_3 in state 3 accordingly increases, the principal finds it optimal to have the agent reduce uncertainty less.

The table also shows that the agent's expected utility is above his participation constraint, $\bar{u} = 0.10$. It is in the principal's self-interest to pay the agent more. The reason is that the principal wishes to incentivize the agent to acquire information, and for this there needs to be a penalty for investing in the bad state, which, in turn, requires that the salary is not too low. As was argued above, the agent's incentives, once he has accepted a contract, depends entirely on the transfers t_k (bonuses and penalties) after investment, and not at all on the salary. However, the limited-liability constraint puts an upper bound on penalties, the strongest feasible penalty being to withhold the whole salary. Hence, it may be in the interest of the principal to offer a high salary.

See Figure 4 below for an illustration of the optimal contract for Projects 3, where the dashed straight line is the pay after non-investment, y_0, and the solid curve is the pay after investment, $y_k = \max\{0, y_0 + \phi(x_k)\}$, where ϕ is the optimal transfer function in Proposition 2.[17] We note that bonuses (penalties) are given for investments with net return rates above (below) $\sigma + \tau \approx 0.201$, and that the agent is paid nothing if investing when the net return rate is below approximately 0.057.

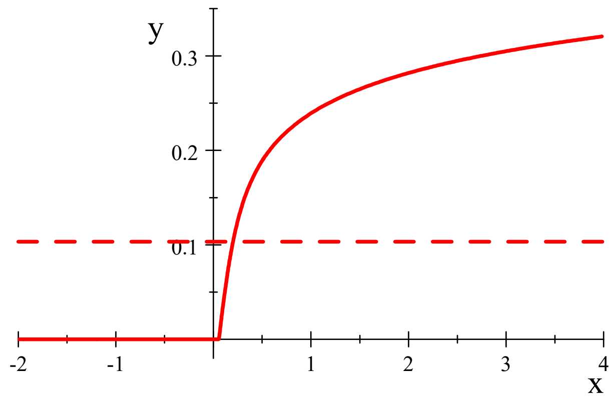


Figure 4: The optimal contract for Project 3, for a risk-neutral agent with $c = 0.05$ and $\bar{u} = 0.1$.

Numerical estimates of the parameters σ and τ in Proposition 1, for each of the five projects, yield σ-values falling (approximately) from 0.051 to 0.044 as ε falls from 1/4 to 1/118, and τ-values approximately rising from 0.150 to 0.159. Hence, these parameters are quite stable across the different projects, with approximate values $\sigma \approx 0.05$ and $\tau \approx 0.155$ (and thus with graphs close to that in Figure 4).[18]

The agent's optimally chosen uncertainty reduction, about the state of nature, rises somewhat, from approximately 0.40 to about 0.30. Hence, the rarer the black swan, the less effort does the agent make. However, being aware of the possibility of a black swan, the principal's contract changes according to its nature. Had the true project been project 5, say, and had the principal been unaware of its black swan, then she would have proposed the optimal contract for project 1, with its higher salary and slightly higher bonus for investment in the intermediate state 2.

We finally note that, when expressed as a share of the investment's net return, the agent receives much less from the best outcome than from the intermediate outcome: The pay share y_2/x_2 is approximately 38% for all projects in the table, while the pay share y_3/x_3 falls from about 24% to about 1.4% in Project 5. Hence, the optimal contracts are quite non-linear. Indeed, they are strongly regressive. As a consequence,

[17] While there are only three possible net return rates, the optimal transfer function ϕ is defined for all net return rates $x \in \mathbb{R}$.

[18] Since the agent's investment probability q is approxmimately one half, we are not surprised to find that σ is close to $c = 0.05$.

contracts based on stock shares or options, being (piecewise) linear, are suboptimal. Their disadvantage for the principal is that they pay the agent unnecessarily much for catching the black swan. While roughly 40% of the net return from investment goes to the agent in the intermediate outcome, only about 1.5% of a "black swan's" net return are paid to the agent. This makes a huge difference when compared with conventional stock- or option contracts.[19]

The next table shows how the optimal contract, expected profit and utility, depend on the agent's unit information cost, c, for Project 3.

TABLE 3: Optimal contracts for agents with different information costs.

c	y_0	y_2	y_3	Π	U
0.01	0.0386	0.0764	0.0982	0.4259	0.0523
0.05	0.1032	0.1891	0.3209	0.2646	0.1275
0.10	0.0415	0	0.5167	0.1665	0.0577
0.15	0.0487	0	0.6503	0.1361	0.0674
0.25	0.0570	0	0.8065	0.0916	0.0764
0.50	0.0564	0	0.8073	0.0303	0.0674

Not surprisingly, the principal's profit is falling as the agent's information cost rises. (The agent's participation constraint is assumed to be slack.) Less evident is the dependency of the agent's expected utility upon his information cost. From first principles (Berge's maximum theorem) we know that the principal's expected profit under optimal contracting, Π, is continuous with respect to the agent's unit cost of information, c. However, the optimal contract need not be continuous with respect to variations in the agent's unit information cost, c. Indeed, as the latter crosses over the value 0.1, the contracts change drastically. For agents with higher costs, the principal pays nothing to the agent if he invests in the intermediate state of nature. Instead, the principal's optimal contract rewards the agent significantly more when he invests in the best state of nature (more the worse the agent is). Such discontinuity may cause the agent's expected utility U, when acting optimally under a contract that is optimal for the principal, to be discontinuous with respect to the agent's unit cost. And this seems indeed to be the case.

See the two diagrams below, showing numerically how the principal's expected profit (at optimum) declines continuously with the agent's unit information cost, while the agent's expected utility (when acting optimally under an optimal contract) has two local maxima, and an possible discontinuity at $c \approx 0.969$ (see Appendix for

[19] Numerical simulations of even more extreme projects in the same class result in similar optimal contracts. For example, if $x_3 \approx 25000$ and $\varepsilon \approx 0.00001$, then $y_0 \approx 0.1017$, $y_2 \approx 0.1896$, $y_3 \approx 0.8506$, $\Pi^* \approx 0.2741$, and $U^* \approx 0.1181$. Hence, $y_3/x_3 \approx 3.4 \cdot 10^{-5}$, a remarkably low share.

details).

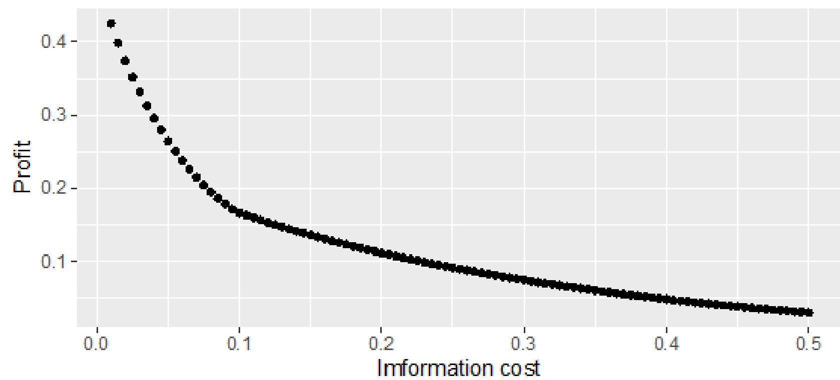


Figure 5: The principal's optimal profit as a function of the agent's unit information cost.

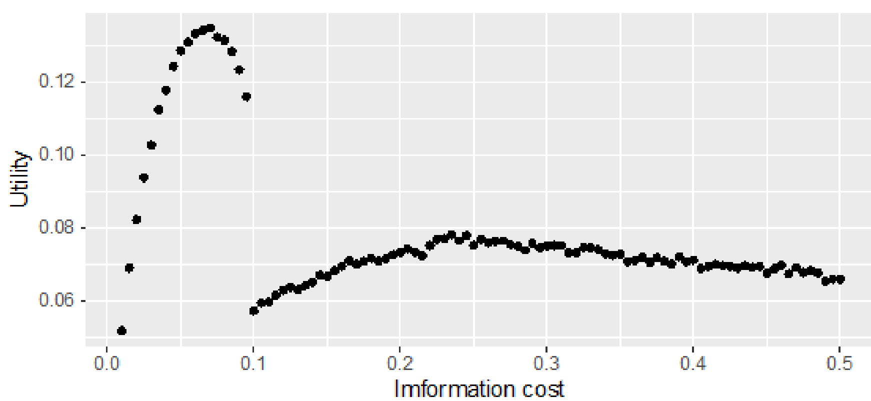


Figure 6: The agent's expected utility as a function of his unit information cost.

Such a continuity may be due to the lack of quasi-concavity of the principal's expected profit with respect to the contract. It may well be that there are two local maxima of this function. For low unit costs (those below 0.1), the local maximum has y_2 positive and sizeable; it is worthwhile for the principal to reward very able agents (those with low information costs) also for investing in the intermediate state of nature. It is as if that will not distract their attention from the possibility of the very good state of nature. The other locally optimal contract has $y_2 = 0$ and results in lower expected profit to the principal for relatively unable agents, those with high unit costs of information. At around $c = 0.969$, the two local maxima take the same value, and for higher unit information costs, it is the local optimum with $y_2 = 0$ that becomes the global maximizer for the principal. But such a discontinuous jump from one type of contract to another hits the agent discontinuously.[20]

5. Profit-share and stock-based contracts

We here compare the optimal contracts in Proposition 2 with contracts much used in practice and discussed in the literature (see, e.g. Swedroe, 2014). To this end, we here calculate the principal's profit when restricting herself to either profit-share contracts or to stock-based contracts (still under the limited-liability constraint). We also briefly comment on options-based contracts.

5.1. Profit-share contracts. With Ix_k representing the gross profit from investing in state of nature $k \in \Omega$, by a profit-share contract we mean a contract $\vec{y} \in \mathbb{R}_+^{m+1}$ of the form

$$y_k = \max\{0, w_1 + a_1 I x_k\} \quad \text{for} \quad k = 0, 1, ..., m,$$

where $w_1 \geq 0$ is a fixed wage and $a_1 \in (0,1)$ the fixed profit share given to the agent.[21] We here set $x_0 = 0$; this is the net return rate after non-investment. In other words, the pay to the agent is w_1 after non-investment, it is zero after investment in states $k \in \Omega$ with $a_1 I x_k \leq -w_1$. After investment in all other states of nature the pay to the agent is $w_1 + a_1 I x_k$.

Choosing w_1 and a_1 optimally (among all profit-sharing contracts) for each of Projects 1-5, for a risk-neutral agent with reservation utility $\bar{u} = 0.1$, results in contracts, expected profits and utilities shown in Table 4 below. We see that the profit share is between 10 and 13 percents in all projects (varying non-monotonically with the riskiness of the project).

TABLE 4: Optimal profit-share contracts.

project	a_1	w_1	Π	U	entropy red.
1	0.1293	0.1101	0.2493	0.1446	0.4019
2	0.1159	0.1022	0.2530	0.1324	0.3451
3	0.1002	0.0841	0.2466	0.1109	0.2219
4	0.1186	0.0847	0.2398	0.1202	0.2064
5	0.1232	0.0849	0.2381	0.1227	0.2032

[20] Think of the two bumps on the back of a camel, where one is higher than the other, depending on the slope of the ground on which the camel stands. A rider who wants to sit on the highest bump will move from one bump to the other. This results in a continuous altitude effect for the rider, but quite a change in burden for the camel.

[21] We thus here model the agent's pay as a share of the principal's gross profit. The net profit to the principal after investment in state k is $(x_k - y_k) I$, and after non-investment it is $-y_0 I$.

When comparing with Table 2, we see that, except for Project 1, the principal's expected profits are reduced when restricted to profit-share contracts. The reason why Project 1 is special is that it has only two outcomes with positive probability, and is thus insensitive to the linearity constraint inherent in profit-share contracts. The difference in expected profits is particularly stark for the most risky Project 5, for which the principal's profit loss from using profit-share contracts is about 12.5%. The reason is that under profit-sharing the agent is paid far too much for "catching the black swan". Despite this big prize, the agent makes less work efforts, and obtains less information under profit-share contracts, in all of Projects 2-5, as can be seen by comparing the entropy reductions in Tables 2 and 4.

5.2. Stock-based contracts. Having investigated profit-share contracts, we next consider stock-based contracts. Taking the stock value after investment in state of nature k to be $V_k = \max\{0, (1 + x_k) I\}$, and after non-investment to be $V_0 = I$, a stock-based contract is a contract $\vec{y} \in \mathbb{R}_+^{m+1}$ of the form

$$y_k = w_2 + a_2 V_k \quad \text{for} \quad k = 0, 1, ..., m,$$

where $w_2 \geq 0$ is a fixed wage and $a_2 \in (0, 1)$ the stock share given to the agent.[22] In other words, under a stock-based contract, the pay to the agent is $y_0 = w_2 + a_2 I$ after non-investment, it is $y_k = w_2$ after investment that results in a loss exceeding the money invested (that is, $x_k < -1$, if such outcomes exist). After investment in all other states of nature the pay to the agent is $y_k = w_2 + a_2 I (1 + x_k) = y_0 + a_2 I x_k$.

Choosing w_2 and a_2 optimally (among all stock-based contracts) for each of Projects 1-5, for the same agent as in Tables 2 and 4, results in contracts, expected profits and utilities as shown in Table 5 below.

TABLE 5: Optimal stock-based contracts.

project	a_2	w_2	Π	U	entropy red.
1	0.1171	0	0.2469	0.1456	0.3960
2	0.1070	0	0.2516	0.1330	0.3421
3	0.0830	0	0.2449	0.1027	0.2007
4	0.0906	0	0.2291	0.1145	0.1598
5	0.1012	0	0.2234	0.1292	0.1696

We note that in optimal stock-based contracts, the agent receives no salary, but is given between 9% and 12% of the stock value. The principal's expected profits are uniformly lower under these contract than under profit-share contracts, which in turn yield uniformly lower profits that the globally optimal contracts in Proposition 2.

[22] Like in the case of profit-based contracts, we here neglect the potential effect of the pay to the agent on the stock value. In addition, we disregard potential equilibrium effects on stock values when hiring agents.

Remark 1. *Some's managers are offered a fixed salary and the option to buy stocks at a future time but at today's price. In order to briefly consider such contracts within the present model, let V_0 denote the stock value at the time when the contract is signed, and V_k the future stock value in state of nature k. A (call) option-based contract can then be written as a payment vector $\vec{y} \in \mathbb{R}_+^{m+1}$ with*

$$y_k = w_3 + a_3 \cdot \max\{0, V_k - V_0\} \quad \text{for} \quad k = 0, 1, ..., m,$$

where $w_3 \geq 0$ is a fixed wage and $a_3 \in (0,1)$ is the stock-share that the contract offers the option to buy. Since call options only have an upside value, ex post, these contracts are clearly suboptimal, since under such a contract it is a dominant strategy for the agent to invest blindly.[23]

6. Bargaining

The analysis above concerns situations in which the principal makes a take-it-or-leave-it contract proposal to the agent, well aware of both parties' outside option. For the agent, the expected utility of his outside option is $\bar{u}$. For the principal it could be to make the investment decision herself, abstain from investing, or to (try to) hire another agent. We here consider the case when the principal's outside option has value zero. For the sake of brevity, we here focus on the case of a risk-neutral agent.

Imagine that the two parties bargain over the terms of the contract (that is, the payment vector $\vec{y}$, or, equivalently over the salary, bonuses and penalties). Let $S \subset \mathbb{R}^2$ be the set of feasible utility-profit pairs, that is, pairs $(U(\vec{y}), \Pi(\vec{y}))$ associated with any contract $\vec{y}$ that meets the limited-liability condition $\vec{y} \geq 0$. The Nash bargaining solution would be the contract that solves the following maximization program:

$$\max_{\vec{y} \geq 0} \quad [U(\vec{y}) - \bar{u}]^{\beta} \cdot [\Pi(\vec{y})]^{1-\beta} \tag{18}$$

where $\beta \in [0,1]$ is the agent's bargaining power and $1-\beta$ the principal's bargaining power. Here $\bar{u}$ is the agent's reservation utility, and the principal's reservation profit is set at zero. Giving the principal the right to make a take-it-or-leave-it offer amounts to giving the agent no bargaining power, $\beta = 0$. The situation is illustrated in the diagram below, drawn for Project 3 in the preceding section, for a risk-neutral agent with $c = 0.05$ and $\bar{u} = 0.10$.

[23] In the terminology used in Section 2, these contracts makes his investment climate "sweet".

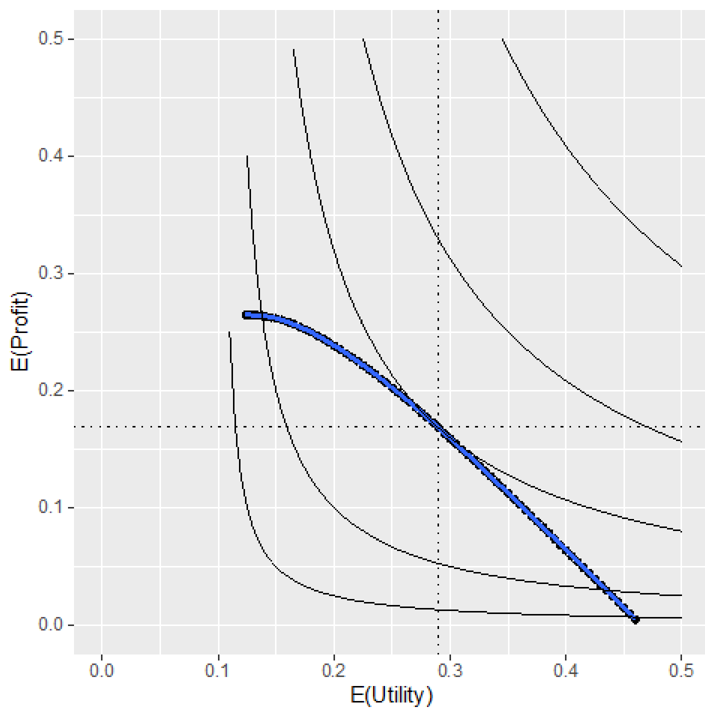


Figure 7: The Nash bargaining solution.

The cluster of (blue) points (obtained by numerical simulation) represent the Pareto frontier of the feasible set S, and the family of hyperbolae constitute a contour map for the Nash product when both parties have equal bargaining power ($\beta = 0.5$). The Nash bargaining solution is then the tangency point, indicated with a dotted horizontal and vertical line. That point is approximately $\hat{\Pi}^{(0.5))} \approx 0.1690$ and $\hat{U}^{(0.5)} \approx 0.2891$, to be compared with the outcome when the principal has all the bargaining power, $\Pi^* \approx 0.2646$ and $U^* \approx 0.1275$. The latter pair, obtained in Section 3.1, here corresponds to the top of the Pareto frontier in the diagram. Not surprisingly, the principal's expected profit is lower when the agent has bargaining power.

It is also interesting to compare the negotiated contract, when $\beta = 0.5$ with that when the principal has all bargaining power. From Section 3.1 we know that in the latter case the contract is approximately $\vec{y}^* = (0.1032, 0, 0.1891, 0.3209)$. By contrast, under equal bargaining power, the equilibrium contract is $\vec{y}^{(0.5)} = (0.2143, 0, 0.4134, 0.5410)$, that is, roughly doubled salary, more than double the pay after investment in the intermediate state of nature, and less than double the pay after investment in the best state of nature. In terms of bonus rates, defined as $b_k = (y_k - y_0)/y_0$, we have, expressed as percentages $b_2^* \approx 83\%$, $b_3^* \approx 211\%$, and $b_2^{(0.5)} \approx 93\%$ and $b_3^{(0.5)} \approx 152\%$.

In the above calculations, we have here fixed the agent's reservation utility (at $\bar{u} = 0.10$). However, this utility is arguably higher the more competition there is for talented agents. The effect of increased competition for agents with a given level of talent, here unit information cost c, may enter the bargaining through two channels; it may increase the agent's bargaining power β and it may also increase his reservation utility. Analysis of the effects on the bargaining outcome is readily obtained from (18). All contracts on the Pareto frontier has maximal punishment after investment in state 1 ($y_1 = 0$).

TABLE 6: Contracts for Project 3 as Nash bargaining solutions.

β	y_0	y_2	y_3	Π	U
0	0.1032	0.1891	0.3209	0.2646	0.1275
0.25	0.1621	0.3104	0.4352	0.2306	0.2126
0.50	0.2143	0.4134	0.5410	0.1690	0.2891
0.75	0.2729	0.5259	0.6600	0.0887	0.3746
1	0.3351	0.6412	0.7985	0	0.4648

We note that the more bargaining power the agent has, the higher is his salary and are his bonuses. The two parties always agree that the penalty for investing in the bad state should be maximal, even when the agent has all the bargaining power. In that extreme case, the principal's participation constraint, to earn a nonnegative profit, is binding, and so the agent has a self-interest in efficiency.

It is remarkable that if the agent has a lot of bargaining power, say $\beta = 0.75$, then his contract contains a high salary, 0.2729 (to be compared with the size of the investment, $I = 1$), and he is paid above the net return if he invests in the intermediate state of nature; he receives $y_2 = 0.6412$ while the net return rate is $x_2 = 0.5$. We saw in Section 3, that such lavish bonuses are never paid (in the class of projects we study) when the principal has all the bargaining power; for Project 3, that we here analyze, we see that $y_2 \approx 0.1891$ when $\beta = 0$. The present model can thus, in principle, be used to infer from contracts whether or not the agent has had significant bargaining power.

7. Risk-averse agent

We here briefly study the case of a risk averse agent. The analytically most elegant case in the present model is logarithmic utility of income. Indeed, this also has some empirical support, and was suggested already by Daniel Bernoulli in 1738. In this case, when the agent's utility is logarithmic, we require all payments to the agent to be strictly positive.

It follows from Lemma 1 that investment conditions under a contract $\vec{y} > 0$ are normal for an agent with logarithmic utility if and only if

$$\frac{1}{\mathbb{E}\left[Y^{-1/c}\right]} < y_0^{1/c} < \mathbb{E}\left[Y^{1/c}\right]. \tag{19}$$

Under normal investment conditions for the agent, the probability that he will invest, q^*, is the unique solution in $(0,1)$ to the equation

$$\sum_{k\in\Omega} \frac{\left(y_k/y_0\right)^{1/c} \mu_k}{1-q+\left(y_k/y_0\right)^{1/c} q} \;=\; 1 \tag{20}$$

This equation shows how the conditional investment probabilities, given the true state of nature $k \in \Omega$, are tilted in favor of those where the agent's pay is above his salary, and tilted away from those with pay below his salary. Indeed, all investment probabilities are functions of the vector of payment ratios, $r_k = y_k/y_0$. Write $\vec{r}$ for the vector $(r_1, .., r_m)$. Thus, while for a risk-neutral agent only the payment *differences*, the net transfers $t_k = y_k - y_0$, matter for his information acquisition and investment decision, here only the payment *ratios* matter.

Under any contract $\vec{y} > 0$ that provides normal investment conditions for the agent, the agent's expected utility is

$$U(\vec{y}) \;=\; \sum_{k\in\Omega} \frac{\mu_k q^* r_k^{1/c}}{1-q^*+q^* r_k^{1/c}} \ln r_k \;-\; c\cdot\left(H(q^*) - \sum_{k\in\Omega}\mu_k H(p_k^*)\right) \;+\; \ln y_0,$$

and the principal's expected profit is

$$\Pi(\vec{y}) \quad=\quad \sum_{k\in\Omega} \frac{\mu_k q^* r_k^{1/c}}{1-q^*+q^* r_k^{1/c}} \left(I x_k - r_k y_0\right) \quad-\quad (1-q^*)\, y_0.$$

We see that, given all payment ratios (which uniquely determine the agent's *ex ante* investment probability q^*), the agent's expected utility is (logarithmically) increasing in his salary, y_0, while the principal's expected profit is (linearly) strictly decreasing in the same salary. Hence, for any given reservation utility $\bar{u} \in \mathbb{R}$ that the agent may have, the optimal contract for the principal can be obtained by (a) first finding an optimal payment-ratio vector $\vec{r}$, for any fixed and given salary $y_0 > 0$, and (b) then adjusting this salary, without changing the payment ratios, so that the agent's participation constraint is met with equality. The second step, (b) amounts to setting the salary, for any given vector $\vec{r}$ and reservation utility $\bar{u}$, such that

$$y_0 \;=\; \exp\left[\bar{u} \;-\; \sum_{k\in\Omega} \frac{\mu_k q^* r_k^{1/c}}{1-q^*+q^* r_k^{1/c}} \ln r_k \;+\; c\cdot\left(H(q^*) - \sum_{k\in\Omega}\mu_k H(p_k^*)\right)\right] \tag{21}$$

Applied to the class of projects with mean-preserving spreads in Section 3.2, and with an agent with the same unit information cost ($c = 0.05$) and reservation wage ($\bar{w} = 0.1$) as there, we now have $\bar{u} = \ln 0.1$. The simulation results are given in Table 7 below (c.f. the optimal contracts for a risk-neutral agent given in Table 2). Perhaps surprisingly, we note that the principal obtains higher profits than when meeting the risk-neutral agent. Under the optimal contract for a risk averse agent, it is sufficient to pay a somewhat lower salary than to the risk neutral agent, and also the payments to the risk averse agent after profitable investments are lower. The risk averse agent values zero pay at utility minus infinity. Under the optimal contract, the agent is given a low but positive pay after investment in the bad state of nature. This pay is only a small fraction of his salary. Evidently the threat of this low pay incentivizes the agent to become well-informed, which is seen in the last column in the table, showing significantly larger entropy reduction than for the risk neutral agent.

TABLE 7: Optimal contracts for logarithmically risk-averse agent with reservation utility $\bar{u} = \ln(0.1)$.

Project	y_0	y_1	y_2	y_3	Π	entropy red.
1	0.0869	0.0386	–	0.1233	0.3944	0.6894
2	0.0873	0.0410	0.1202	0.1254	0.3945	0.6886
3	0.0878	0.0130	0.1205	0.1313	0.3946	0.6873
4	0.0880	0.0425	0.1211	0.1373	0.3947	0.6872
5	0.0881	0.0141	0.1212	0.1439	0.3947	0.6870

We conjectured before that perhaps optimal contracts for risk-averse agents are less regressive than those for risk-neutral agents. We now have the numbers. Like in the case of a risk-neutral agent, the pay share y_2/x_2 to a risk averse agent is also approximately constant across all projects. However, the level, approximately 0.24, is lower. The pay share y_3/x_3 for the risk-averse agent falls from about 0.12 to about 0.0047. Hence, again quite a regressive contract, so our conjecture did not hold up; also optimal contracts for risk averse agents are highly regressive.

To elicit the robustness of these observations we briefly consider agents with CRRA Bernoulli functions, $u(y) = \left(y^{1-\rho} - 1\right) / (1 - \rho)$, for various coefficients of relative risk-aversion ρ, and all with the same reservation wage as in Section 3, and hence $\bar{u} = u(0.1)$. See table below, with results for Project 3, where $\rho = 0$ is the case of a risk-neutral agent, and $\rho = 1$ is the case of a logarithmically risk-averse agent.[24]

[24] As is well-known, the CRRA formula gives logarithmic utility in the limit as $\rho \to 1$. For $\rho = 0$, we obtain $u(y) = y - 1$ and hence $\bar{u} = u(0) = 0.1 - 1 = -0.9$, for comparison with the results in Section 3.

TABLE 8: Optimal contracts for Project 3 for CRRA agents.

ρ	y_0	y_1	y_2	y_3	Π
0	0.1032	0	0.1891	0.3209	0.2646
0.25	0.0695	0	0.1655	0.2315	0.3580
0.50	0.0722	0	0.1537	0.1897	0.3828
0.75	0.0816	0.0071	0.1336	0.1526	0.3901
1	0.0878	0.0130	0.1205	0.1313	0.3946

The earlier observations for the logarithmic case appear robust. We note, however, that the pay after non-investment, y_0, is non-monotonic with respect to the coefficient of relative risk-aversion, and that not so risk-averse agents ($\rho \leq 0.5$) are paid nothing if investing in the worst state of nature. The more risk-averse the agent, the lower is the bonus for the best outcome, thus deviating even further from linear contacts. Hence, the very high bonuses we some's see in practice cannot be explained within this model by allusion to agents' risk aversion (the need for very high pay for success, to compensate for the agent's declining marginal utility of income).

Remark 2. *In realistic settings, the pay from the contract is only part of the agent's life-' income. Hence, if the agent is risk-averse and there is a well-functioning credit market, presumably the agent can smooth his consumption over '. Effectively, this will amount to rendering the agent less risk-averse. This topic, although of great practical importance, will not be addressed here. For an analysis of dynamic contracting with risk-averse agents, see e.g. Bergemann and Pavan (2015).*

8. Contracts with a firing clause

We have, so far, allowed contracts to pay less after poor investments than after non-investment. However, real-life contracts for CEOs typically do not include such pecuniary penalties. Instead, they often allow the principal to fire the CEO with short notice. The threat of being fired may enhance the agent's incentive on the job, since, arguably, an agent who has been fired in general will face worse future career prospects than an agent who voluntarily leaves his or her position for another job. In other words, being fired may be a penalty for the agent. In practice, however, the agent is given a severance payment, or "golden parachute" when fired. We here sketch a modified version of the above model in which the principal cannot pay the agent less than the non-investment pay (the "salary").

Consider, thus, an (risk neutral or risk averse) agent for whom being fired is equivalent to a utility loss of Δ. This utility loss is larger the more patient the agent is and the larger is the reputational damage of being fired, for his future career prospects. For brevity we call Δ the agent's *firing cost*. The difference from the above sketched career concerns is two-fold: first, firing is usually more visible than the realized returns from investment. Second, firing is an active decision by the principal.

We proceed by replacing the contracts $\vec{y} \in \mathbb{R}_+^{m+1}$ studied so far by contracts without penalties but instead with a firing clause that takes the following form: (i) the agent is fired if he has invested and the realized net return rate is below a pre-specified threshold $\xi \in \mathbb{R}$, and (ii) he is given a pre-specified severance payment $v \geq 0$ if fired. Hence, a feasible contract is now a triplet $\langle \vec{y}, \xi, v \rangle$, under which the agent is paid $y_0 \geq 0$ if not investing, $y_k \geq y_0$ if investing in a state k with $x_k \geq \xi$, and he is fired and paid severance $v \geq 0$ if investing in a state k with $x_k < \xi$. If the agent, in addition, has career concerns of the type mentioned above, then his utility after non-investment is $u(y_0) + \gamma_0$, it is $u(y_k) + \gamma_k$ after investment in a state k with $x_k \geq \xi$, and it is $u(v) + \gamma_k - \Delta$ after investment in a state k with $x_k < \xi$.

We briefly analyze this setting under the simplifying assumptions (a) that the agent's career is not influenced by the realized returns from investment apart from the possibility of being fired (that is, $\vec{\gamma} = 0$), and (b) that there is no cost for the principal of firing the agent, and hence no commitment issue if and when firing is called for according to the contract.

Suppose, first, that the agent is risk neutral. The analysis then follows the same lines as in Section 3, *mutatis mutandis*. The equivalent of (4) writes as before, but with $y_k = v \geq 0$ if $x_k < \xi$ and $y_k \geq y_0$ if $x_k \geq \xi$, and the same is true for the agent's expected utility. We illustrate the nature of optimal contracts with a firing clause by comparing them with the purely financial contracts for the class of projects in Section 4. In particular, we investigate if and when the principal can do better by offering a contract under which penalties are replaced by firing. We first consider a risk-neutral agent with unit information cost $c = 0.05$ and reservation utility $\bar{u} = 0.1$, just as in Table 2. The only difference is that now $\Delta = 0.2$.[25] It is then optimal for the principal to fire the agent if and only if the agent invests in state of nature 1. The next table presents the optimal contracts for all five projects in Section 4.

TABLE 9: Optimal contracts with a firing clause, for a risk neutral agent with reservation utility $\bar{u} = 0.1$.

project	y_0	v	y_2	y_3	Π
1	0.0160	0	–	0.2522	0.3560
2	0.0192	0	0.2122	0.2855	0.3569
3	0.0247	0	0.2250	0.3532	0.3580
4	0.0284	0	0.2314	0.4191	0.3585
5	0.0298	0	0.2342	0.4663	0.3588

[25] If one views Δ as the present value of a reduction by one half of the future reservation wage, paid perpetually, with per project-period discount factor $\delta = 4/5$, then $\Delta = (0.1 - 0.05)\,\delta/\cdot(1-\delta) = 0.2$.

The optimal contracts contain a positive, but low,salary, provides larger bonuses after investments with positive returns than in the absence of the firing clause, and result in higher expected profits to the principal. The expected utility to the agent is lower than in contracts without a firing clause. The agent being risk-neutral, we are not surprised to find that the agent receives no severance payment after being fired. We next briefly study optimal contracts with a firing clause for a risk-averse agent, the same as in Section 7, with logarithmic utility from income and reservation utility $\bar{u} = \ln(0.1)$. We focus on Project 3, and vary the the firing cost for the agent.[26]

TABLE 10: Optimal contracts for a logarithmically risk-averse agent with different firing costs.

Δ	y_0	v	y_2	y_3	Π
0	0.0876	0.0165	0.1206	0.1305	0.3946
0.7	0.0878	0.0369	0.1206	0.1311	–”–
1.4	0.0878	0.0784	0.1207	0.1309	–”–
2.1	0.0878	0.1589	0.1206	0.1313	–”–
2.8	0.0877	0.3232	0.1208	0.1315	–”–
3.5	0.0877	0.6128	0.1207	0.1311	–”–

We see that optimal contracts do contain sizeable severance payments after firing, "golden parachutes". While these payments increase monotonically with the agent's firing cost, the other parts of the contract are remarkably constant, as is the principal's expected profit. It thus seems that it is enough for the principal needs to mitigate the agent's loss of future income in case he makes a bad investment and is fired. Under these contracts, the probability that the agent will invest in the bad state of nature is extremely low (far below 10^{-6}) so variations in the severance payment v has a very small effect on the principal's expected profit and the agent's expected utility (while keeping the agent's utility at his participation constraint). One and the same contract, for example $y_0 = 0.877$, $v = 0.05$, $y_2 = 0.121$ and $y_3 = 0.131$, will work just as well, from a practical viewpoint, as the optimally fine-tuned contracts, independently of the size of the agent's utility loss Δ from being fired. We also see that the principal's expected profit is the same as when the contract does not contain any firing clause at all (see Table 7, and compare also with the case $\Delta = 0$ in Table 10). We finally note that the principal makes a slightly higher expected profit when hiring the risk-averse agent than when hiring the risk-neutral agent; approximately 39.5% versus 35.8%.

[26] Again viewing Δ as the present value of a reduction by one half of the future reservation wage, paid perpetually, one obtains for $\delta = 4/5$: $\Delta = (\ln 0.1 - \ln 0.05)\,\delta/(1-\delta) \approx 2.77$.

9. Related literature

The papers in the literature that, to the best of our knowledge, are closest to ours are, in chronological order: Lambert (1986), Lewis and Sappington (1997), Levitt and Snyder (1997), Aghion and Tirole (1997), Crémer, Khalil and Rochet (1998), Malcomson (2009), Zermeño (2011, 2012), Terovitis (2018), Yang and Zeng (2019), and Carroll (2019).

The approach in Lambert (1986) differs from ours in that the agent's information acquisition decision and return distribution are binary. Conclusions about contract forms for more general situations are not possible to obtain. By contrast, in our model the agent faces a continuum of information-acquisition choices, and the asset return distribution has arbitrary finite support, which permits us to identify functional forms for contracts. Lewis and Sappington (1997) also analyze situations in which the agent has a discrete choice; at a given cost, he can choose to be perfectly informed about the state of nature. This is also true of Crémer, Khalil and Rochet (1998).

In the framework of Aghion and Tirole (1997), our model concerns the case of what they call A-formal authority ("agent-formal" authority). However, the setting differs from theirs in a few dimensions. First, while they consider a finite number of ex ante identical project, we here only consider one project.[27] Second, they only consider projects with binary outcome distributions, while we here allow for projects with an arbitrary finite number of outcomes. Third, in their model, the agent is either completely uninformed or completely informed about the true return rates of projects, while our agent is either completely uniformed or arbitrarily but boundedly well-informed on a continuum scale. The latter feature allows us to obtain functional forms for optimal contracts, whereby the dependence of bonuses on realized returns becomes transparent.

Levitt and Snyder (1997) analyze the optimal design of incentive schemes when the agent not only has private information about his own work effort, but also has a private signal (of given precision) about the state of nature. In that model, the agent chooses his effort level, which can be either high or low, and this in turn determines the success probability of the project at hand. The agent then receives a private signal, of given precision, about the state of nature, and makes an announcement about the received signal to the principal, who makes the investment decision.[28] By contrast, in our model the agent's effort, which is not binary but a continuous variable, does not affect the success probability of the project at hand. Moreover, in our model it is the agent who makes the investment decision on behalf of the principal.

Malcomson (2009) considers a principal-agent model in which the agent has to take

[27] The present approach easily generalizes to a finite number of project. However, this comes at a relatively high price in terms of notation and algebra without, we feel, adding much insight.

[28] See also Friebel and Raith (2004).

an action, such as information acquisition, and make a decision, such as investment in a risky asset. The sequencing is as in our model: the principal first makes a take-it-or-leave it contract offer, the agent accepts or rejects this offer, and, in the case of acceptance, takes his action, receives a signal that provides information about the return distributions to the possible decisions, and then makes his decision, after which the return to the decision materializes and payments are made. The principal is risk neutral and the agent risk averse. Like in our model, the agent's action and signal are unobservable to the principal, and the materialized returns to the agent's action are verifiable. However, unlike in our model, the agent's decision is taken to be non-contractible, and there no limited-liability constraint on contracts. While the first difference is not essential, the second is. Indeed, much of our analysis turns around the limited-liability constraint. Despite this difference, we note that also in Malcomson's model is it usually optimal for the principal to not fully align the agent's incentives, at the moment of decision-making, with those of the principal (had she then had access to the agent's private information).

The two papers Zermeño (2011, 2012) provide a general methodology for analysis of principal-agent models in which the agent's task is to acquire information, just as in our model. In Zermeño (2011), the agent-cum-expert is risk neutral and there are only two states of nature. In Zermeño (2012) there is an arbitrary finite set of potential outcomes, and the agent can affect the realization probabilities. This second paper analyzes delegation and authority along the lines of Aghion and Tirole (1997).

Terovitis (2018) analyzes a risk neutral principal who delegates to a risk neutral agent the task of choosing between two actions. There are only two states of nature, and the payoff structure for the principal is symmetric with respect to taking the right or wrong action in each state. The agent may at a cost obtain a noisy signal, of fixed precision, before taking his action. Thus, the agent faces a binary information choice. By contrast, in our model his information choice set constitutes a continuum and we allow for an arbitrary number of states of nature. Another difference between the two models is that Terovitis (2018) assumes that the state of nature is always verifiable ex post, while we assume that only the returns to realized investments are verifiable.

A recent paper that is methodologically similar to ours is Yang and Zeng (2019). In their model, an entrepreneur with a production idea, but without money, proposes a security to an investor (say, a venture capitalist). Upon receiving this offer, the latter may acquire information about the entrepreneur's project. This information acquisition is modelled by way of a rational-inattention approach like ours. However, the underlying economic situation, and its payoff structure, is quite different from ours. For while our investor, like their entrepreneur, has only exogenous prior information, our investor has to risk the money, while in their model it is the other party, the venture capitalist, who both acquires information and risks the money. Moreover, in Yang's and Zeng's model, both parties are risk neutral, while our agent may be risk averse, and, moreover, we use our model to discuss other issues than they do (their focus is on debt versus equity).

To the best of our knowledge, the closest model to ours is that of Carroll (2019), who builds upon Zermeño (2011, 2012). However, there are several important differences: (1) Carroll does not use the rational-inattention approach to model information acquisition, (2) instead of a binary investment decision, Carroll allows for any non-empty compact set of alternative investment decisions, (3) the principal does not know the agent's set of information-acquisition technologies, (4) the agent is risk-neutral, (5) the agent can send a report with decision-relevant information to the principal, (6) the principal makes the investment decision, (7) the remunerations specified in the contract may be conditioned not only on the decision made and the realized return, but also on the report sent by the agent and on the full state of nature (in particular also on what the return would have been from investments not made), and (8), the principal chooses contract according to a maxmin ("worst case") criterion.[29] Also the results in the two papers differ. For while Carroll shows that a class of affine contracts are optimal in his model, optimal contracts in our model are non-linear. The different assumptions made in Carroll's and our model fit more or less well in different applications, and thus shed complementary light on delegation of information acquisition in connection with decision-making, a topic arguably of relevance for many real-life situations.

10. Conclusion

We have developed a model of delegation of investment decisions, where the agent's information acquisition is endogenous and modelled within the rational-inattention framework. The model is highly operational and can be used in applications and empirical work. We here use it to shed light on a number of issues surrounding lavish bonus schemes to managers in large corporations.

Assuming limited liability, we find that optimal contracts require a balance between the pay when not investing, and bonuses and penalties when investing. Optimal contracts are non-linear, and reward very profitable investments less than proportionally. Thus, the very high bonuses one some's sees in practice do not seem to have

[29] The author remarks (op. cit. p. 384) that the second main point of the paper is methodological, to "... show how using a maxmin objectie leads to a tidy and tractable model. By contrast, a traditional Bayesian approach, where the principal knows the expert's information acquisition technology (or has a probabilistic belief about it), is unlikely to be tractable without much more specific functional form assumptions, e.g. binary state and one-dimensional effort choice by the expert." Our model does not presume binary states or one-dimensional effort, and yet does, arguably, obtain a "tidy and tractable model". The reason is our modelling of information acquisition in terms of entropy reduction, admittedly an important functional-form assumption.

support in our model. However, we find that the expected utility for the agent may well exceed his reservation utility. The reason is limited liability; in order to make the agent avoid investing in bad states of nature it may be necessary to pay him well if not investing. Hence, it may be in the principal's interest to offer apparently lavish contracts.

Our analysis builds upon many heroic simplifications. A relevant extension would be to go deeper than here into cases of incomplete information about the agent's type, that is, his unit information cost, risk attitude, and/or outside option. When can a principal in such situations let the agent self-select from a menu of contracts? What if some agents are overconfident in the sense of underestimating their unit costs of information acquisition? It would also be interesting to extend the model to allow for agents with more complex motivation, such as loyalty towards their principal, morality etc. Another interesting extension would be to consider competition among principals for agents. We hope that the present model and analysis can serve as a first step in such and other extensions.

11. Appendix

11.1. The Lambert W function. The Lambert W multi-function (or correspondence), is implicitly defined by $y = \mathcal{W}(z)$ if and only if $ye^y = z$, where z is any complex number, see Corless et al. (1996). For real numbers z, this defines a correspondence $\mathcal{W}$ from the interval $[-1/e, +\infty)$ to $\mathbb{R}$, a correspondence that is singleton-valued for all $z \geq 0$. See diagram.

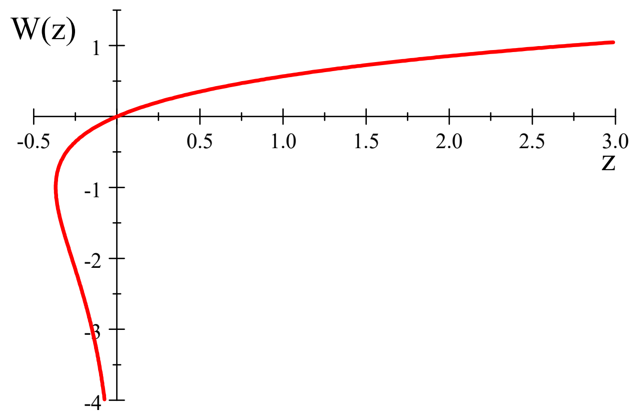


Figure A1: The Lambert W correspondence.

The upper branch of this correspondence is called its principal branch, and we denote it $W : (-1/e, +\infty) \to \mathbb{R}_+$. This function is differentiable and satisfies $W(0) = 0$, $W(e) = 1$, $W'(0) = 1$ and $W' > 0$.

11.2. Proof of Lemma 1. Write $f(q)$ for the left-hand side of (4). This defines $f : [0,1] \to \mathbb{R}$ as a twice differentiable function with

$$f'(q) = -\sum_{k \in \Omega} \frac{\mu_k e^{u(y_k)/c} \left(e^{u(y_k)/c_a} - e^{u(y_0)/c}\right)}{\left[q e^{u(y_k)/c} + (1-q) e^{u(y_0)/c}\right]^2} \tag{22}$$

and

$$f''(q) = 2\sum_{k \in \Omega} \frac{\mu_k e^{u(y_k)/c} \left(e^{u(y_k)/c} - e^{u(y_0)/c}\right)^2}{\left[q e^{u(y_k)/c} + (1-q) e^{u(y_0)/c}\right]^3} \geq 0 \tag{23}$$

for all $q \in [0,1]$. Hence, f is strictly convex. Since $f(1) = 1$, we have that $f(q) < 1$ for all $q \in (0,1)$ if $f(0) < 1$, and $f(q) > 1$ for all $q \in (0,1)$ if $f(0) > 1$ and $f'(1) \leq 0$, where

$$f(0) = \mathbb{E}\left[e^{[u(Y)-u(y_0)]/c}\right] = e^{-u(y_0)/c}\mathbb{E}\left[e^{u(Y)/c}\right]$$

and

$$f'(1) = -\sum_{k\in\Omega} \mu_k \frac{e^{u(y_k)/c} - e^{u(y_0)/c}}{e^{u(y_k)/c}} = e^{u(y_0)/c}\mathbb{E}\left[e^{-u(Y)/c}\right] - 1$$

which establishes all claims.

11.3. Proof of Proposition 1. Let $\vec{y} \in \mathbb{R}_+^{m+1}$ be a contract that solves program (12) when the agent's participation constraint is slack. Consider the associated constrained optimization program,

$$\max_{\vec{y},q} \quad \left[\sum_{k=1}^{m} \frac{\mu_k q e^{u(y_k)/c}\left(Ix_k - y_k + y_0\right)}{q e^{u(y_k)/c} + (1-q)\, e^{u(y_0)/c}} \quad - \quad y_0\right] \tag{24}$$

subject to the inequality constraints

$$0 \leq y_0, .., y_m \text{ and } 0 < q < 1, \tag{25}$$

and the equality constraint

$$\sum_{k=1}^{m} \frac{\mu_k e^{u(y_k)/c}}{q e^{u(y_k)/c} + (1-q)\, e^{u(y_0)/c}} = 1. \tag{26}$$

Writing λ for the Lagrangian associated with the latter, we obtain the following necessary first-order conditions for all $k > 0$ with $y_k > 0$:

$$\frac{\partial}{\partial y_k}\left[\frac{q\left(Ix_k - y_k + y_0\right) + \lambda}{q + (1-q)\, e^{[u(y_0)-u(y_k)]/c}}\right] = 0$$

or, equivalently,

$$\left[q\left(Ix_k - y_k + y_0\right) + \lambda\right](1-q)\,\frac{u'(y_k)}{c} e^{[u(y_0)-u(y_k)]/c} = q\left[q + (1-q)\, e^{[u(y_0)-u(y_k)]/c}\right]$$

or

$$q\left(Ix_k - y_k + y_0\right) + \lambda = \frac{cq}{(1-q)\, u'(y_k)}\left[q e^{[u(y_k)-u(y_0)]/c} + 1 - q\right]$$

or

$$Ix_k - y_k + y_0 = \frac{c}{(1-q)\, u'(y_k)}\left[q e^{[u(y_k)-u(y_0)]/c} + 1 - q\right] - \frac{\lambda}{q}$$

or

$$y_k = Ix_k + y_0 + \frac{\lambda}{q} - \frac{c}{u'(y_k)}\left(\frac{q}{1-q} e^{[u(y_k)-u(y_0)]/c} + 1\right). \tag{27}$$

For any given value of $q \in (0,1)$, the right-hand side is continuous and decreasing in y_k, from the value

$$Ix_k + y_0 + \frac{\lambda}{q} - \frac{c}{u'(0)}\left(\frac{q}{1-q}e^{[u(0)-u(y_0)]/c} + 1\right)$$

when $y_k = 0$. Hence, for all $x_k > \xi$, where

$$\xi = \left[\frac{c}{u'(0)}\left(\frac{q}{1-q}e^{[u(0)-u(y_0)]/c} + 1\right) - y_0 - \frac{\lambda}{q}\right]\frac{1}{I}$$

$y_k > 0$ is uniquely determined by (27), while for all $k \in \Omega$ with $\xi \leq 0$, $y_k = 0$. Moreover, when positive, y_k is continuous and strictly increasing in x_k. Finally, we note that if $y_1 > y_0$ then (4) has no solution in $(0,1)$. Likewise if $y_m < y_0$. The equations in the proposition are obtained for $\alpha = y_0 + \lambda/q$.

11.4. Proof of Proposition 2. We here assume that $u(y) \equiv y$.

Part I. Suppose first that the agent's participation constraint is slack. From the proof of Proposition 1, the necessary first-order condition for $\vec{y}$ to be optimal is

$$y_k - y_0 = Ix_k + \frac{\lambda}{q} - \frac{cq}{1-q}e^{(y_k-y_0)/c} - c \quad \forall k \in \Omega \text{ with } y_k > 0. \tag{28}$$

Moreover, for a risk-neutral agent (and for a given project X, investment I and unit information cost c), q and λ are functions of the transfer vector $\vec{t} = (t_1, ..., t_m)$ (with $t_k = y_k - y_0$ $\forall k \in \Omega$), irrespective of the value of y_0. Hence, we may write this equation in the form

$$t_k = Ix_k - \sigma e^{t_k/c} - \tau, \tag{29}$$

where $\sigma = cq/(1-q)$ is positive and $\tau = c - \lambda/q$, and both are independent of y_0 and I. Hence, if $\vec{y}$ is optimal and the agent's participation constraint is slack, then $t_k = \max\{0, \phi(x_k)\}$, where $\phi : \mathbb{R} \to \mathbb{R}$ is uniquely defined for all $x \in \mathbb{R}$ by the functional equation

$$\phi(x) = Ix - \sigma e^{\phi(x)/c} - \tau. \tag{30}$$

We have obtained, for each $x \in \mathbb{R}$, a fixed-point equation in $\phi(x) \in \mathbb{R}$ that (for given $I > 0$, $c > 0$ and $\tau \in \mathbb{R}$) uniquely determines $\phi(x) \in \mathbb{R}$. We also note that this equation defines $\phi : \mathbb{R} \to \mathbb{R}$ as a twice differentiable function with the following properties:

(i) $\phi(x) = 0$ iff $x = (\sigma + \tau)/I$

(ii) $\lim_{x\to-\infty} \phi(x) = -\infty$ and $\lim_{x\to+\infty} \phi(x) = +\infty$

(iii) $\phi(x) < Ix - \tau$ for all $x \in \mathbb{R}$, with $\lim_{x\to-\infty}(Ix - \phi(x)) = \tau$

(iv) $\phi' \in (0, I)$ and $\phi'' < 0$, with $\lim_{x\to-\infty} \phi'(x) = I$ and $\lim_{x\to+\infty} \phi'(x) = 0$

While (i)-(iii) follow immediately from (30), property (iv) follows from (ii) and differentiation of both sides of (30), which gives

$$\phi'(x) = \frac{Ic}{c + \sigma e^{\phi(x)/c}} \quad \forall x \in \mathbb{R}$$

This establishes the claims made concerning the function ϕ when the agent's participation constraint is not binding.

We proceed to show that this function ϕ can be expressed in terms of the Lambert W function. First, rewrite (29) as

$$\frac{\sigma}{c} \cdot e^{\phi(x)/c} + \frac{\phi(x)}{c} = \frac{Ix - \tau}{c}$$

The the equation can be written in the form

$$\psi(x) + \ln \psi(x) = \ln \frac{\sigma}{c} + \frac{Ix - \tau}{c}, \tag{31}$$

where

$$\psi(x) = \frac{\sigma}{c} e^{\phi(x)/c}.$$

Noting that the Lambert-W function satisfies $W(y) + \ln W(y) = \ln y$, we obtain that

$$\psi(x) = W\left(\frac{\sigma}{c} e^{(Ix-\tau)/c}\right) \quad \forall x \in \mathbb{R}$$

satisfies (31) for $\forall x \in \mathbb{R}$. Hence,

$$e^{\phi(x)/c} = \frac{c}{\sigma} W\left(\frac{\sigma}{c} e^{(Ix-\tau)/c}\right) \quad \forall x \in \mathbb{R}$$

or

$$\phi(x) = c \cdot \left[\ln W\left(\frac{\sigma}{c} e^{(Ix-\tau)/c}\right) - \ln \frac{\sigma}{c}\right] \quad \forall x \in \mathbb{R}$$

as claimed.

Part II. Now we allow for the possibility that the agent's participation constraint may be binding. The agent's expected utility under any contract $\vec{y} \geq 0$ can be written in the form

$$U(\vec{y}) = \sum_{k \in \Omega} \mu_k \left[\frac{q t_k}{q + (1-q) e^{-t_k/c}} + cH\left(\frac{q}{q + (1-q) e^{-t_k/c}}\right)\right] + y_0 - cH(q).$$

Let $\rho \geq 0$ be the Lagrangian associated with the agent's participation constraint. Consider any state of the world $k \in \Omega$ with $y_k > 0$. The associated Karush-Kuhn-Tucker condition, with λ denoting the Lagrangian associated with the equation for q, is

$$\frac{\partial}{\partial t_k}\left(\frac{qIx_k - q(1-\rho)t_k - \lambda}{q + (1-q)e^{-t_k/c}} + \rho c H\left[\frac{q}{q + (1-q)e^{-t_k/c}}\right]\right) = 0 \tag{32}$$

The derivative of the entropy function, H, is $H'(p) = \ln(1-p) - \ln p$ for all $p \in (0,1)$, so

$$\frac{\partial}{\partial t_k}\left[H\left(\frac{q}{q + (1-q)e^{-t_k/c}}\right)\right] = \left[\ln\left(\frac{1-q}{q}\right) - \frac{t_k}{c}\right] \cdot \frac{\partial}{\partial t_k}\left[\frac{q}{q + (1-q)e^{-t_k/c}}\right]$$

Therefore, (32) can be written as

$$\frac{\partial}{\partial t_k}\left[\frac{qIx_k - q(1-\rho)t_k - \lambda}{q + (1-q)e^{-t_k/c}}\right] + \rho\left[c\ln\left(\frac{1-q}{q}\right) - t_k\right] \cdot \frac{\partial}{\partial t_k}\left[\frac{q}{q + (1-q)e^{-t_k/c}}\right] = 0$$

or

$$\begin{aligned} 0 \;=\; & \frac{(1-q)\left[qIx_k - q(1-\rho)t_k - \lambda\right]e^{-t_k/c}}{c\left[q + (1-q)e^{-t_k/c}\right]^2} - \frac{q(1-\rho)}{q + (1-q)e^{-t_k/c}} + \\ & + \rho\left[c\ln\left(\frac{1-q}{q}\right) - t_k\right] \cdot \frac{q(1-q)e^{-t_k/c}}{c\left[q + (1-q)e^{-t_k/c}\right]^2} \end{aligned}$$

or

$$0 = Ix_k - t_k - \frac{\lambda}{q} - c(1-\rho)\frac{q}{1-q}e^{t_k/c} + c(1-\rho) + \rho \cdot \ln\left(\frac{1-q}{q}\right).$$

Equivalently:

$$\sigma^* e^{t_k/c} = Ix_k - t_k - \tau^* \tag{33}$$

for

$$\sigma^* = (1-\rho)\frac{qc}{1-q} \text{ and } \tau^* = (1-\rho)c - \frac{\lambda}{q} + \rho \cdot \ln\left(\frac{1-q}{q}\right)$$

We note that $\sigma^* > 0$ if $\rho < 1$, while otherwise $\sigma^* \leq 0$. We recall that the Lambert W function's domain is $(-1/e, +\infty)$. Applied to (33), this requires

$$(1-\rho)\frac{q}{1-q}e^{(Ix-\tau)/c} > -\frac{1}{e}$$

or

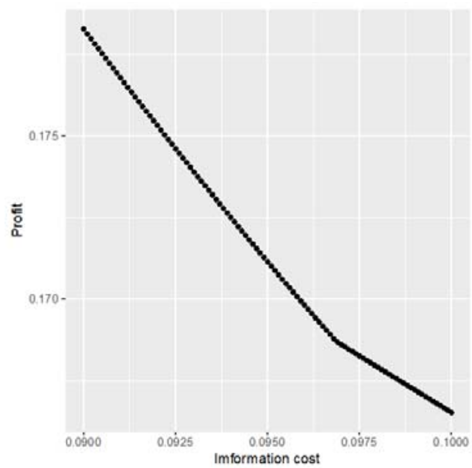
$$\rho < 1 + \frac{1-q}{q}e^{-(Ix-\tau+c)/c}.$$

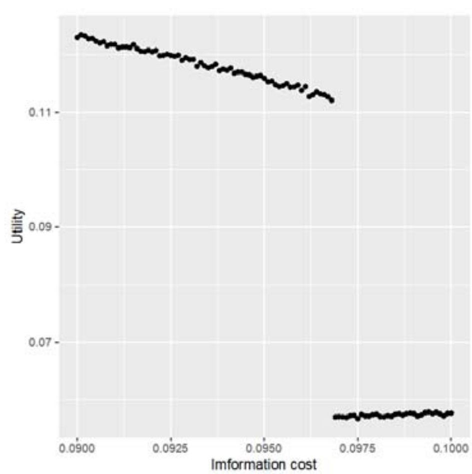
Hence, the right-hand side is an upper bound on the Lagrangian ρ for the Lambert-W representation to be valid.[30]

[30] We conjecture, but have not been able to prove, that $\rho \leq 1$ always holds.

11.5. Additional results.

Discontinuity of the agent's expected utility. The next diagrams show how the principal's expected profit and the agent's expected utility change as the agent's unit information cost, c, varies between 0.09 and 0.10. The diagrams suggests that the expected profit is indeed continuous (which theoretically follows from Berge's maximum theorem), and that the agent's expected utility has a discontinuity near $c = 0.96875$.





Career concerns. We here consider the effect of the agent's career concerns on optimal contracts in terms of Projects 3 in Section 4. Suppose that investment in the best state of nature enhances the agent's career prospects and that investment in the worst state of nature diminishes the agent's career prospects, while investment in the intermediate state, like non-investment, has no career effect. Let δ be the reduction of the present value of the agent's future earnings when investing in the bad state of nature, and let αx_3, for $\alpha > 0$, be the increase of the present value of the agent's future earnings when investing in the best state. Hence, in the notation of Section 2.2, $\gamma_1 = -\delta$, $\gamma_0 = \gamma_2 = 0$ and $\gamma_3 = \alpha x_3$. The tables below show the effects of these parameters on optimal contracts for the same risk-neutral agent as in Table 2. The first table in obtained without regard to the agent's participation constraint, the second table with due regard to it. We note that, not surprisingly, the principal's expected profit is higher when the agent is partly motivated by career concerns, and that the bonuses to the agent for investment in the best state of nature are lower.

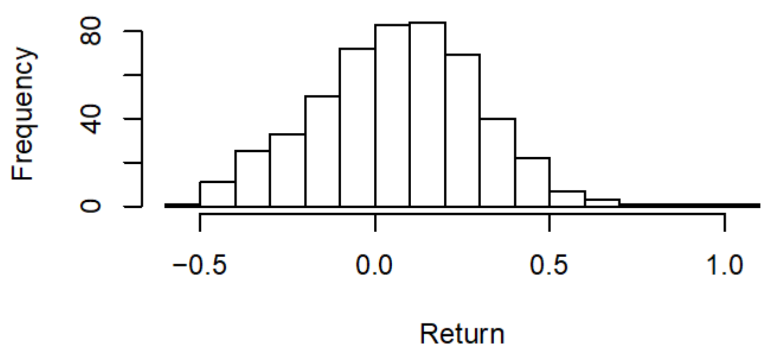
TABLE A1: Optimal contracts for Project 3 when the agent has career concerns but no participation constraint.

α	δ	y_0	y_1	y_2	y_3	Π	U
0	0	0.1032	0	0.1891	0.3209	0.2646	0.1275
0.01	0.01	0.0937	0	0.1797	0.2721	0.2769	0.1181
0.03	0.03	0.0746	0	0.1604	0.1746	0.3015	0.0989
0.03	0.05	0.0555	0	0.1420	0.1552	0.3205	0.0800
0.05	0.03	0.0747	0	0.1606	0.0953	0.3072	0.0991
0.05	0.05	0.0555	0	0.1418	0.0759	0.3261	0.0800
0.1	0.1	0.0113	0	0.0984	0	0.3794	0.0478

TABLE A2: Optimal contracts for Project 3 when the agent has career concerns and a participation constraint.

α	δ	y_0	y_1	y_2	y_3	Π	U
0	0	0.1032	0	0.1891	0.3209	0.2646	0.1275
0.01	0.01	0.0937	0	0.1797	0.2721	0.2769	0.1181
0.03	0.03	0.0752	0	0.1624	0.1749	0.3015	0.1
0.03	0.05	0.0697	0	0.1713	0.1812	0.3182	0.1
0.05	0.03	0.0751	0	0.1622	0.0964	0.3072	0.1
0.05	0.05	0.0696	0	0.1714	0.1019	0.3238	0.1
0.1	0.1	0.0514	0	0.1811	0	0.3638	0.1

11.6. Application to empirical return distributions. Proposition 2 enables operational application of the present model to arbitrary asset distributions with finite support when the agent is risk neutral. This is because every optimal contract belongs to a parametric family with only three parameters (y_0, σ and τ), irrespective of the number of potential outcomes and their values. To illustrate this, we applied the model to a stylized quasi-empirical return rate distribution, see histogram in the diagram below. This diagram is based on the return rates for the S&P 500 companies, as registered over the last 52 weeks, approximately July 1, 2018 until June 30, 2019. The horizontal axis is the return rate on individual stocks, while the vertical axis is the number of companies.



Annual return rates for S&P 500 companies, July 2018 - June 2019.

Transforming this to a probability distribution, we obtain a prior probability distribution μ for a fictitious random variable X, a representative S&P 500 company (with no claims that this is empirically valid). The expected value of this random variable is $\mathbb{E}[X] \approx 0.0659$, that is, an average annal return of about 6.6%. The variance is 0.0567. The probability for a positive net return rate, $\Pr[X > 0]$ is approximately 0.6190, and the conditionally expected net return rate, given that it is positive, $\mathbb{E}[X \mid X > 0]$ is about 0.2122. Hence, an upper bound for the expected profit for the principal (should he invest $I = 1$ and delegate the investment decision to an agent with zero information costs and zero reservation utility), is approximately $0.2122 \cdot 0.6190 = 0.13154$.

Consider a risk-neutral agent with unit information cost $c = 0.01$ and zero reservation utility. Setting $r = 0$, the optimal contract is then of the form (15), for

$y_0 \approx 0.01995$, $\sigma \approx 0.01533$, and $\tau \approx 0.02636$, and it results in expected profit 0.0839 to the principal and expected utility 0.0269 to the agent. It follows from Proposition 2 that the agent is paid a bonus (malus) if investing in states of nature with net return rate x above (below) $\sigma + \tau \approx 0.042$. In particular, no bonuses for ex post bad investments. For agents with higher unit information cost, the expected profit to the principal is even lower. Hence, the gains from trade are very small. For $c \gtrsim 0.0174$ it is optimal for the principal to instead invest single-handedly and blindly (thereby obtaining an expected profit of 0.0659). This suggests that index funds may be better than reliance on less than excellent investment advisors.

References

[1] Aghion, P., and J. Tirole (1997): "Formal and real authority in organizations", *Journal of Political Economy* 105, 1-29.

[2] Bergemann, D. and A. Pavan (2015), "Introduction to symposium on dynamic contracts and mechanism design", *Journal of Economic Theory* 159, 679-701.

[3] Cabrales, A., O. Gossner, and R. Serrano (2013): "Entropy and the value of information to investors", *American Economic Review* 103, 360-377.

[4] Caplin, A., M. Dean, and J. Leahy (2019): "Rational Inattention, optimal consideration sets, and stochastic choice", *Review of Economic Studies* 86, 1061–1094.

[5] Carroll, G. (2019): "Robust incentives for information acquisition", *Journal of Economic Theory* 181, 382-420.

[6] Corless, R., G. Gonnet, D. Hare, D. Jeffrey, and D. Knuth (1996): "On the Lambert W function", *Advances in Computational Mathematic*s 5, 329-359.

[7] Crémer, J., F. Khalil, and J.-C. Rochet (1998): "Contracts and productive information gathering", *Games and Economic Behavior* 25, 174-193.

[8] Demski, J., and D. Sappington (1987): "Delegated expertise", *Journal of Accounting Research* 25, 68-89.

[9] Fosgerau, M., E. Melo, A. de Palma, and M. Shum (2020): "Discrete choice and rational inattention: a general equivalence", International Economic Review, June issue.

[10] Friebel, G., and M. Raith (2004): "Abuse of authority and hierarchical communication", *RAND Journal of Economics* 35, 224-244.

[11] Gabaix, X., and A. Landier (2008): "Why has CEO pay increased so much?", *Quarterly Journal of Economics, February*, 49-100.

[12] Hobson, A. (1969): "A new theorem of information theory", *Journal of Statistical Physics* 1, 383-391.

[13] Ibert, M., R. Kaniel, S. Van Nieuwerburgh, and R. Vestman (2018): "Are mutual fund managers paid for investment skill?", *Review of Financial Studies* 31(2), 715-772.

[14] Jaynes, E. (1957): "Information theory and statistical mechanics I", *Physical Review* 106, 171-190.

[15] Kaplan, S. and P. Strömberg (2002): "Financial contracting theory meets the real world: An empirical analysis of venture capital contracts", *Review of Economic Studies* 100, 1-35.

[16] Kaplan, S. and P. Strömberg (2004): "Characteristics, contracts, and actions: Evidence from vernture capitalist analyses", *Journal of Finance* 59, 2177-2210.

[17] Kaplan, S. (2013): "CEO pay and corporate governance in the U.S.: Perceptions, facts and challenges", *Journal of Corporate Finance* 25, 8-25.

[18] Kaplan, S., and J. Rauh (2010): "Wall Street and Main Street: What constitutes to the rise in the highest incomes?", *Review of Financial Studies* 23, 1004-1050.

[19] Kullback, S. (1959): *Information Theory and Statistics.* John Wiley & Sons: New York.

[20] Lambert, R. (1986): "Executive effort and the selection of risky projects", *The RAND Journal of Economics* 17, 77-88.

[21] Levitt, S., and C. Snyder (1997): "Is no news bad news? Information transmission and the role of "early warning" in the principal-agent model", *RAND Journal of Economics* 28, 641-661.

[22] Lewis, T., and D. Sappington (1997): "Information management in incentive problems", *Journal of Political Economy* 105, 796-821.

[23] Lindbeck, A. and J. Weibull (2017): "Investment, rational inattention, and delegation", IFN Institute for Industrial Economics, Stockholm, WP No. 1171.

[24] Ma, L., Y. Tang and J.-P. Gómez (2019): "Portfolio manager compensation in the U.S. mutual fund industry", *Journal of Finance* 74 (2), 587-638.

[25] Malcomson, J. (2009): "Principal and expert agent", *B.E. Journal of Theoretical Economics* 9, Issue 1 (Contributions), Article 17.

[26] Matějka, F., and A,. McKay (2015): "Rational inattention to discrete choices: a new foundation for the multinomial logit model", *American Economic Review* 105, 272-298.

[27] Murphy, K.J. (2012): "Executive compensation: where we are and how we got there", in G. Constantinides, M.Harris and , R. Stulz (eds), *Handbook of the Economics of Finance.* Amsterdam: North-Holland.

[28] Sims, C. (2003): “Implications of rational inattention”, *Journal of Monetary Economics* 50, 665-690.

[29] Sims, C. (2006): “Rational inattention: beyond the linear-quadratic case”, *American Economic Review* 96, 158-163.

[30] Shannon, C. (1948): “A mathematical theory of communication”, *The Bell System Technical Journal* 27, 379-423.

[31] Shannon, C., and W. Weaver (1949): *The Mathematical Theory of Communication*. University of Illinois Press, Urbana, IL.

[32] Swedroe, L. (2014): "What’s the best incentive scheme for fund managers?", *Moneywatch*, May 22.

[33] Terovitis, S. (2018): "Motivating information acquisition under delegation", mimeo, University of Amsterdam, Finance Group.

[34] Terviö, M. (2008): “The difference that CEOs make: An assignment model approach”, *American Economic Review*, 98:3, 642-668.

[35] Woodford, M. (2008): “Inattention as a source of randomized discrete adjustment”, mimeo., Columbia University.

[36] Yang, M. (2015): “Coordination with flexible information acquisition”, *Journal of Economic Theory* 158, 721-738.

[37] Yang, M., and Y. Zeng (2019): “Financing entrepreneurial production: security design with flexible information acquisition”, *Review of Financial Studies* 32, 819-863.

[38] Zermeño, L. (2011): “A principal-expert model and the value of menues”, mimeo., MIT.

[39] Zermeño, L. (2012): “The role of authority in a general principal-expert model”, mimeo., MIT.