RESEARCH INSTITUTE OF INDUSTRIAL ECONOMICS

IFN Working Paper No. 1171, 2017

Delegation to a Rationally Inattentive Agent

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 to a rationally inattentive agent

Assar Lindbeck* and Jörgen Weibull†

November 29, 2017.‡

Abstract. In the political debate, lavish bonus schemes to managers in large corporations are often questioned on moral grounds as well as in terms of economic efficiency. Can high bonuses, sometimes even after *ex post* unprofitable decisions, be motivated from first principles in economics? We analyze this question through the prism of a model in which an investor delegates information acquisition and investment decisions to a rationally inattentive agent. The agent's effort to acquire information is private, as is the information obtained and its quality. We identify a class of optimal contracts for such situations, and show that not even in the case of a risk-neutral agent is it in the principal's interest to perfectly align the agent's incentives with her own. Morover, unlike practice, optimal contracts include penalties as well as bonuses. Due to limited-liability, risk-neutral agents' contracts will be asymmetric in the direction of rewarding profitable investments more than penalizing unprofitable investments, and may in fact reward the agent after unprofitable investments.

Keywords: Delegation, 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.

†Stockholm School of Economics, Institute for Advanced Study in Toulouse, and KTH Royal Institute of Technology, Stockholm.

‡This is a substantially revised and extended version of WP 1171, "Investment, rational inattention, and delegation". The authors thank Atahan Afsar, Ingela Alger, Bo Becker, Jun Chen, Lee Dinetan, Sven-Olof Fridolfson, Jungsuk Han, Bengt Holmström, Gustav Karreskog, Augustin Landier, Filip Matějka, Paul Milgrom, Francois Salanié, Francesco Sangiorgi, Per Strömberg, Thomas Tangerås, Peter Wikman and Ming Yang for helpful comments to different versions of this manuscript. We thank Atahan Afsar for excellent research assistance. Assar Lindbeck thanks the Jan Wallander and Tom Hedelius Foundation, and the Marianne and Marcus Wallenberg Foundation for financial support. 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.

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 equally important task for many agents, such as CEOs of large corporations and pension-fund managers, is to make well-informed choice of which project, if any, to choose. This was noted many years ago by Demski and Sappington (1987), and this is the theme we here focus on. While powerful analytical tools for this task were lacking at the time of Demski's and Sappington's work, the situation is quite different today, in particular after the development of the rational-inattention approach to information acquisition (Sims 1998, 2003, 2006, and followers). Instead of pushing the technical potential of the approach further, we here explore its usefulness for economics.[1] In order to check the robustness of the results with respect to modelling approach we also develop a canonical signal-extraction model and compare the predictions.

We here take the agent to be someone who has a comparative advantage over the principal in acquiring and evaluating pertinent information about projects or investment opportunities. Such information acquisition and processing usually requires the agent to exert (non-contractible) effort, and the information obtained is often private and difficult to communicate to the principal. There are two obstacles to communication. First, the principal may not be qualified, or have the time needed, to understand and assess the contents, reliability and relevance of the information that the agent has acquired. Second, it may be in the agent's self-interest not to share all information—for instance, if he has negative information about a project for which he would be well paid. The agent may, more generally, opportunistically misreport or suppress information. For this reason, not only information acquisition but also the investment decision itself (once information has been acquired) is either formally or in practice delegated to the agent. An important issue is thus how to motivate the agent to first acquire relevant and reliable information and then, when making the investment decision, use the information in line with the principal's interest.

The topic being rich and complex, we abstract from many important real-life factors and focus only on a few key elements. We assume that the agent is purely self-interested and only cares about his own remuneration and work effort. We also assume that neither the quality of his information nor its contents can be communicated to the principal. Any contract between the two parties can thus only be conditioned on whether investment is made, and if made, its realized 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

[1] For technical extensions concerning principal-expert models, see Zermeño (2011, 2012).

[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,

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 to face some penalty after unsuccessful investments. To have an agent who abstains from investment when prospects (about which he has private information) are not good can be just as important for the principal as to have an agent who invests when prospects are good. We will call the "non-investment pay" the agent's *salary*. To choose the agent's salary is a delicate matter. It needs to be balanced against potential bonuses and penalties. All contracts have to meet the limited-liability constraint that the net pay from the principal to the agent has to be nonnegative in all states of nature. Under this constraint, a low salary limits the size of penalties, since the maximal penalty then is to withdraw the whole salary. This causes an asymmetry between bonuses and penalties, even for risk neutral agents facing investment projects that are ex ante symmetric in terms of potential losses and gains.

Canonical examples of what we have in mind is a CEO of a large corporation, a division manager, or a manager of a pension fund. Because of the personal cost to the manager, say, in terms of long hours at work, there is here a major moral hazard issue, an issue we now tackle. Other examples are given by the consulting industry, where firms, institutions and sometimes single individuals decide whether or not to delegate information acquisition and, in effect, also decision-making to a consultant or expert whose effort and information cannot be monitored.

Although the paper emphasizes the interaction between principal and agent, we begin by analyzing an investor in autarky, in order to set the stage and have a benchmark. The investor-cum-principal is assumed to be risk neutral, while the agent may be risk neutral or risk averse. The investment decision is binary, such as whether or not to undertake a risky project—say, purchase an asset, start a new product line, build another production unit or research unit, enter a new market, or buy another company. The project's return is random and unknown at the time of investment. The principal and agent have the same prior beliefs about its probability distribution. The realized return from the project, if undertaken, is verifiable. A contract between the two parties specifies a payment to the agent under every possible outcome, including non-investment. Having signed such a contract, the agent decides how much, if any, effort to make in order to acquire information about the project at hand. The more effort he makes, the higher is the precision of his obtained information. The terms of the contract influence the agent in two distinct ways. First, it motivates the agent to acquire information. Second, it guides the agent's investment decision, once his information has been obtained. As will be seen, there is, in general, a tension between these two goals, a possibility pointed out by Demski and Sappington (1987).

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 (2017).

They provide sufficient conditions for when this tension, which they call *induced moral hazard*, may arise (their Proposition 1), and illustrate this possibility with a numerical example.[3] We here provide methods that permit further analysis of this tension.

We report five main findings. The first is that optimal contracts contain not only bonuses but also penalties, that is, reduction of the payment to the agent after unprofitable investments, although such reductions appear to be non-existent or rare in practice. Optimal bonuses and penalties are monotonic and non-linear in the realized net return of the investment, and are precisely characterized (Corollary 4 and Proposition 3).

Our second finding is that even when the agent is risk-neutral, it is not optimal for the principal to fully align the agent's incentives, at the moment of the investment decision, with those of the principal (had the principal then had the agent's private information). In particular, a contract consisting of a salary and a fixed share of the investment's net return (say, stocks) is suboptimal. The reason is limited liability in combination of the need to incentivize the agent to acquire information. Had the quality or precision of the agent's information been exogenous (as in standard principal-agent models), there would have been no such tension. However, in order to induce the agent to gather information, which in general requires both bonuses and penalties, the limited-liability constraints makes penalties more costly to the principal than bonuses. The reason is that an increase of penalties requires a raise both of the agent's fixed salary and also of all bonuses. A way to obtain this in practice could be to let the contract consist of three parts; a salary, some stocks, and also some options (to buy future stocks at their current price). While making the agent somewhat too "trigger happy" at the moment of the investment decision, such a contract may incentivize him to be well informed to such an extent that the combination is optimal from the principal's viewpoint. Hence, while real-life bonuses may be exaggerated, an upwards-biased bonus, salary and penalty scheme may in fact be optimal—even when the agent is risk neutral.

Our third finding is that it may be optimal for the principal to offer the agent a contract under which the agent's participation constraint is not binding. The reason for this is, again, the need to incentivize the agent for information acquisition. Had the quality or precision of the agent's information been exogenous, as in standard models of asymmetric information, such a slack would not arise. However, in order to motivate the agent to acquire information, it may be necessary to raise his salary in order to increase penalties, which are bounded by the limited-liability constraint, and bonuses, resulting in a contract that is strictly preferred by the agent over his

[3] In addition, they show that induced moral hazard cannot arise when there are only two possible outcomes (their Proposition 2). That condition is not met in our model, since there are at least three outcomes in the present model (non-investment, a bad and a good return).

outside option. Hence, in some situations lavish contract may in fact be in the best interest of the principal. We do not deny the importance of real-life agents' potential bargaining power over the terms of their own contracts. Indeed, such power is an obvious consequence of agents' outside options, which is part and parcel of the present model, but it may also depend on other factors not accounted for here (such an incumbent manager's potential better information about the internal and external conditions of the company).

Our fourth finding is that optimal contracts for rationally inattentive agents needs to avoid creating either "sweet" or "sour" investment conditions for the agent. By "sweet" investment conditions we mean conditions (defined by the terms of the contract, the size of the investment, and the project's return distribution) under which it is optimal for the agent to invest "blindly", that is, without acquiring any information at all. The opposite case, "sour" investment conditions, is when a rationally inattentive agent decides to not acquire any information and also not to invest. In the remaining case, which we call "normal" investment conditions and which fall between these two extremes, the agent will acquire *some* information and make the investment decision conditional upon the received information.

Our fifth and final finding is that the rational-inattention approach makes the delegation problem more tractable than under a canonical signal-extraction model, and yet the two approaches yield qualitatively similar results (at least) in simple examples. While the nesting of the agent's optimization program within the principal's optimization program is analytically non-trivial (and non-convex) in the signal-extraction model, it is easy (and convex) in the rational-inattention model.

To the best of our knowledge, we are among the first to use the rational-inattention approach to principal-agent analysis. The only such study we know of is Yang and Zeng (2017). They analyze optimal contracts between an investor with money and an entrepreneur with ideas but no money. The investor trades off resources spent on collecting costly information about the entrepreneur's project against resources spent on financing the project. In their model, the principal takes the financing decision—a choice between debt and equity—and the entrepreneur, who is the agent, takes production decisions.

Apart from the above-mentioned paper, and the pioneering paper Demski and Sappington (1987), the papers closest to our seem to be Lewis and Sappington (1997), Levitt and Snyder (1997), and Malcomson (2009). The first paper analyzes situations in which the agent, at a given cost, can choose to be perfectly informed about the state of nature. By contrast, our agent chooses his degree of information on a continuum scale. The second paper analyzes 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.[4] 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, but enhances the agent's (private) information about the state of nature. Moreover, in our model the agent who makes the investment decision on behalf of the principal. The third paper, Malcomson (2009), also analyzes a principal who hires an agent to acquire information and make a decision. The agent first takes a (for him costly) action that affects the distribution of a signal, then observes the signal, and thereafter makes the decision. Malcomson's model allows for a continuum range of decisions, but disallows the possibility that contracts be conditioned on the agent's decision. Despite these differences, also in Malcomson's model it is usually optimal for the principal to not fully align the agent's incentives, at the moment of decision-making, with those of the principal herself, had she then had access to the agent's private information. By assuming that contracts can be conditioned on the agent's decision, and by using the rational-inattention approach to information acquisition, we obtain a number of new results.

Our approach differ from that in Crémer and Khalil (1992) and Crémer, Khalil and Rochet (1998a). These papers analyze situations in which the agent can, at a cost, acquire information about the state of nature before signing a contract, while we here assume that the agent can acquire information only after the contract has been signed. The latter assumption is also made in Crémer, Khalil and Rochet (1998b), but in that model the agent has to choose to either be completely uninformed, or, at a fixed cost, obtain full information, while in our model the agent chooses from a continuum scale of degrees of information. Our model also differs from that in Carroll (2017), where a principal hires an expert (or agent) who can obtain relevant information by exerting costly effort, just like in our model. However, in Carroll's model the agent reports his information to the principal, who makes the decision based on the reported information, and both parties will know, *ex post*, the realized returns also to projects not chosen (and hence the veracity of the agent's reported information). Moreover, the principal does not know the agent's information acquisition technology and chooses contract according to its worst-case performance with respect to information technology.

The presentation of the material is organized as follows. Our model of a rationally inattentive investor is detailed in Section 2. In Section 3 the investor considers the possibility of delegating information acquisition and the investment decision to an agent. Since the high remuneration of CEO:s in large corporations often are criticized by the general public (see discussion in Murphy, 2012), we pay extra attention to

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

examine social efficiency and the nature of optimal contracts. Section 4 studies the robustness of the model in Section 3 by developing an alternative model in which the agent observes a signal about the true state of nature, with additive normal noise and endogenous signal precision.[5] Section 5 concludes. Mathematical proofs are provided in an appendix.

2. A rationally inattentive investor

Before embarking on the main topic of this study, delegation, we briefly consider a risk-neutral and rational investor who in autarky faces an indivisible investment opportunity, or *project*. The project requires a lump-sum investment, $I > 0$, and gives a random return, Y. The project's net return is thus $Y - I$ and its *net return rate* is the random variable

$$X = Y/I - 1. \tag{1}$$

The investor's prior belief about the net return rate X has finite support $M = \{x_1, ..., x_m\}$, where $x_1 < x_2 < ... < x_m$, and the prior probability for each such potential realization x_i is positive and denoted $\mu(x_i)$ or simply μ_i. The vector $\mu = (\mu_1, ..., \mu_m)$ is thus the investor's prior. This prior may be based on public information or knowledge about the economy at large, the industry in question, and on freely available information about the project at hand. Given her prior, the investor first decides how much effort, if any, to make in order to obtain more information about the project before she decides whether or not to invest in it.

If the investor chooses not to invest, she obtains a risk free net-return rate r. Hence, Ir is her opportunity cost for investing, or her outside option. For instance, r can be the risk-free interest rate in a credit market to which the investor has access. If she decides not to try to acquire further information about the project, then she will invest in the project if and only if its net return rate is non-negative, $\mathbb{E}[X] \geq r$. If she instead decides to acquire information, which is costly for her (in terms of time and effort), she may subsequently change her beliefs about the project's future return, depending on the information obtained. She then chooses to invest if and only if the conditionally expected net return from investing, given all her information, exceeds the risk-free rate r. We assume throughout that $x_1 < r < x_m$ (otherwise the decision problem is trivial).

We analyze this scenario by applying the rational-inattention model version due to Matějka and McKay (2015). Accordingly, we treat 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 must equal the decision-maker's

[5] This section builds upon an earlier working paper, see Lindbeck and Weibull (2017).

prior belief, and with information costs represented in terms of entropy reduction.[6] This will lead to investment probabilities that are conditional on the true state of nature, and depend on the investor's choice of how well-informed she wants to be. By spending more time and effort on information acquisition, she can reduce the risk of investing in bad states of nature and enhance the chances for investing in good states.

By applying Theorem 1 and Lemma 2 in Matejka and McKay (2015), we obtain that the investor's optimal information-*cum*-investment strategy induces the following conditional investment probabilities:

$$\hat{p}_{|X=x} = \frac{\hat{q}e^{xI/c}}{\hat{q}e^{xI/c} + (1-\hat{q})\, e^{rI/c}} \quad \forall x \in M, \tag{2}$$

where

$$\hat{q} \in \arg\max_{q\in[0,1]} \sum_{x\in M} \mu(x) \cdot \ln\left[qe^{xI/c} + (1-q)\, e^{rI/c}\right], \tag{3}$$

and $c > 0$ is the investor's unit cost of information, see below. The maximand in (3) is continuous and strictly concave in q, and the constraint set is convex and compact, so the maximizer $\hat{q}$ is uniquely determined. If $\hat{q}$ lies strictly between zero and one, then it necessarily satisfies the associated first-order condition, which can be written in the form (for a proof, see Appendix A):

$$f(q) = 1, \tag{4}$$

where $f : [0,1] \to \mathbb{R}$ is defined by

$$f(q) = \sum_{x\in M} \frac{\mu(x) \cdot e^{xI/c}}{qe^{xI/c} + (1-q)\, e^{rI/c}}. \tag{5}$$

It follows from (2) and (5) that the product $\hat{q}f(\hat{q})$ equals the *ex ante* investment probability, $\mathbb{E}\left[\hat{p}_{|X}\right]$. From this we conclude that $\hat{q} = \mathbb{E}\left[\hat{p}_{|X}\right]$ if $0 < \hat{q} < 1$ (since then $f(\hat{q}) = 1$).[7] In fact, it is easily verified that $\hat{q} = \mathbb{E}\left[\hat{p}_{|X}\right]$ also when $\hat{q} = 0$ and $\hat{q} = 1$. In sum: $\hat{q}$ is the *ex ante* investment probability, the probability of investment before information has been acquired: $\hat{q} = E\left[\hat{p}_{|X}\right]$.

---

[6] The information-theoretic interpretation of entropy is due to Shannon (1948), and further developed by Shannon and Weaver (1949), Jaynes (1957), Kullback (1959) and Hobson (1969). For applications to economics and decision theory, see Mattsson and Weibull (2002), Gossner, Hernandez and Neyman (2006), Woodford (2008), Cabrales, Gossner and Serrano (2013), Yang (2015), Steiner, Stewart, and Matějka (2017), and Fosgerau, Melo, and Schum (2017). A recent comprehensive study is given in Caplin, Dean, and Leahy (2017).

[7] By definition of the function f it is clear that (4) always has a trivial corner solution, namely $q = 1$, which, however, is "out of bounds" since the equation is required to hold only when $\hat{q} \in (0,1)$.

The investor's conditionally expected profit, given that the investment's true net return rate is $x \in M$, consists of three terms. The first term is the net return xI from investing, multiplied by the conditional investment probability in that state of nature. The second term is the net return from not investing, multiplied by the conditional probability for not investing in that state. The third term is the investor's information cost. Formally, for all $x \in M$:

$$\Pi(x,c) = \frac{xI \cdot \hat{q}e^{xI/c}}{\hat{q}e^{xI/c} + (1-\hat{q})e^{rI/c}} + \frac{rI \cdot (1-\hat{q})e^{rI/c}}{\hat{q}e^{xI/c} + (1-\hat{q})e^{rI/c}} - c \cdot \left[H(\hat{q}) - H\left(\hat{p}_{|X=x}\right)\right]. \tag{6}$$

Here $c > 0$ is the investor's *unit cost of information*, mentioned above, and H is the entropy function for a binary probability distribution, defined by

$$H(p) = -\left[p \ln p + (1-p)\ln(1-p)\right], \tag{7}$$

for all $p \in [0,1]$ (with the convention $0 \ln 0 = 0$).[8]

In (6), $H(\hat{q})$ is thus the entropy of the ex ante investment probability and $H\left(\hat{p}_{|X=x}\right)$ the entropy of the conditional investment probability in state $x \in M$. Consequently, the difference $H(\hat{q}) - H\left(\hat{p}_{|X=x}\right)$ is the entropy reduction when moving from the prior investment probability $\hat{q}$ to the posterior investment probability $\hat{p}_{|X=x}$ in state of nature x. Maximal entropy reduction would be obtained if the investor would almost surely invest precisely in those states where the net return rate x exceeds the interest rate r. However, to obtain such precise information is prohibitively costly and hence not optimal. The investor thus has to trade off information costs against information benefits for her subsequent investment decision.

Taking expectations according to the investor's prior μ, we obtain the following expression for the investor's *ex ante expected* profit:

$$\hat{\Pi}(I,c) = I \cdot \left[\sum_{x \in M} \frac{\hat{q} \cdot \mu(x) \cdot x}{\hat{q} + (1-\hat{q})e^{(r-x)I/c}} + (1-\hat{q})r\right] - c \cdot \left[H(\hat{q}) - \mathbb{E}\left(H\left(\hat{p}_{|X}\right)\right)\right], \tag{8}$$

[8]Entropy represents the uncertainty embodied in a probability distribution. Shannon (1948), characterizes axiomatically entropy as a quantitative measure of the uncertainty inherent in any probability distribution with finite support. More precisely, he shows that entropy is the unique measure (up to scaling) with the following three properties. First, continuity in the probabilities. Second, for equally likely outcomes, the measure should be larger the more outcomes there are. Third, if the random mechanism generating the outcomes consists of two steps, where first is a random draw of a cell from a partitioning of the set of outcomes, and the second step is a random draw from within the selected cell. Then the uncertainty measure should be the probabilistically weighted sum of the uncertainty measures of the random draws within each cell.

Cabrales, Gossner, and Serrano (2013) consider an investor who can purchase information in the form of "information structures", and they show that entropy represents the informativeness ordering of such information structures.

where $\mathbb{E}\left[H\left(\hat{p}_{|X}\right)\right]$ is the *ex ante* expected entropy of the conditional investment probability,

$$\mathbb{E}\left[H\left(\hat{p}_{|X}\right)\right] = \sum_{x\in M} \mu(x) \cdot H\left(\hat{p}_{|X=x}\right). \tag{9}$$

The difference $H(\hat{q}) - \mathbb{E}\left[H\left(\hat{p}_{|X}\right)\right]$ in (8) thus represents the *expected* reduction in entropy when moving from the prior $\hat{q}$ to the random posterior $\hat{p}_{|X}$.

Going back to how $\hat{q}$ is determined in (4), one sees that $\hat{q} = 1$ if all net returns $x_i > r$ (since then $f(q) > 1$ for all $q < 1$). Likewise, $\hat{q} = 0$ if $x_i < r$ (since then $f(q) < 1$ for all $q < 1$). In the first case, the investor always invests, $\hat{p}_{|X} = 1$ (a.s.), while in the second case she never invests, $\hat{p}_{|X} = 0$ (a.s.). In both cases, she wastes no resources on information acquisition: $\mathbb{E}\left[H\left(\hat{p}_{|X}\right)\right] = H(\hat{q}) = 0$. In other words, she then makes her investment decision "blindly". The phenomenon of "blind" decisions, to decide without acquiring information, occurs also in less stark situations. We will say that investment conditions are *sweet* when it is optimal to invest blindly, $\hat{q} = 1$, and that investment conditions are *sour* when it is optimal to blindly not invest, $\hat{q} = 0$. In all other cases, $0 < \hat{q} < 1$, investment conditions will be called *normal*. The following result characterizes the three investment conditions:[9]

Proposition 1. *The ex-ante investment probability $\hat{q} = \mathbb{E}\left[\hat{p}_{|X}\right]$ is uniquely determined by (2), (3) and (4). Investment conditions are normal if*

$$-c\ln\mathbb{E}\left[e^{-XI/c}\right] < rI < c\ln\mathbb{E}\left[e^{XI/c}\right], \tag{10}$$

sour if

$$rI \geq c\ln\mathbb{E}\left[e^{XI/c}\right], \tag{11}$$

and sweet if

$$rI \leq -c\ln\mathbb{E}\left[e^{-XI/c}\right]. \tag{12}$$

The sourness condition (11) says that for high enough interest rates it is not worthwhile for a rational investor to even consider the project. The sweetness condition (12) says that if the interest rate is low enough, then it is rational to invest without bothering to acquire further information about the project (beyond the information represented by the prior). The normality condition (10) identifies the intermediate range of interest rates at which it is worthwhile for the investor to acquire some information about the project and, if this information is favorable enough, to invest. Moreover:

[9] For other results that provide conditions under which a rationally inattentive decision-maker will not acquire any information, see Lemma 2 in Woodford (2008) and Proposition 1 in Yang (2015).

Corollary 1. *For every project X and any unit information cost $c > 0$ there exists a nonempty interval of interest rates r under which investment conditions are normal.*

In sum, the investor's expected profit, when using her optimal information-*cum*-investment strategy, is given by

$$\hat{\Pi}(I,c) = \begin{cases} I \cdot r & \text{if (11)} \\ c \cdot \pi(I/c) & \text{if (10)} \\ I \cdot \mathbb{E}[X] & \text{if (12)} \end{cases} \tag{13}$$

where, for any $v > 0$,

$$\begin{aligned} \pi(v) \quad = \quad & \left[(1-\hat{q})\, r + \sum_{x \in M} \frac{\hat{q} \cdot \mu(x) \cdot x}{\hat{q} + (1-\hat{q})\, e^{(r-x)v}}\right] \cdot v \\ & - H(\hat{q}) \; + \; \sum_{x \in M} \mu(x)\, H\left(\frac{\hat{q}}{\hat{q} + (1-\hat{q})\, e^{(r-x)v}}\right), \end{aligned} \tag{14}$$

and, under normal investment conditions, $\hat{q} \in (0,1)$ satisfies

$$\sum_{x \in M} \frac{\mu(x)}{\hat{q} + (1-\hat{q})\, e^{(r-x)v}} = 1. \tag{15}$$

The last equation implies that $\hat{q} \to \Pr[X > r]$ as $v \to +\infty$.[10] In other words, if the investment I is very large and/or the unit information cost c is very low, then the *ex ante* probability of investment, $\hat{q}$, is close to the probability that the net return exceeds the interest rate. In the limit, the investor acts as if she knew the true state of nature.

Having examined the investor's investment problem when acting in autarky, we now turn to the main topic of this study.

3. DELEGATION TO A RATIONALLY INATTENTIVE AGENT

Suppose now that the investor considers the possibility of hiring a rationally inattentive agent who has some comparative advantage in obtaining and processing relevant information. Suppose that the principal knows the agent's unit information cost and risk attitude. However, the principal does not know the agent's effort, the quality of the agent's information, or what information this is. The only verifiable information is whether or not investment was made, and, if made, its realized return. In view of

[10] We here assume that X has probability zero of being exactly equal to the interest rate r.

these informational difficulties, the principal delegates the investment decision to the agent, if hired. The agent is either risk neutral or risk-averse, and we take him to be purely self-interested.

To be more precise, we consider contracts of the form $\langle y, \boldsymbol{w} \rangle$, where $y \in \mathbb{R}$ is the agent's pay if he does not invest, and $\boldsymbol{w} : M \to \mathbb{R}$ is a *payment scheme*, a function that specifies the reimbursement $w_i = \boldsymbol{w}(x_i)$ to the agent for every possible realized net return rate $x_i \in M$ from investment in the project. We focus on contracts that meet the limited-liability constraint of never requiring the agent to make a net payment to the principal. More precisely: $y \geq 0$ and $\boldsymbol{w}(x_i) \geq 0$ in all states of nature $i = 1, ..., m$. We will refer to y as the agent's *salary* and the difference $\boldsymbol{w}(x) - y$ as his *bonus*, if positive, or *penalty*, if negative. The agent's Bernoulli function of income, u, is taken to be strictly increasing and continuous. His unit cost of information acquisition is $c_a > 0$, and his outside option has expected utility $\bar{u}$.

We obtain similar expressions for the agent, once hired, as were obtained for the investor in autarky. The agent's salary will play the same role as the investor's opportunity cost rI. To be more precise, let q^* be the *a priori* probability that the agent will invest, once he has signed the contract, and let $p^*_{|X=x}$ his conditional investment probability, when the state of nature is $x \in M$. We then have

$$p^*_{|X=x} = \frac{q^* e^{u[\boldsymbol{w}(x)]/c_a}}{q^* e^{u(\boldsymbol{w}(x))/c_a} + (1 - q^*) e^{u(y)/c_a}} \, , \tag{16}$$

where

$$q^* \in \arg\max_{q \in [0,1]} \sum_{x \in M} \mu(x) \cdot \ln \left[q e^{u[\boldsymbol{w}(x)]/c_a} + (1 - q) e^{u(y)/c_a} \right] \, , \tag{17}$$

and, if $q^* \in (0, 1)$,

$$\sum_{x \in M} \frac{\mu(x) \cdot e^{u[\boldsymbol{w}(x)]/c_a}}{q^* e^{u[\boldsymbol{w}(x)]/c_a} + (1 - q^*) e^{u(y)/c_a}} = 1. \tag{18}$$

Just as for the investor in autarky, the agent's investment conditions may be "sour", "normal" or "sweet". Necessary and sufficient conditions for each of these three cases parallel those for the investor in autarky:

Corollary 2. *Investment conditions are normal for the agent if*

$$-c_a \ln \mathbb{E}\left[e^{-u[\boldsymbol{w}(X)]/c_a}\right] < u(y) < c_a \ln \mathbb{E}\left[e^{u[\boldsymbol{w}(X)]/c_a}\right], \tag{19}$$

sour if

$$u(y) \geq c_a \ln \mathbb{E}\left[e^{u[\boldsymbol{w}(X)]/c_a}\right], \tag{20}$$

and sweet if

$$u(y) \leq -c_a \mathbb{E}\left[e^{-u[\boldsymbol{w}(X)]/c_a}\right]. \tag{21}$$

In other words, once employed, the agent will make no effort to acquire information and not invest if his salary is too high, as expressed by inequality (20). He will also make no effort to acquire information, but nevertheless invest, if his salary is too low, as expressed in (21). For intermediate salaries, those that satisfy (19), he will acquire some information and thereafter invest if and only if the obtained information is sufficiently favorable for investment, for him personally under his contract. We finally note that the agent's behavior, once employed by the principal, is uniquely determined by his (positive or negative) utility gain from investing, $u[\boldsymbol{w}(x)] - u(y)$, under each possible outcome $x \in M$.[11]

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. Second, the contract must be such that it provides normal investment conditions for the agent (otherwise the principal is better off not hiring the agent). Third, if there are contracts that meet these two requirements, then such a contract should yield the highest possible expected profit to the principal among these. Fourth and finally, the principal's expected profit from hiring the agent should (weakly) exceed the expected profit from not hiring the agent. We will analyze these four conditions in turn.

For this purpose, we begin by noting that the agent's *ex ante* expected utility under any contract $\langle y, \boldsymbol{w} \rangle$ can be written as a convex combination **of** the utility from the salary (weighted by the *ex ante* probability that the agent will not invest) and the utility from payment after investment (weighted by the ex ante probability that the agent will invest), net of the agent's information cost:

$$\begin{aligned} U(y, \boldsymbol{w}) \quad &= \quad (1 - q^*)\, u(y) \; + \; q^* \sum_{x \in M} \frac{\mu(x)\, u[\boldsymbol{w}(x)]}{q^* + (1 - q^*)\, e^{(u(y) - u[\boldsymbol{w}(x)])/c_a}} \\ &\quad - c_a \cdot \left(H(q^*) - \mathbb{E}\left[H\left(p^*_{|X} \right) \right] \right) \end{aligned} \tag{22}$$

The first condition mentioned above, to meet the agent's participation constraint, is simply

$$U(y, \boldsymbol{w}) \geq \bar{u}. \tag{23}$$

The second condition, that the agent' investment conditions should be normal, is precisely (19). In essence, this condition requires that the salary, and bonuses and penalties should be well balanced. In particular, the salary should be neither too low nor too high, and there should be enough penalties and bonuses to incentivize the agent to become well-informed before making the investment decision.

In order to pin down the third and fourth conditions, we first need to express the principal's *ex ante* expected profit from hiring the agent under any contract $\langle y, \boldsymbol{w} \rangle$.

[11] This is evident after some algebraic manipulation of all equations and inequalities above, which shows that all that matters are the two quantities $\mathbb{E}\left[e^{[u(\boldsymbol{w}(X)) - u(y)]/c_a}\right]$ and $\mathbb{E}\left[e^{-[u(\boldsymbol{w}(X)) - u(y)]/c_a}\right]$.

This expected profit is the convex combination of two terms, where the first term has weight $1 - q^*$, the *ex ante* probability that the agent will not invest, and represents the principal's "net savings" in this case, $Ir - y$, the interest payment minus the agent's salary. The second term has weight q^*, the ex ante probability that the agent will invest, and represents the principal's expected "net earnings", in this case. For each possible net return rate $x \in M$, the principal earns $xI - w(x)$, the difference between the net return from the investment and the pay to the agent. Formally:

$$\Pi(y, \boldsymbol{w}) \quad = \quad (1 - q^*)(rI - y) \quad + \quad q^* \sum_{x \in M} \frac{\mu(x) \cdot [xI - \boldsymbol{w}(x)]}{q^* + (1 - q^*) e^{(u(y) - u[\boldsymbol{w}(x)])/c_a}}. \tag{24}$$

If the contract provides normal investment conditions for the agent, $0 < q^* < 1$, then q^* is uniquely determined by (18). If instead the agent's investment conditions are sour under the contract, then he acquires no information and does not invests. In this case (24) gives $\mathbb{E}[\Pi(y, \boldsymbol{w})] = rI - y$. Likewise, if the contract turns the agent's investment sweet, then the agent invests "blindly", resulting in $\mathbb{E}[\Pi(y, \boldsymbol{w})] = I \cdot \mathbb{E}[X - \boldsymbol{w}(X)]$. Clearly, it is unprofitable for the principal to propose the agent a contract under which he collects no information, since the principal would earn at least as much by making the investment decision herself without any information acquisition; $rI \geq rI - y$ and $\mathbb{E}[X] \geq \mathbb{E}[X - \boldsymbol{w}(X)]$.

We can now state the third condition, that the contract should be optimal for the principal among all feasible contracts, if any, that meet the agent's participation constraint. Formally, this condition can be summarized as the requirement that the contract, $\langle y^*, \boldsymbol{w}^* \rangle$, should be a solution of the program

$$\max_{\langle y, \boldsymbol{w} \rangle \text{ s.t. } (23) \;\&\; (19)} \Pi(y, \boldsymbol{w}) \tag{25}$$

This brings us to the fourth condition, that all of this should be worthwhile for the principal. Suppose, thus, that $\langle y^*, \boldsymbol{w}^* \rangle$ solves program (25) and results in profit $\Pi(y^*, \boldsymbol{w}^*)$. If the principal has unit information cost $c > c_a$ in autarky, then it is profitable for her to hire the agent under the said contract if and only if $\Pi(y^*, \boldsymbol{w}^*) > \hat{\Pi}(I, c)$. If the principal's unit information cost, c, is so high that she would in autarky be in a sour investment condition, then $\hat{\Pi}(I, c) = I \cdot \max\{r, \mathbb{E}[X]\}$. We will henceforth focus on the last case, that is, situations in which the principal's unit cost of information acquisition is prohibitively high, so that the principal's alternative to hiring the agent is to decide single-handedly on investment without acquiring any information.[12]

[12]When the principal's information cost is relatively low, there is the additional issue whether the principal can commit not to acquire information alongside the agent. A complex issue that we do not enter.

4. Risk-neutral agent

Suppose, first, that the agent is risk neutral (risk averse agents are analyzed in Section 5). The principal's expected profit from any contract $\langle y, \boldsymbol{w} \rangle$ can then be written in the form

$$\Pi(y, \boldsymbol{w}) \quad = \quad q^* \cdot \sum_{i=1}^{m} \frac{\mu_i \left[(x_i - r) I + y - w_i \right]}{q^* + (1 - q^*) e^{(y - w_i)/c_a}} \quad - \quad y \tag{26}$$

where $\mu_i = \mu(x_i)$ and $w_i = \boldsymbol{w}(x_i)$ for each state of nature i. Under normal investment conditions for the agent, $0 < q^* < 1$ and

$$\sum_{i=1}^{m} \frac{\mu_i}{q^* + (1 - q^*) e^{(y - w_i)/c_a}} = 1. \tag{27}$$

The agent's behavior, once hired, is thus driven entirely by the *net transfers* in the different states of nature, the differences $t_i = w_i - y$. The vector of net transfers, $t = (t_1, ..., t_m)$, uniquely determines q^*, the *ex ante* probability that the agent will invest, according to equation (27), and, given q^*, this in turn determines all conditional investment probabilities. The agent's expected utility under a contract $\langle y, \boldsymbol{w} \rangle$ takes the form

$$\begin{aligned} U(y, \boldsymbol{w}) \quad = \quad & y \; + \; \sum_{i=1}^{m} \frac{q^* \mu_i \cdot (w_i - y)}{q^* + (1 - q^*) e^{(y - w_i)/c_a}} \; - \; c_a H(q^*) \\ & + c_a \sum_{i=1}^{m} \mu_i H\left(\frac{q^*}{q^* + (1 - q^*) e^{(y - w_i)/c_a}} \right) \end{aligned} \tag{28}$$

An application of the Karush-Kuhn-Tucker theorem leads to the following result:

Proposition 2. *If a contract $\langle y, \boldsymbol{w} \rangle$ solves (25), then $q^* \in (0, 1)$ is uniquely determined by (27). Moreover, for generic parameter values*

$$I x_i - w_i - c_a \frac{q^*}{1 - q^*} e^{(w_i - y)/c_a} = I x_j - w_j - c_a \frac{q^*}{1 - q^*} e^{(w_j - y)/c_a} \tag{29}$$

for all states of nature i and j where $w_i, w_j > 0$.

Each side of (29) is strictly decreasing in the payment to the agent after investment, and strictly increasing in the agent's salary. Hence, if the payments to the agent after investment in any two states are positive, then the pay will be higher in the state with a higher net return. Consequently, any optimal contract $\langle y, \boldsymbol{w} \rangle$ under which the agent's participation constraint is not binding has the following monotonicity property: *either* all payments are positive and $0 < w_1 < w_2 < ... < w_m$ *or* $w_1 = w_2 = ... = w_k = 0$ for some positive integer $k < m$, and $0 < w_{k+1} < w_{k+2} < ... < w_m$. Is the first case possible? The answer is "no" if the agent's participation constraint is slack. The reason can be seen in the definition (26) of the principal's expected profit. Suppose that $w_1 > 0$. Then $y > w_1$, since otherwise the agent would invest blindly. We may thus subtract some small amount ε (where $0 < \varepsilon < w_1$) from all payments w_i and also from the salary, y. This does not affect the first term in (26), since the agent's behavior, once hired, is driven entirely by the net transfers, $t_i = w_i - y$. Nor does it violate the agent's participation constraint if $\varepsilon > 0$ is small enough. However, such a subtraction increases the principal's expected profit (by precisely $\varepsilon > 0$), contradicting the hypothesis that the contract contract $\langle y, \boldsymbol{w} \rangle$ is optimal. In sum:

Corollary 3. *If a contract $\langle y, \boldsymbol{w} \rangle$ is optimal and the agent's participation constraint is slack, then there exists a positive integer $k < m$ such that*

$$w_1 = ... = w_k = 0 < w_{k+1} < ... < w_m. \tag{30}$$

Moreover, irrespective of whether the agent's participation constraint is binding or not, the pay schedule $\boldsymbol{w}$ is monotonic in the realized return rate x. Indeed, it follows from (29) that the pay schedule takes a particular form that can be expressed in terms of the so-called Lambert-W function used in physics and biochemistry. (This function is uniquely defined for positive z by the equation $W(z)\, e^{W(z)} = z$, see e.g. Corless et al., 1996).

Corollary 4. *If a contract $\langle y, \boldsymbol{w} \rangle$ is optimal, then $\boldsymbol{w}(x) = \max\{0, g(x)\}$ for all $x \in M$, where, for generic parameter values,*

$$g(x) = I(x - r) - c_a W\left(Ae^{Ix/c_a}\right) - B, \tag{31}$$

for constants $A > 0$ and $B \in \mathbb{R}$. The function g is continuous, strictly increasing and strictly concave.

The result in Corollary 4 is illustrated in Figure 1 below. The diagram shows a contract for a project that requires unit investment, $I = 1$, and that has a symmetric return rate distribution around zero, so its ex ante expected value is zero, $\mathbb{E}[X] = 0$. The interest rate r is zero. The agent's reservation wage is $\bar{u} = 0.1$ (or lower), and is not shown in the diagram. The horizontal line is the salary, here $y = 0.11$, and the kinked (red) curve is the payment schedule w. Hence, the agent has to pay a penalty if he invests and the realized return rate x is below approximately 12.5% (indicated by the right-most dashed vertical line), and he receives a bonus if the realized return rate x is higher. His payment after investment indeed is strictly increasing and strictly concave in the realized return rate for all return rates x above approximately -3.4% (indicated by the left-most dashed vertical line), below which his penalty is maximal, that is, his whole salary is withheld. The upward sloping (blue) line is the principal's gross return from investment, Ix. The principal's realized net profit when the agent invests is thus the vertical distance between this straight line and the curve representing the payment schedule w to the agent. Hence, the principal's realized profit after investment is negative if and only if the realized gross return rate x is below approximately $+9.5\%$, indicated by a dashed vertical line. The diagram is consistent with the optimal contract reported in the second row in Table 1 in the next subsection.[13]

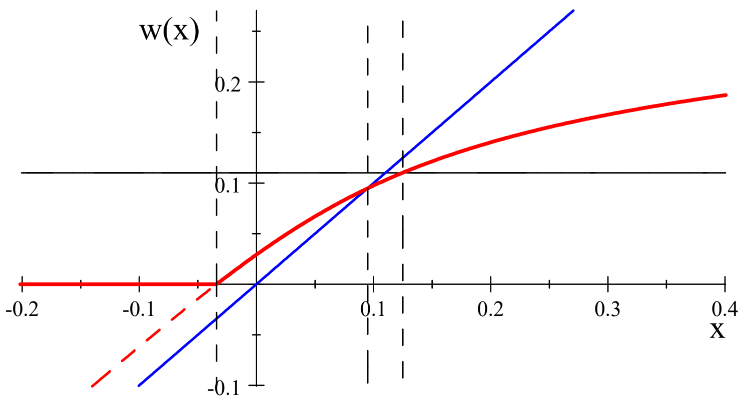

Figure 1: An example of an optimal contract.

Remark 1. *Since the payment schedule* $\mathbf{w}$ *is strictly concave wherever positive, affine contracts, such as a CEO contract composed of a fixed salary and some stocks, are suboptimal—even when the agent is risk neutral. The reason is limited liability. In the absence of this constraint, it would be optimal for the principal to sell (or rent) the whole asset in question to the agent (franchising). However, under limited liability the agent cannot be required to make a net payment to the principal. Suppose, thus, that a contract stipulates a fixed salary* y *and, in addition, a share* $\theta \in (0,1)$ *of the realized net gain,* $(x-r)\,I$*, from investment if made. Thus,*

$$w_i = y + \theta\,(x_i - r)\,I \quad \text{for } i = 1, ..., m,$$

[13] Those numbers define the optimal contract when $I = 1$, $r = 0$, $m = 2$, $r = 0$, $x_1 = -1$, $x_2 = 1$, $c_a = 0.05$, and $\bar{u} < 0.11$.

where the second term may be positive or negative, depending on the realized return rate. Limited liability thus amounts to the constraint

$$\theta \leq \frac{y}{(r - x_1)\, I}$$

(where $r > x_1$ by hypothesis). The lower the salary and the larger the investment, the tighter is this constraint. Information being a public good for the two parties, where the agent takes the whole cost but only reaps part of the benefit, the agent will under-invest in information acquisition, in comparison with first-best. The principal may thus want to boost the agent's incentive for acquisition beyond what the proposed affine contracts admit. This cannot be done by increasing penalties (after unprofitable investments), since this would break the limited-liability constraint, but can be done by means of enhanced bonuses (after profitable investments). In practice, this may be obtained by offering a three-part contract: a fixed salary, a share of the realized net gain, and, in addition, an option to purchase an additional share of the realized net gain at a predetermined price.

4.1. Comparative statics. The agent's expected utility if hired, may be *lower* the *more* able the agent is (with high ability defined as low unit information cost). This may first appear counter-intuitive. However, in order to obtain information about the project at hand, a more able agent needs to exert less effort and so may need to be paid less than a less able agent in order to be incentivized. In particular, smaller bonuses and penalties may be needed for a more able agent. This possibility is shown in Table 1 below. In this numerical example, the interest rate is zero and the project is a double-or nothing lottery, with equal probability for both outcomes: $r = 0$, $I = 1$, $m = 2$, $x_1 = -1$, $x_2 = 1$, and $\mu = 0.5$. Hence, $\mathbb{E}[X] = 0$, so the principal's expected profit in autarky is zero. The optimal contracts, conditional upon hiring a more or less able agent, are given below, under the presumption that no agents' participation constraint is binding. As before, "salary" refers to y, "bonus" refers to the difference $b = w_2 - y$, and "penalty" to the difference $p = y - w_1$.

TABLE 1

c_a	*salary*	*bonus*	*penalty*	*expected profit*	*expected utility*
0.10	0.141	0.158	0.141	0.122	0.171
0.05	0.110	0.129	0.110	0.250	0.144
0.01	0.039	0.045	0.039	0.424	0.055

We see that the principal's expected profit increases with the agent's increased ability, while the agent's expected utility decreases. Social welfare, defined as the sum of the two, is increasing with the agent's ability. These calculations were based on the premise that the agent's participation constraints are slack. How do the optimal contracts change if this is not the case? As seen in (26) and (27), the agent's behavior under a contract depends only on the bonus and penalty, not the salary. Hence, the optimal contact under a binding participation constraint is obtained from the above contracts by keeping the bonus and penalty constant but raising the salary, until the expected utility equals the value of the agent's outside option, granted this new contract is also profitable for the principal.

We also note the slight asymmetry, despite the project's symmetry, in the optimal contracts in Table 1; the bonuses are higher than the penalties. This asymmetry is in fact general, and can be seen in Corollary 4. The reason is that, although any symmetric contract would perfectly align the agent's incentives with those of the principal at the moment of investment (had the principal had access to the agent's private information about the project at hand), such contracts will not provide enough incentive for the agent's information acquisition. The agent carries all the burden of information acquisition but reaps only part of the benefit of enhanced information. As the bonus and penalty are increased in order to enhance the agent's incentive for information acquisition, the penalty will sooner or later hit the limited-liability constraint. Any increase of the penalty is therefore more costly for the principal than the same increase of the bonus. This is because an increase of the penalty can only be made by raising the salary, and if the bonus is to be kept constant, also the payment after a successful investment needs to be raised.

Figure 2, below (drawn for the same parameters as in Table 1), shows indifference curves for both parties in contract space.

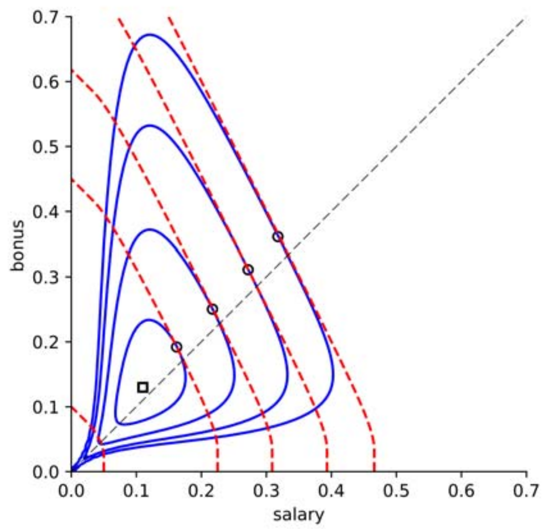

Figure 2: Iso-profit curves and iso-utility curves in contract space.

Here the salary y (which is equal to the penalty $p = y - w_1$ in the present binary case since $w_1 = 0$) is on the horizontal axis and the bonus, $b = w_2 - y$, on the vertical. The principal's iso-profit curves are the solid closed loops, while the agent's iso-utility curves are the dashed negatively sloped curves. The principal's global optimum—that is, in the absence of the agent's participation constraint—is indicated by the small square at approximately, $y^* \approx 0.11$ and $b^* \approx 0.13$, where the principal's expected profit is approximately 0.25 and the agent's expected utility approximately 0.14. Hence, if the agent's reservation utility is below the latter level, then his participation constraint is slack. The largest iso-profit contour in the diagram corresponds to zero expected profit, the profit the principal would earn in autarky.

The diagram also shows that the agent's expected utility is *not* quasi-concave in contract space; the non-convex upper contour sets for his expected utility is the areas above the dashed curves. Despite this, the principal's maximization program (25) has unique solutions, indicated by little circles, for each of the different levels of his reservation utility $\bar{u}$. All optimal contracts in this example includes a higher bonus than penalty—they lie above the diagonal—although the project in question is perfectly symmetric. This asymmetry of the optimal contracts, discussed in general above, implies that the agent, at the investment decision after information has been acquired, is willing to take a little more risk than the investor would then like. In particular, the agent may invest even when his private information is not sufficient to make investment profitable for the investor.

Remark 2. *The present analysis is premised on the assumption that the principal knows the agent's type (risk attitude, information costs, and reservation utility). What can be said in cases of incomplete information? Suppose that there are three types of agent, those in Table 1, and assume that the principal does not know what type a given agent has. The principal would prefer to hire the most able type of agent, granted such agents' reservation utilities are not too high. However, agents do not necessarily want to appear more able than they are, since their salary and bonus would then be lower (as would their penalty). For instance, a low-ability agent* ($c_a = 0.10$)*, if hired under the contract for the high-ability type* ($c_a = 0.01$)*, would obtain expected utility* 0.043*, and a medium-ability agent* ($c_a = 0.05$) *under that contract would obtain expected utility* 0.045*; in both cases far below their proper contracts. In order for the principal to let agents of unknown ability self-select contracts, reservation utilities will matter.*

4.2. Almost perfectly informed agent. Let us briefly consider contracts for risk-neutral agent with extremely low unit costs of information. It is easily verified from Corollary 4 that the optimal contract then is approximately piece-wise linear, see Figure 3 below. This diagram has been drawn for the same data as used in Figure 2, except for the agent's unit information cost, which has been reduced by a factor 25 (from $c_a = 0.05$ to $c_a = 0.002$). The kinked and piece-wise almost linear curve is the associated optimal pay scheme. The agent, who has very precise information about the actual return rate, will invest almost surely if and only if the return rate exceeds the interest rate (zero in this example), since this is precisely when his pay $w(x)$ from investing exceeds or equals y, his pay if not investing. For return rates below -11% (indicated by a dashed vertical line) the penalty is maximal, while for higher, but still negative return rates, the pay grows linearly with the return rate. For high return rates the bonus is small and virtually independent of the return rate. The principal's gross return rate, x, is indicated by the (blue) upward-sloping straight line. The principal's profit when the agents invests is thus the vertical difference between this straight line and the pay schedule. We see in the diagram that the principal (almost surely) makes a positive profit in those states of nature where the realized return rate exceeds approximately $+11.5\%$ (indicated by a dashed vertical line).

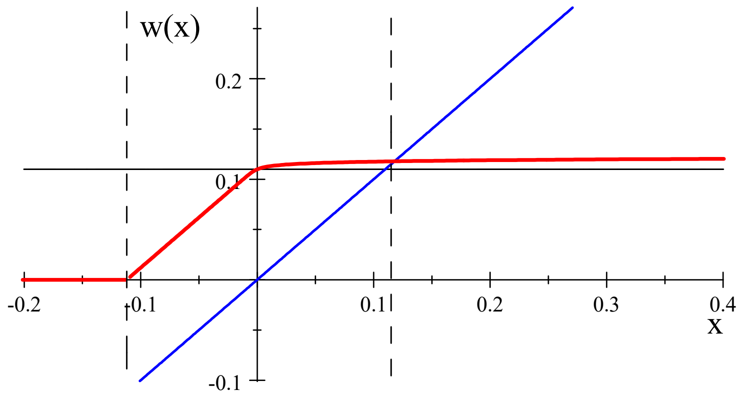

Figure 3: An optimal contract for an agent with very low information costs.

More generally and precisely:

Corollary 5. *Consider a sequence of risk neutral agents with smaller and smaller unit cost $c_a > 0$ of information. Let $\bar{u} > 0$ be the reservation wage of the limit agent with zero information cost. As $c_a \to 0$, the optimal contracts tend (point-wise) to the limit contract $\langle y_0, \boldsymbol{w}_0 \rangle$, where $y_0 = \bar{u}$ and*

$$w_0(x) = \begin{cases} 0 & \text{if } x < r - \bar{u}/I \\ I(x-r) + \bar{u} & \text{if } r - \bar{u}/I \leq x \leq r \\ y & \text{if } x > r \end{cases} . \tag{32}$$

In order to get an intuition for this result, consider first an agent who has no cost at all of information acquisition (a situation that falls outside the domain of our model since we assume $c_a > 0$). Such an idealized agent knows the true state of the world. For this agent it is optimal to invest under a contract $\langle y, \boldsymbol{w} \rangle$ if and only if $w(x) \geq y$. The principal, knowing this, does not have to pay the agent more than his salary, y, in order to induce him to invest. Hence, it is optimal for the principal, who wants to economize on her payment to the agent, to pay the agent $w(x) = y$ if he invests when $x \geq r$, and any amount $w(x) < y$ if he invests when $x < r$. The exact payment $w(x)$ in the latter case is irrelevant, since the agent, knowing x, will anyhow not invest. However, as claimed in the above corollary, in the limit as the agent's unit information cost is positive but tends to zero, the payment scheme takes the above piece-wise linear form.

Hence, our idealized agent, with zero cost of information acquisition, earns y in all states of nature. The principal hence chooses y as low as possible, that is, such that the agent's participation constraint is just met, or $y = \bar{u}$. The agent's utility is thus $\bar{u}$, while the principal's profit is random, with mathematical expectation

$$\Pi^o = \mathbb{E}\left[(X - r) I \mid X > r\right] - \bar{u}.$$

It is thus optimal for the principal to hire the agent if and only if Π^o is no lower that her profit in autarky, that is, if and only if

$$\mathbb{E}\left[X \mid X > r\right] - \bar{u}/I \quad \geq \quad \max\left\{r, \mathbb{E}\left[X\right]\right\}. \tag{33}$$

This gains-from-trade condition is more easily met the larger is the investment I and the worse is the agent's outside option, $\bar{u}$.

We finally note that in the limit case of a virtually costlessly informed agent, the optimal contract (32) is socially efficient under condition (33); the agent is then hired and investment takes place if and only if its net return rate, x, exceeds the interest rate, r, and the agent receives income y in all states of the world. If condition (33) is violated, then no trade between the two parties takes place. Instead, the principal will then make the investment decision herself on the basis of her prior belief about the state of the world.

Remark 3. *If the principal's unit cost of information, c, is not much higher than that of the agent, c_a, then the right-hand side in condition (33) should be replaced by $\hat{\Pi}(I, c)$, the principal's expected profit from her optimal information acquisition and investment decision in autarky.*[14]

[14] In this case, when the investor can also acquire information, it is important that she, under delegation to an agent, commits not to interfer in the agent's work or decsion-making.

5. RISK-AVERSE AGENT

Having considered risk-neutral agents we now turn to risk-averse agents. More specifically, the agent's Bernoulli function $u : \mathbb{R}_+ \to \mathbb{R}$ is now assumed to be continuous, strictly increasing and concave, and to be twice differentiable on $\mathbb{R}_{++}$with $u' > 0$, $u'' \leq 0$, and $\lim_{z \to 0} u'(z) = +\infty$. The following result is derived in the Appendix:

Proposition 3. *If a contract $\langle y, \boldsymbol{w} \rangle$ is optimal and the agent's participation constraint is not binding, then $\boldsymbol{w}(x) = \max\{0, h(x)\}$ for all $x \in M$, where $h : [P, +\infty) \to \mathbb{R}$, for any given $y > 0$, is implicitly defined by*

$$Ix = h(x) + \frac{c_a}{u'(h(x))} \cdot \left(1 + Qe^{[u(h(x)) - u(y)]/c_a}\right) + P, \tag{34}$$

for constants $Q > 0$ and $P \in \mathbb{R}$. The function h is continuous and strictly increasing. Moreover, $Q = q^* / (1 - q^*)$, where $q^* \in (0, 1)$ is the unique solution to (18).

It turns out that the optimal payment scheme, $\boldsymbol{w}$, is qualitatively similar to that for a risk-neutral agent with very low information costs; the bonus is small for profitable investments, and it tapers off almost linearly for mediocre return rates. But now it is for a different reason; it is not because the agent is very well informed but because he is risk averse. It is then not worthwhile to pay him a high bonus since his utility gain from increased remuneration is less than his utility loss from an equally big decrease in remuneration. While a risk neutral agent with very low information costs faces almost no risk—he almost surely earns close to his salary y in all states of nature—this is not the case for a risk-averse agent with high information costs. Such an agent will sometimes make mistakes and invest when he should not have done so.

See Figure 4 below, drawn for an agent with constant relative risk aversion $\rho = 0.8$ (solid curve) and for an agent with constant relative risk aversion $\rho = 1.6$ (dashed curve), and otherwise the same parameters as before. In this example, the principal makes a net profit if the net return rate x from investment exceeds approximately $+13.3\%$ for the less risk-averse agent and $+11.5\%$ for the more risk-averse agent.

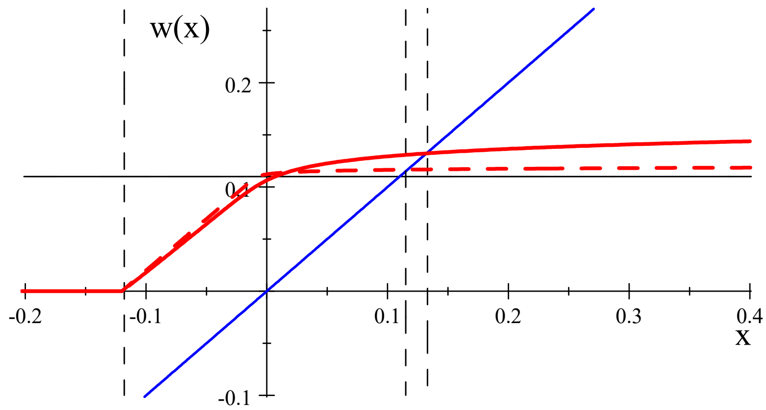

Figure 4: A contract for a more (dashed) or less (solid) risk-averse agent.

More generally, for an agent with arbitrary constant relative risk aversion $\rho > 0$, equation (34) takes the form

$$Ix = h(x) + c_a h(x)^\rho \cdot \left(1 + Q \exp\left[\frac{h(x)^{1-\rho} - y^{1-\rho}}{(1-\rho)\, c_a}\right]\right) + P.$$

where $Q > 0$ satisfies

$$\sum_{x \in M} \frac{\mu(x) \cdot (1+Q) \exp\left[\frac{h(x)^{1-\rho} - y^{1-\rho}}{(1-\rho) c_a}\right]}{1 + Q \exp\left[\frac{h(x)^{1-\rho} - y^{1-\rho}}{(1-\rho) c_a}\right]} = 1.$$

The last equation, being a version of (18), expresses the consistency requirement that the ex ante expected value of the posterior investment probability (after information has been acquired) equals the prior investment probability.

6. A canonical signal-extraction model

Are the qualitative results obtained above specific to the rational-inattention model of information acquisition? In order to investigate this we here compared the results with those obtained from a canonical model of signal-extraction with endogenous signal precision. A more complete analysis of this alternative model of information acquisition is provided in Lindbeck and Weibull (2015).[15] We here briefly discuss this alternative model for the case of binary return distributions, normally distributed noise, and a risk neutral agent. We begin by quickly considering an investor in autarky and then study delegation to an agent.

[15] For a decision-theoretic signal-extraction model with endogeneous signal precision, see Weibull, Mattsson, and Voorneveld (2007).

6.1. Investor in autarky. We normalize the investment to unity, $I = 1$, and focus on a binary return distribution, with net return rates $x_1 < 0 < x_2$, and zero interest rate ($r = 0$). Consider, first, an investor who acts in autarky. If she chooses to acquire information, she will receive a noisy signal S about the net return rate,

$$S = X + \varepsilon, \tag{35}$$

where the noise term ε is statistically independent of X, and is normally distributed with mean zero and variance $1/\tau > 0$. We call $\tau > 0$ the *signal precision*, and this is chosen by the decision maker. Hence, the decision maker can, if she wants, affect the noisiness of the signal. Let $C(\tau)$ be the investor's cost or disutility of obtaining signal precision $\tau > 0$, where C is a continuous, strictly increasing and convex function.

The investor's *profit*, if she decides to invest, is the (random) net return minus the information cost (which she has to pay even if she subsequently decides not to invest). Under what conditions will the investor decide to acquire information? If she decides to acquire information, what signal precision $\tau > 0$ will she choose? Once her information has arrived, what investment decision will she then make? Will the results be qualitatively similar to those obtained in the rational-inattention model? We answer these questions in reversed order.

Suppose, thus, that the investor has obtained information, the realization $s \in \mathbb{R}$ of the signal S, with precision $\tau > 0$. Her information costs being sunk, it is now optimal for her to invest if and only if $\mathbb{E}[X \mid S = s] \geq r = 0$. Since the signal structure (35) meets the monotone likelihood ratio property (MLRP), this inequality can be written equivalently as

$$S \geq \hat{s}(\tau) \tag{36}$$

for some $\hat{s}(\tau) \in \mathbb{R}$; the optimal *signal threshold*.[16] This threshold defines an optimal *investment strategy*, for any signal precision $\tau > 0$ that the investor may have chosen.

We now take one step backwards in time and consider the investor's information acquisition decision. Anticipating that she will use her optimal signal threshold thereafter, she chooses her signal precision $\tau \geq 0$ so that it maximizes her *ex ante* expected profit

$$\tilde{\Pi}(\tau) = V(\tau) - C(\tau), \tag{37}$$

where the first term is the expected net return from using the optimal investment strategy (given τ):

$$V(\tau) \;=\; I \cdot \mathbb{E}[X \mid S > \hat{s}(\tau)] \cdot \Pr[S \geq \hat{s}(\tau)] \tag{38}$$

[16] Even under the MLRP, $\hat{s}(\tau)$ may be infinite, in which case $\hat{s}(\tau) = -\infty$ means "always invest" and $\hat{s}(\tau) = +\infty$ "never invest". However, because of its thin tails, the normal distribution has $\hat{s}(\tau) \in \mathbb{R}$ for all $\tau > 0$, see Lindbeck and Weibull (2017).

It is the product of three factors: the size of the investment, the conditionally expected net return if investment is made, and the probability for investment (recall that the opportunity cost, the interest rate, is here normalized to zero). The choice $\tau = 0$ represents the case of no information acquisition (or, equivalently, a costless uninformative signal).

If the investor chooses not to acquire any information, $\tau = 0$, then it is optimal for her to invest in the project if and only if its *a priori* expected return rate is nonnegative (since $r = 0$). By contrast, if she chooses to acquire information, $\tau > 0$, then the chosen precision τ has to meet the first-order condition that its marginal financial value equals its marginal cost,

$$V'(\tau) = C'(\tau). \tag{39}$$

We will write $\mu \in 0, 1$ for her prior probability for the good state, $x = +1$ (just as in the binary example studied in the rational inattention model). It can be shown that the investor's optimal signal threshold for any $\tau > 0$ then is

$$\hat{s}(\tau) = \frac{x_2 + x_1}{2} - \frac{\ln \hat{\rho}}{(x_2 - x_1)\tau}, \tag{40}$$

where

$$\hat{\rho} = \frac{\mu}{1-\mu} \cdot \frac{x_2}{|x_1|}, \tag{41}$$

a parameter that we call the *risk balance* of the project, being the ratio between the project's "expected upside", θx_2, and "expected downside", $(1-\mu)|x_1|$. Equation (40) tells us that the more favorable the risk balance, the lower the investor's optimal signal threshold, that is the wider is the range of signals for which she is willing to invest. Moreover, the project is *a priori* profitable (unprofitable) if the risk balance exceeds (falls short of) unity; $\hat{\rho} > 1 \Leftrightarrow \mathbb{E}[X] > r = 0$.

It can also be shown that the marginal value of information satisfies

$$V'(\tau) = \frac{x_2 - x_1}{2} \cdot \frac{I \cdot \hat{\kappa}}{\sqrt{2\pi\tau}} \cdot \exp\left[-\frac{1}{2\tau}\left(\frac{\ln \hat{\rho}}{x_2 - x_1}\right)^2 - \frac{\tau}{2}\left(\frac{x_2 - x_1}{2}\right)^2\right], \tag{42}$$

where

$$\hat{\kappa} = I\sqrt{\mu(1-\mu)x_2|x_1|}, \tag{43}$$

a quantity we refer to as the *riskiness* of the project. The marginal value of information is thus positive at all positive signal precisions. However, when $\hat{\rho} \neq 1$ it tends to zero as the signal precision either tends to zero or to plus infinity. A poorly informed investor is not much helped by a small bit of information, and very well informed investor does not benefit much by a small piece of additional information. Hence, the marginal value of information is non-monotonic with respect to signal precision in the generic case when $\hat{\rho} \neq 1$. More precisely , if $\hat{\rho} \neq 1$, then $V'(\tau) > 0$ for all $\tau > 0$, $\lim_{\tau \to +\infty} V'(\tau) = 0$, and $\lim_{\tau \to 0} V'(\tau) = 0$. By contrast, if $\hat{\rho} = 1$, then still $V'(\tau) > 0$ for all $\tau > 0$, and $\lim_{\tau \to +\infty} V'(\tau) = 0$, but now $\lim_{\tau \to 0} V'(\tau) = +\infty$. Hence, in this knife-edge case, but only then, the marginal value of information is infinite at zero signal precision and declines monotonically towards zero as signal precision rises.

Moreover, we see in (42) that, at any given positive signal precision, the marginal value of information is increasing in the riskiness $\hat{\kappa}$ of the project, but non-monotonic in its risk balance $\hat{\rho}$. At any given signal precision, the marginal value of information is largest for projects with unit risk balance. This holds precisely when the uninformed investor is indifferent between investing and not investing (when $\mathbb{E}[X] = 0$). The more the risk balance deviates from unity, the lower is the marginal value of information for the investor.

In order to close the model, we need to specify the cost of information acquisition. We consider a certain parametric form that turns out to be analytically convenient, namely cost functions of the form

$$C(\tau) = \gamma \cdot \int_0^\tau \frac{1}{\sqrt{t}} \exp\left(\alpha t - \frac{\beta}{t}\right) dt, \tag{44}$$

for $\alpha, \beta, \gamma > 0$. The associated marginal information cost, $C'(\tau)$, is positive at all positive signal precisions. It tends to zero as the signal precision either tends to zero or towards plus infinity.[17] For cost functions of the form (44), the optimal signal precision can be solved explicitly. Let

$$\hat{K} = \ln \hat{\kappa} + \ln\left(\frac{x_2 - x_1}{2}\right) - \ln\left(\gamma\sqrt{2\pi}\right), \tag{45}$$

a (positive or negative) quantity that is increasing in the riskiness of the project, $\hat{\kappa}$, and decreasing in the information cost parameter γ. Let

$$\hat{D} = \hat{K}^2 - \left[\left(\frac{x_2 - x_1}{2}\right)^2 + 2\alpha\right] \cdot \left[\left(\frac{\ln \hat{\rho}}{x_2 - x_1}\right)^2 - 2\beta\right] \tag{46}$$

This quantity is positive when the risk balance $\hat{\rho}$ is close to unity. When $\hat{D} > 0$, let

$$\hat{\tau} = \frac{\hat{K} + \sqrt{\hat{D}}}{(x_2 - x_1)^2/4 + 2\alpha}. \tag{47}$$

[17] For suitable parameter values and moderate ranges of signal precision, the graph of these cost functions are almost indistinguishable from those of more conventional cost functions, such as $C(\tau) = \tau^2$.

The following result combines the necessary first-order condition $V'(\tau) = C'(\tau)$ for a positive signal precision τ to be optimal with the also necessary requirement that this should result in a higher expected profit to the investor than choosing signal precision $\tau = 0$ (and then going for the best alternative use of her money).

Proposition 4. *Suppose that the cost function is of the form (44). If $\hat{D} > 0$, $\hat{\tau} > 0$, and $\Pi(\hat{\tau}) > I \cdot \max\{r, \mathbb{E}[X]\}$, then $\hat{\tau}$ is the optimal signal precision. If $\hat{D} \leq 0$, the optimal signal precision is zero.*

In other words, the condition $\hat{D} > 0$ is necessary, but not sufficient, for the investor to bother to acquire information. Suppose that $\hat{D}$ is positive. It is not difficult to verify that the investor's optimal signal precision is then positive and given by (47) if either $\hat{K} \geq 0$ or

$$\hat{K} < 0 \quad \text{and} \quad e^{-(x_2-x_1)\sqrt{2\beta}} < \hat{\rho} < e^{(x_2-x_1)\sqrt{2\beta}}. \tag{48}$$

Hence, there are two distinct conditions under which the investor will not acquire any information, just as in the rational-inattention model. Under certain investment conditions, the investor will not bother to acquire any information and will not invest. Under other investment conditions, she will invest "blindly", that is without bothering to acquire any information. Under intermediate investment conditions, the investor will acquire some information and invest if and only if the received signal is sufficiently favorable. This is the first point where we note a similarity with the rational-inattention model.

In force of Proposition 4, the three investment conditions can be precisely identified, a task to which we now turn in a special case, namely a double-or-nothing investment project. Let $I = 1$, $x_1 = -1$, and $x_2 = 1$. Then the *risk balance* is simply the odds ratio, $\hat{\rho} = \mu/(1-\mu)$, and the *riskiness* is the geometric mean of the probabilities for the two states of nature, $\hat{\kappa} = \sqrt{\mu(1-\mu)}$. Moreover, $K \geq 0$ if and only if $\gamma \leq \sqrt{\mu(1-\mu)/2\pi}$. This inequality holds if the information-cost parameter γ is small and/or the prior μ is close to one half. In particular, $\hat{K} < 0$ for all $\mu \in [0,1]$ if and only if $\gamma > 1/\sqrt{8\pi} \approx 0.2$. Suppose that γ meets this condition. Then $\hat{\tau} > 0$ if and only if $\ln[\mu/(1-\mu)] < \sqrt{8\beta}$. For $\beta = 1/2$, this is identical with the definition of normal investment conditions in the rational-inattention model when applied this example (for $c = 0.5$). The solid curve in Figure 5 below, is the investor's expected profit as a function of the prior μ for the good state of nature in the present signal-extraction model (for $\alpha = 0.1$ and $\beta = 0.5$, and $\gamma = 0.25$). The dashed curve is the investor's profit function in the rational-inattention model.

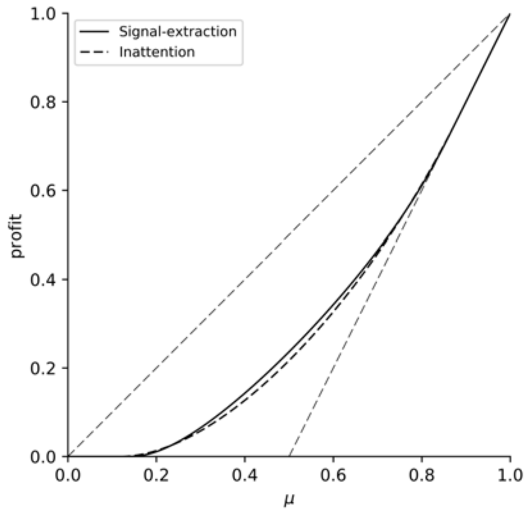

Figure 5: The investor's expected profit in the signal-extraction model (solid curve) and in the rational-inattention model (dashed curve).

In this example, the range of probabilities μ for which the investment conditions are normal is approximately $0.12 < \mu < 0.88$.[18] Not surprisingly, investor's expected profit curve is increasing in the prior μ for the good state of nature. Its horizontal line segment, for low μ, represents her expected profit, zero, in sour investment conditions ($\mu \lesssim 0.12$). For $\mu \gtrsim 0.88$, investment conditions are sweet, which results in the steep straight line part of the curve (the net return from blind investment). The dashed 45-degree line is the *ex ante* expected profit to a perfectly informed investor, who will invest if and only if the state of nature is good. By contrast, the expected profit to a completely uninformed investor is zero for all $\mu \leq 1/2$, since it is then optimal for her not to invest. It is $2\mu - 1$ for all $\mu > 1/2$, since it is then optimal for her to invest. This is the dashed steep straight line in the diagram.

6.2. Delegation. Suppose now that the investor considers the possibility of hiring an agent who has a comparative advantage in acquiring and processing relevant information about investment projects, just as in the rational inattention model. If hired by the principal, in this model the agent will choose his signal precision and make the investment decision accordingly. To be more specific, the agent, if employed,

[18] More exactly, the normality condition writes

$$\frac{1 - e^{-2}}{e^2 - e^{-2}} < \mu < \frac{e^2 - 1}{e^2 - e^{-2}}.$$

will either make no effort to acquire information or he will decide to acquire some information and receive a noisy *private signal* S_a about the net return rate of the project,

$$S_a = X + \varepsilon_a, \tag{49}$$

where the *noise* ε_a is statistically independent of X. This noise term is again normally distributed with mean zero, but now with variance $1/\tau_a > 0$, where $\tau_a > 0$ is the *signal precision* chosen by the agent.

A *contract* $\langle y, \boldsymbol{w} \rangle$ between the principal and agent takes the same form as in the rational-attention model. Just as in the rational-inattention model, the agent's utility is additively separable in utility from income and disutility from effort to gather information. His utility is here the random variable

$$\tilde{U} = u\left(\tilde{Y}\right) - C_a\left(\tau_a\right), \tag{50}$$

where the first term is his utility from his income $\tilde{Y}$, evaluated in terms of his Bernoulli function u for income, and the second term is his disutility from the effort needed to obtain signal precision $\tau_a \geq 0$. The agent's participation constraint is $\mathbb{E}\left[\tilde{U}\right] \geq \bar{u}$. The principal anticipates all of this, and will thus offer a contract that maximizes her expected profit—the residual that remains after the agent has been paid his due—among all contracts that meet the agent's participation constraint.

We solve the model backwards in time. Suppose, thus, that the agent has already signed the contract and made his effort to gather information about the project. For what signal realizations will he invest? Being rational and self-interested, he will base this decision on his expected remuneration utility under the given contract. The disutility of his effort to collect information is now bygone, as is his outside option. It is thus optimal for the agent to invest if and only if his conditionally expected utility from remuneration when investing, given his signal realization, exceeds his remuneration utility from not investing,

$$\mathbb{E}\left[\tilde{U} \mid S_a = s\right] \geq u\left(y\right). \tag{51}$$

Equality in this condition uniquely determines a signal threshold, $s^*\left(\boldsymbol{w}, \tau_a\right) \in \mathbb{R}$, above which it is optimal for the agent to invest and below which it is optimal for him to not invest. For any contract $\boldsymbol{w} \in W$ and any any positive signal precision τ_a that he may have chosen (by way of his effort to acquire information), this optimal investment threshold can be shown to be

$$s^*\left(\boldsymbol{w}, \tau_a\right) = \frac{x_2 + x_1}{2} - \frac{\ln \rho\left(\boldsymbol{w}\right)}{\left(x_2 - x_1\right)\tau_a}. \tag{52}$$

where

$$\rho(\boldsymbol{w}) = \frac{\mu}{1-\mu} \cdot \frac{u(w_2) - u(y)}{u(y) - u(w_1)}, \tag{53}$$

is what we call the *carrot-stick ratio* of the contract, a useful concept in this type of analysis (see Appendix). The latter is the probability-weighted ratio between the agent's two utility gains from "doing the right thing" in each state of nature (to invest in the good state and not invest in the bad). The carrot-stick ratio, $\rho(\boldsymbol{w})$, plays the same role for the agent as the risk balance, $\hat{\rho}$, played for the investor in autarky. We note that the carrot-stick ratio of a contract is independent of the agent's choice of signal precision.

We now take one step backwards in time, to the moment when the agent has signed a contract and is about to decide how much effort to gather information. It is not difficult to show that the agent will either make no effort at all, choose signal precision $\tau_a = 0$, or else make a positive effort and subsequently obtain a signal with positive precision, $\tau_a > 0$. In the latter case, the agent's optimal signal precision has to meet the following first-order condition:

$$\frac{x_2 - x_1}{2} \cdot \frac{\kappa(y, \boldsymbol{w})}{\sqrt{2\pi\tau_a}} \cdot \exp\left[-\frac{(x_2 - x_1)^2 \tau_a}{8} - \left(\frac{\ln \rho(\boldsymbol{w})}{x_2 - x_1}\right)^2 \frac{1}{2\tau_a}\right] = C_a'(\tau_a), \tag{54}$$

where

$$\kappa(y, \boldsymbol{w}) = \sqrt{\mu(1-\mu)(u(w_2) - u(y))(u(y) - u(w_1))} \tag{55}$$

We call $\kappa(y, \boldsymbol{w})$ the contract's *power*. It is increasing in the product of the agent's two utility gains from "doing the right thing" in each state of nature, and more steeply increasing the more uncertain the project is *ex ante*. In particular, this factor is maximal when both states of nature are equally likely. We also note that, like the carrot-stick ratio, the power of a contract is independent of the agent's signal precision. We also note that equation (54) is formally identical with the necessary first-order condition for the investor's choice of signal precision in autarky, (42). The only difference is that the role of $\hat{\kappa}$ is now played by $\kappa(y, \boldsymbol{w})$ and the role of $\hat{\rho}$ by $\kappa(y, \boldsymbol{w})$.

Under every contract $\langle y, \boldsymbol{w} \rangle$ that is worthwhile for the principal, the agent will choose a positive signal precision $\tau_a^*(y, \boldsymbol{w})$ and subsequently use his optimal signal threshold $s^*((y, \boldsymbol{w}), \tau_a)$, defined for all signal precisions $\tau_a > 0$ in (52). Anticipating this, the principal expects the profit

$$\tilde{\Pi}(y, \boldsymbol{w}) = I(x_2 - w_2 + y) \cdot p_G(y, \boldsymbol{w}) + I(x_1 - w_1 + y) \cdot p_B(y, \boldsymbol{w}) - y, \tag{56}$$

where $p_G(\boldsymbol{w})$ is the probability that the state of nature is good and the agent invests, and $p_B(\boldsymbol{w})$ is the probability that the state of nature is bad and the agent invests.

Suppose that the agent is risk neutral. Then the agent's signal threshold (once hired) is a function of the bonus and penalty only, and so is the agent's optimal signal precision. Hence, once hired, the agent's behavior is completely determined by the bonus and penalty. Like in the rational-inattention model, the salary is a pure transfer from the principal to the agent, with no behavioral consequences, given the bonus and penalty. In order to induce the agent to be better informed, though he has a comparative advantage, it is thus needed that the contract has a sufficient bonus and penalty. Given the moral hazard involved, do contracts exist that are profitable to the investor. The isoquants in Figure 6 below, solid for the principal, dashed for the agent, have been created for the same project as used in the rational-attention model (Figure 2).[19] Yet another qualitative similarity between the two models.

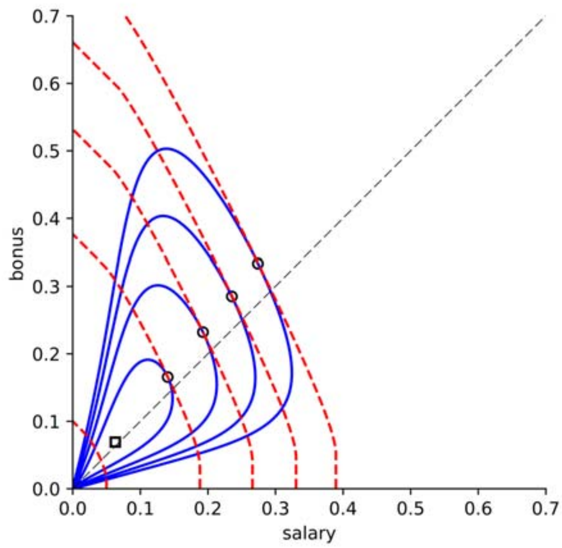

Figure 6: Iso-profit and iso-utility curves in contact space for the signal-extraction model.

7. Discussion

We have here developed a model of delegation of investment decisions, where the agent is endogenously informed about the profitability of investment projects. The agent's information acquisition was modelled within the rational-inattention framework for a single project with finitely many potential outcomes. The analysis can be readily extended to more general outcome distributions and to multiple investment options.

[19] The information-cost parameters for the agent are here $\gamma_a = 0.025$, $\alpha_a = 0.1$ and $\beta_a = 0.5$.

By way of comparing the results with results obtained in a signal-extraction model with endogenous signal precision, we found that, at least in some simple special cases, the results are robust to how information acquisition is modelled.

Our model of delegation builds upon many heroic simplifications. A relevant extension would be to allow for incomplete information about the agent's type, that is, his talent or ability (unit information cost), risk attitude (Bernoulli function), and/or outside option. Can a principal then use screening to let agents self-select contracts? This depends crucially upon agents' outside options and whether agents know their own type or not. It would be particularly risky to let agents self-select contracts if some of them would be overconfident in the sense of underestimating their unit costs of information acquisition or overestimating the quality or precision of their obtained information. Another interesting extension would be to consider agents with career concerns. In such settings, bonuses and penalties could be, at least partly, replaced by future career prospects. It would also be interesting, and arguably realistic, to consider agents who are partially motivated by social esteem, morality, or loyalty towards the principal. We hope that the present analysis can serve as a first step towards such richer and more realistic settings.

8. Appendix

We first establish that equation (4) is necessary for optimality. This claim agrees with Corollary 2 in Matejka and McKay (2015). Taking the derivative of the maximand in (17) with respect to q, one obtains

$$\sum_{x\in M}\frac{\mu(x)\cdot e^{xI/c}}{\hat{q}e^{xI/c}+(1-\hat{q})\,e^{rI/c}}=\sum_{x\in M}\frac{\mu(x)\cdot e^{rI/c}}{\hat{q}e^{xI/c}+(1-\hat{q})\,e^{rI/c}}. \tag{57}$$

Multiplication of both sides by $1-\hat{q}$ gives

$$(1-\hat{q})\sum_{x\in M}\frac{\mu(x)\cdot e^{xI/c}}{\hat{q}e^{xI/c}+(1-\hat{q})\,e^{rI/c}}=\sum_{x\in M}\mu(x)\cdot\frac{(1-\hat{q})\,e^{rI/c}}{\hat{q}e^{xI/c}+(1-\hat{q})\,e^{rI/c}}. \tag{58}$$

The left-hand side equals $(1-\hat{q})\,f(\hat{q})$. As for the right-hand side, we note that

$$\frac{(1-\hat{q})\,e^{rI/c}}{\hat{q}e^{xI/c}+(1-\hat{q})\,e^{rI/c}}=1-\hat{p}_{|X=x}. \tag{59}$$

for all $x\in M$. Multiplying both sides of the latter equation by $\mu(x)$ and summing over all $x\in M$, and using the identity $\sum_{x\in M}\mu(x)\,\hat{p}_{|X=x}=\hat{q}f(\hat{q})$, we obtain that the right-hand side of (58) equals $1-\hat{q}f(\hat{q})$. Hence, $(1-\hat{q})\,f(\hat{q})=1-\hat{q}f(\hat{q})$, or $f(\hat{q})=1$.

8.1. Proof of Proposition 1. Clearly f is a smooth function with

$$f'(q) = -\sum_{x\in M} \frac{\mu(x)\, e^{xI/c}\left(e^{xI/c} - e^{rI/c}\right)}{\left[qe^{xI/c} + (1-q)\, e^{rI/c}\right]^2} \quad \forall q \in [0,1] \tag{60}$$

and

$$f''(q) = 2\sum_{x\in M} \frac{\mu(x)\, e^{xI/c}\left(e^{xI/c} - e^{rI/c}\right)^2}{\left[qe^{xI/c} + (1-q)\, e^{rI/c}\right]^3} \geq 0 \;\forall q \in [0,1]\,. \tag{61}$$

In particular, f is a continuous and strictly convex (since M is not a singleton). Since $f(1) = 1$, $f(q) = 0$ for at most one $q \in (0,1)$. If $f(0) \leq 1$ then $f(q) < 1$ for all $q \in (0,1)$, which establishes (11), since $f(0) = \mathbb{E}\left[e^{(X-r)I/c}\right]$; investment conditions are sour if $\mathbb{E}\left[e^{(X-r)I/c}\right] \leq 1$. Likewise, $f(q) < 1$ for some $q \in (0,1)$ if $f(0) > 1$ and $f'(1) > 0$, which establishes (10), since $f'(1) = \mathbb{E}\left[e^{(r-X)I/c}\right] - 1$; investment conditions are normal iff $\mathbb{E}\left[e^{(X-r)I/c}\right] > 1$ and $\mathbb{E}\left[e^{(r-X)I/c}\right] > 1$. Investment conditions are sweet in the residual case, that is, when $f'(1) \leq 0$, or $\mathbb{E}\left[e^{(r-X)I/c}\right] \leq 1$.

8.2. Proof or Corollary 1. For any $c > 0$, there exist r meeting (10) if and only if

$$\ln \mathbb{E}\left[e^{XI/c}\right] + \ln \mathbb{E}\left[e^{-XI/c}\right] > 0. \tag{62}$$

By assumption, X has non-singleton support, and the logarithm is a strictly concave function, so by Jensen's inequality

$$\ln \mathbb{E}\left[e^{XI/c}\right] + \ln \mathbb{E}\left[e^{-XI/c}\right] > \mathbb{E}\left[\ln e^{XI/c}\right] + \mathbb{E}\left[\ln e^{-XI/c}\right] = I\cdot\left(\mathbb{E}\left[X/c\right] - \mathbb{E}\left[X/c\right]\right) = 0. \tag{63}$$

8.3. Proof of Corollary 2. Write $g(q)$ for the left-hand side of (18), with $\hat{q}$ replaced by an arbitrary $q \in [0,1]$. This defines $g : [0,1] \to \mathbb{R}$ as a smooth function with

$$g'(q) = -\sum_{x\in M} \frac{\mu(x)\, e^{u(\boldsymbol{w}(x))/c_a}\left(e^{u(\boldsymbol{w}(x))/c_a} - e^{u(y)/c_a}\right)}{\left[qe^{u(\boldsymbol{w}(x))/c_a} + (1-q)\, e^{u(y)/c_a}\right]^2} \tag{64}$$

and

$$g''(q) = 2\sum_{x\in M} \frac{\mu(x)\, e^{u(\boldsymbol{w}(x))/c_a}\left(e^{u(\boldsymbol{w}(x))/c_a} - e^{u(y)/c_a}\right)^2}{\left[qe^{u(\boldsymbol{w}(x))/c_a} + (1-q)\, e^{u(y)/c_a}\right]^3} \geq 0 \tag{65}$$

for all $q \in [0,1]$. Hence, similarly as for f in the case of an investor in autarky, g is continuous and strictly convex, and $g(1) = 1$. Hence, $g(q) < 1$ for all $q \in (0,1)$ if $g(0) < 1$, and $g(q) > 1$ for all $q \in (0,1)$ if $g(0) > 1$ and $g'(1) \leq 0$, where $g(0) = \mathbb{E}\left[e^{[u(\boldsymbol{w}(X)) - u(y)]/c_a}\right]$ and $g'(1) = \mathbb{E}\left[e^{[u(y) - u(\boldsymbol{w}(X))]/c_a}\right] - 1$, which establishes all claims.

8.4. Proof of Proposition 2. Let $\langle y, \boldsymbol{w} \rangle$ be a contract that solves program (25) for a risk neutral agent. Formally, we may rephrase the optimization program so that the principal also chooses $q \in (0,1)$ under constraint (27). Hence, we imagine a decision-maker who chooses y, $\boldsymbol{w}$ and q under the following constraints: $y, w_1, ..., w_m \geq 0$, $q \in (0,1)$, (27) and (23). By the Karush-Kuhn-Tucker theorem applied to this program, there exists scalars λ and ρ (the Lagrangians associated with conditions (27) and (23), respectively) such that the following equation holds for all states of nature j where $w_j > 0$:

$$\frac{\partial \Pi (y, \boldsymbol{w})}{\partial w_j} - \lambda \cdot \frac{\partial}{\partial w_j} \left[\sum_{i=1}^{m} \frac{\mu_i}{q + (1-q)\, e^{(y-w_i)/c_a}} \right] - \rho \cdot \frac{\partial U (y, \boldsymbol{w})}{\partial w_j} = 0. \tag{66}$$

We note that

$$\frac{\partial U (y, \boldsymbol{w})}{\partial w_j} = \frac{\partial V (y, \boldsymbol{w})}{\partial w_j}, \tag{67}$$

where $V(y, \boldsymbol{w})$ is the sum of the first two terms in (22). Moreover, because of risk-neutrality,

$$\frac{\partial \Pi (y, \boldsymbol{w})}{\partial w_j} + \frac{\partial V (y, \boldsymbol{w})}{\partial w_j} = 0.$$

Hence, the necessary condition (66) can be re-written as

$$(1-\rho) \cdot \frac{\partial \Pi (y, \boldsymbol{w})}{\partial w_j} \quad - \quad \lambda \cdot \frac{\partial}{\partial w_j} \left[\sum_{i=1}^{m} \frac{\mu_i}{q + (1-q)\, e^{(y-w_i)/c_a}} \right] \quad = \quad 0, \tag{68}$$

or, equivalently, as

$$\frac{\partial}{\partial w_j} \left[\sum_{i=1}^{m} \frac{[(1-\rho)\, q\, (Ix_i - w_i) - \lambda]\, \mu_i}{q + (1-q)\, e^{(y-w_i)/c_a}} \right] = 0$$

or, even simpler, as

$$\frac{\partial}{\partial w_j} \left[\frac{(1-\rho)\, q\, (Ix_j - w_j) - \lambda}{q + (1-q)\, e^{(y-w_j)/c_a}} \right] = 0,$$

which is

$$(1-\rho)\, q \left[Ix_j - w_j - c_a \cdot \left(1 + \frac{q}{1-q} e^{(w_j - y)/c_a} \right) \right] = \lambda.$$

Here $q \neq 0$. If $\rho \neq 1$, which holds generically, then

$$Ix_j - w_j - c_a \cdot \frac{q}{1-q} e^{(w_j - y)/c_a} = P$$

for some $P \in \mathbb{R}$, the same for all j with $w_j > 0$. This gives (29).

8.5. Proof of Corollary 4. By Proposition 2 there exists a constant P such that for all $x \in M$ with $w(x) > 0$:

$$Ix = w(x) + Q \cdot e^{w(x)/c_a} + P \tag{69}$$

or, equivalently,

$$e^{w(x)/c_a} = \frac{Ix - P - w(x)}{Q} \tag{70}$$

for

$$Q = \frac{q^* c_a}{1 - q^*} e^{-y/c_a} > 0.$$

Equation (70) defines x as a continuous, strictly increasing and strictly convex function f of the payment $z = w(x)$ at each net return rate x. Clearly f is continuous and strictly increasing. Hence it has an inverse, $g = f^{-1}$, which is also continuous and strictly increasing, as well as strictly concave, and we have $w(x) = \max\{0, g(x)\}$.

We proceed to show that the function g can be expressed in terms of the the so-called principal branch, W_0, of the Lambert-W correspondence, implicitly defined for all $z > -1/e$ by $y = W_0(z)$ for $ye^y = z$. This correspondence is singleton-valued for all $z \geq 0$. First, consider the following simple transcendental algebraic equation in $w \in \mathbb{R}$:

$$e^{-w/c} = a(w - x) \tag{71}$$

for $a, c > 0$ and $x \in \mathbb{R}$. This equation can be written as

$$\frac{w - x}{c} \cdot e^{(w-x)/c} = \frac{e^{-x/c}}{ac},$$

or $ye^y = z$, for $y = (w - x)/c$ and $z = e^{-x/c}/(ac) > 0$. Hence, by definition of the Lambert-W correspondence,

$$W\left(\frac{e^{-x/c}}{ac}\right) = \frac{w - x}{c},$$

or

$$w = x + cW\left(\frac{e^{-x/c}}{ac}\right),$$

which thus solves (71), where we may replace W by W_0, the principal branch of W (since the argument is positive). It follows that the similar equation $e^{w/c} = a(x - w)$ has the solution

$$w = x - cW_0\left(\frac{e^{x/c}}{ac}\right).$$

Accordingly, a solution of (70) is

$$w = Ix - P - c_a W_0 \left(Q \frac{e^{(Ix-P)/c_a}}{c_a} \right)$$

which establishes (31), for A is such that

$$Q \frac{e^{(Ix-P)/c_a}}{c_a} = Ae^{Ix/c_a},$$

or

$$A = \frac{Q}{c_a} e^{-P/c_a} = \frac{q^*}{1-q^*} e^{-(P+y)/c_a}.$$

8.6. Proof of Proposition 3. We here study the optimization program for the principal when the agent's participation constraint is not binding. The agent's Bernoulli function $u : \mathbb{R}_+ \to \mathbb{R}$ is assumed to be either linear or twice differentiable on $\mathbb{R}_{++}$, with $u' > 0$, $u'' \leq 0$, and $\lim_{z \to 0} u'(z) = +\infty$.

The principal faces the program

$$\max_{\boldsymbol{w},y,q} \quad q \cdot \sum_{i=1}^{m} \frac{\mu_i \cdot [(x_i - r) I + y - w_i]}{q + (1-q) e^{[u(y) - u(w_i)]/c_a}} \quad - \quad y, \tag{72}$$

where $\boldsymbol{w} \in \mathbb{R}_+^m$, $y \in \mathbb{R}_+$, $q \in (0,1)$, and

$$\sum_{i=1}^{m} \frac{\mu_i}{q + (1-q) e^{[u(y)-u(w_i)]/c_a}} = 1. \tag{73}$$

Given (73), we may w.l.o.g. rewrite (72) as

$$\max_{\boldsymbol{w},y,q} \quad \sum_{i=1}^{m} \frac{(Ix_i - w_i) \mu_i q}{q + (1-q) e^{[u(y)-u(w_j)]/c_a}} \quad - Irq - (1-q) y.$$

Suppose that $\langle \boldsymbol{w}, y, q \rangle$ is a solution. Let $J \subseteq \{1, ..., m\}$ be those j for which $w_j > 0$. By the Karush-Kuhn-Tucker theorem, there exists a $\lambda \in \mathbb{R}$, the Lagrangian associated with (73), such that for all $j \in J$:

$$\frac{\partial}{\partial w_j} \left[\sum_{i=1}^{m} \frac{q (Ix_i - w_i) \mu_i}{q + (1-q) e^{[u(y)-u(w_i)]/c_a}} \right] - \lambda \cdot \frac{\partial}{\partial w_j} \left[\sum_{i=1}^{m} \frac{\mu_i}{q + (1-q) e^{[u(y)-u(w_i)]/c_a}} \right] = 0.$$

Equivalently:

$$\frac{\partial}{\partial w_j} \left[\frac{Iqx_j - qw_j - \lambda}{q + (1-q) e^{[u(y)-u(w_j)]/c_a}} \right] = 0 \quad \forall j \in J,$$

or

$$Ix_j = w_j + \frac{c_a}{u'(w_j)} \cdot \left(1 + \frac{q}{1-q} e^{[u(w_j)-u(y)]/c_a}\right) + \frac{\lambda}{q} \quad \forall j \in J. \tag{74}$$

We note that, for given λ, y and q, this equation uniquely determines each x_j, for $j \in J$, as a strictly increasing function of w_j, and hence also w_j as a strictly increasing function of x_j, for all $j \in J$, while $w_j = 0$ for all $j \notin J$.

It follows that

$$w(x) = \max\{0, \phi(x)\} \quad \forall x \in M$$

where $M = \{x_1, ..., x_m\}$ and $\phi : X \to \mathbb{R}$ is implicitly defined, on some domain $X \subseteq \mathbb{R}$, by

$$Ix = \phi(x) + \frac{c_a}{u'(\phi(x))} \cdot \left(1 + \frac{q}{1-q} e^{[u(\phi(x))-u(y)]/c_a}\right) + P$$

for some constant $P \in \mathbb{R}$.

To see that this equation indeed uniquely defines $\phi(x) \in \mathbb{R}$ for every $x \in X$, consider the equation $Ix = f(w)$, where

$$f(w) = w + \frac{c_a}{u'(w)} \cdot \left(1 + \frac{q}{1-q} e^{[u(w)-u(y)]/c_a}\right) + P,$$

for $I, c_a > 0$, $y \geq 0$, u as specified above, $q \in (0,1)$, and $b \in \mathbb{R}$ are fixed and given. If u is linear, this equation defines defines $f : \mathbb{R} \to \mathbb{R}$ as continuous and strictly increasing surjection, and thus its inverse f^{-1} exists and coincides with ϕ. In this case, the domain of ϕ is $\mathbb{R}$. If u is non-linear, then u has domain $\mathbb{R}_+$. By assumption, $u'(w) \to +\infty$ as $w \downarrow 0$, so f is then a continuous and strictly increasing surjection from $\mathbb{R}_+$ to $[P, +\infty)$. Hence, for every $x \in X = [P, +\infty)$ there exists a unique $w \geq 0$ satisfying the equation. This defines $\phi : [P, +\infty) \to \mathbb{R}_+$.

REFERENCES

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

[2] Caplin, A., M. Dean, and J. Leahy (2017): "Rationally inattentive behavior: characterizing and generalizing Shannon entropy", mimeo., New York University, Columbia University, and University of Michigan.

[3] Carroll, G. (2017): "Robust incentives for information acquisition", mimeo., Stanford University.

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

[5] Crémer, J., and F. Khalil (1992): "Gathering information before signing a contract", *American Economic Review* 82, 566-578

[6] Crémer, J., F. Khalil, and J.-C. Rochet (1998a): "Strategic information gathering before a contract is offered", *Journal of Economic Theory* 81, 163-200.

[7] Crémer, J., F. Khalil, and J.-C. Rochet (1998b): "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, and M. Schum (2017): "Discrete choice and rational inattention: a general equivalence result", mimeo., Technical University of Denmark, Indiana University, and CalTech.

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

[11] Gossner, O., P. Fernandez, and A. Neyman (2006): "Optimal use of communication resources", *Econometrica* 74, 1603-1636.

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

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

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

[15] 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.

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

[17] Lindbeck, A. and J. Weibull (2015): "Pay schemes, bargaining, and competition for talent", IFN Institute for Industrial Economics, Stockholm, WP No. 1100.

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

[19] 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.

[20] Mattsson, L.-G. and J. Weibull (2002): "Probabilistic choice and procedurally bounded rationality", *Games and Economic Behavior* 41, 61-78.

[21] 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.

[22] Sims, C. (1998): "Stickiness", *Carnegie-Rochester Conference Series on Public Policy* 49, 317-356.

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

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

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

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

[27] Steiner, J., C. Stewart, and F. Matějka (2017): "Rational inattention dynamics: inertia and delay in decision-making", *Econometrica* 85, 521-553.

[28] Weibull J., L.-G. Mattsson, and M. Voorneveld (2007): "Better may be worse: some monotonicity results and paradoxes in discrete choice under uncertainty", *Theory and Decision* 63, 121-141.

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

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

[31] Yang, M., and Y. Zeng (2017): "Financing entrepreneurial production: security design with flexible information acquisition", mimeo., Duke University and University of Washington.

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

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