RESEARCH INSTITUTE OF INDUSTRIAL ECONOMICS

IFN Working Paper No. 1171, 2017

Investment, Rational Inattention, and Delegation

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

Investment, rational inattention, and delegation

ASSAR LINDBECK* AND JÖRGEN WEIBULL†

May 19, 2017.‡

ABSTRACT. We analyze investment decisions when information is costly, with and without delegation to an agent. We use a rational-inattention model and compare it with a canonical signal-extraction model. We identify three "investment conditions". In "sour" conditions, no information is acquired and no investment made. In "sweet" conditions, investment is made "blindly", i.e. without acquiring costly information. In intermediate, "normal" conditions, the decision-maker acquires information and conditions the investment decision upon the information obtained. We investigate if the investor can benefit from employing an agent when the agent's effort and information is private. Not even in the case of a risk neutral agent will the principal perfectly align the agent's incentives with her own at the moment of investment (had the principal known the agent's private information). Optimal contracts for risk neutral agents not only reward good investments but also punishes bad investments. Such contracts include three components: a fixed salary, stocks and options.

Keywords: Investment, rational inattention, signal extraction, principal-agent, information acquisition, 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.

‡The authors thank Atahan Afsar, Ingela Alger, Bo Becker, Jun Chen, 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 for some agreed-upon project. However, an equally important task for many agents, such as CEOs of large corporations and pension-fund managers, is to make well-informed investment decisions. This is a theme we here focus on, following up on pioneering work by Demski and Sappington (1987) and Lewis and Sappington (1997).[1] We take the agent to be someone who has a comparative advantage over the principal in acquiring and evaluating pertinent information about investment opportunities or projects. However, such information is costly for the agent to acquire. It requires him to exert (non-contractible) effort, perhaps over many days or months, 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 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 he has. An important issue is thus how to motivate the agent to acquire relevant and reliable information and then use it in line with the principal's interests when making the investment decision.

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 return. In order to keep back the agent's potential eagerness to invest, because of his hope to earn a bonus, the agent has to be paid also for not investing, and has 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 associated with investment. Moreover, under limited-liability, a low salary limits the size, and hence effect, of penalties, since the maximal penalty then is to withdraw the salary. This causes an asymmetry in contracts, even for risk neutral agents and investment

[1] The motivation behind those studies is similar to ours: "In many relevant settings, however, the agent is not omniscient from the outset, and the principal deems it important to motivate the agent to acquire valuable planning information before he acts" (Lewis and Sappington, 1997, pp. 796-7).

projects that are *ex ante* symmetric in terms of upsides and downsides.

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. Such a manager chooses how much effort to make to acquire access to pertinent information, and assess this, for investment decisions. Because of the personal cost to the manager (say, in terms of long hours at work), there is here a major moral hazard issue. Other examples are given by the consulting industry, where firms and institutions decide whether or not to delegate information acquisition and (in effect) decisions to consultants, whose efforts and information cannot be monitored. The flexibility of the rational inattention approach enables us to capture both the hidden effort and information structure in such situations. Our model is intended to capture some key features of agency relationships of the mentioned sorts.

The focus of our study is on costly information acquisition and subsequent decision-making, both for an investor without agent and for an investor with agent. 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 bench-mark. 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, buy up 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. When there is an agent, the principal delegates the investment decision to the agent, or the agent recommends the principal what investment decision to take. Given the assumed difficulties for the principal to obtain and evaluate the agent's information, this distinction is immaterial. A contract between the two parties specifies a payment to the agent under every possible outcome, including non-investment. We require contracts to meet the limited liability constraint that the net pay from the principal to the agent be nonnegative.[2]

Having signed such a contract, the agent decides how much, if any, effort to make and time to spend 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 agent's effort and information are his private information. The terms of the contract influence the agent in two distinct ways. First, it motivates the agent to acquire information about the project at hand. 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 already by Demski and Sappington (1987). They provide sufficient conditions for when this tension, which they call *induced moral hazard*, may arise (their Proposition 1), and illustrate this

[2] Or, more generally, does not fall below some specified bound.

with a numerical example.[3] It turns out that our rational-inattention formalization provides analytical power that generates important new insights (summarized below).

As a starting point we take the rational-inattention approach, pioneered by Sims (1998, 2003, 2006), in the form developed by Matějka and McKay (2015). A major appeal of that approach is that it does not impose restrictive assumptions on what form of information the decision-maker chooses to acquire.[4] Rationally inattentive decision-makers optimally acquire and process information that is relevant and useful for the decision at hand and ignore information that is not worth the effort of acquiring and processing. This approach turns out to be particularly analytically convenient for analysis of delegation problems. In particular, it is easier to work with than a model of explicit signal extraction. However, the latter model has the advantage of being more transparent in specifying the form of information available. We therefore also develop a canonical signal-extraction model (but only for the case of binary return distributions) in order to check the robustness of the results with respect to choice of model.

We have six main findings. The first is that the inherent moral hazard problem is so substantial that it is not worthwhile for an investor to hire an agent unless the agent has a strong comparative advantage in acquiring and assessing information. The agent's unit cost of information acquisition has to be substantially lower than the principal's.

Our second finding is that optimal contracts with risk neutral agents are monotonic with respect to realized returns, and they contain bonuses as well as penalties (as compared with the pay after non-investment), and, moreover, that they are non-linear. Hence, contracts based on a flat salary and shares in the project are suboptimal, and so are contracts based on a salary plus an option to buy future stocks at today's price. The first type of contract is suboptimal because it provides too weak incentives for information acquisition, and the latter type of contract is suboptimal since it contains no penalty in case of a failed investment. Under such contracts, the agent will always invest. We also show that the limited-liability constraint on contracts is binding for risk-neutral agents. More precisely, the agent's pay after investment in the worst state of nature is zero, that is, to withhold the agent's salary.

Our third finding is that it is typically optimal for the principal to not fully align

[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).

[4] See de Palma and Fosgerau (2016) for a generalized version of the rational-inattention model that allows for exogenous "information filters" that represent that, in practice, some types of information are more easily available than others. See also Fosgerau, Melo, and Shum (2017) for results on the equivalence between generalized rational inattention and discrete choice under additive random utility.

the agent's incentives at the moment of the investment decision (after information has been acquired and those costs sunk) with those of the principal (had the principal had the agent's private information), not even when the agent is risk neutral. The reason is the limited-liability constraint; in order to deter the agent from risky investments, which may be very costly for the principal but less costly to the agent because of the mentioned constraint, it may be optimal to tilt the contract slightly in favor of bonuses and away from penalties. A way to obtain this in practice is to let the contract consist of a flat salary, some stocks, and some options.

Our fourth finding is that it may be optimal for the principal to offer the agent a contract with expected utility above the agent's reservation utility. In other words, the agent's participation constraint may well be slack in an optimal contract. The reason is, again, incentives, since by cutting the pay in all states of the world (including non-investment), the agent's investment incentives are unchanged but the agent may then become insufficiently motivated to acquire information. Hence, while we do not deny the potential importance of real-life agents' potential bargaining power over the terms of their own contracts, in some situations, lavish contract may in fact be in the best interest of the principal.

Our fifth finding is that investment conditions may be such that a rationally inattentive decision-maker, be it the investor herself or the agent, rationally decides to invest without acquiring any information. We call such investment conditions "sweet". This form of "individually *rational* exuberance" may be part of the explanation of times of apparently uncritical investment. In such situations changes in market conditions are experienced only with considerable delay, after investments have been made. The opposite case is also possible, namely, that a rationally inattentive decision-maker decides to completely ignore a project, that is, not acquire any information and not invest. We call such investment conditions "sour". By contrast, under "normal" investment conditions, which fall between these two extremes, the decision-maker acquires some information and then makes the investment decision conditional upon the information received. The identification of "sour" investment conditions is in line with observations in Matějka and McKay (2015) and, in particular, Caplin, Dean, and Leahy (2016). Indeed the latter provide necessary and sufficient conditions for discrete-choice alternatives to be at all considered by rationally inattentive decision-makers (their Proposition 1), thereby giving a precise characterization of what is called a consumer's *consideration set* in the marketing literature.[5]

Our sixth and final finding is that the rational-inattention approach is much easier to work with than an explicit signal-extraction model, and yet the two approaches seem to yield qualitatively very similar results, at least in simple examples. This

[5]For a game-theoretic definition of consideration sets, see Myerson and Weibull (2015).

finding speaks in favor of the rational-inattention model, which also more readily lends itself to generalization. For instance, while we here only consider one investment opportunity, multiple alternative investment opportunities can easily be handled in the rational-inattention model, as can general outcome distributions.[6]

To the best of our knowledge, we are among the first to use the rational-inattention approach for 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. More specifically, they analyze the optimal mix of debt financing and equity financing. Both the investor's information acquisition and subsequent financing decisions are endogenous. In Yang's and Zeng's (2017) 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 closest seems to be Lewis and Sappington (1997), who analyze situations in which the agent, at a cost, can choose to be perfectly informed about the state of nature. In a related paper, Levitt and Snyder (1997) analyze how interventions by the principal may undermine the incentives for the agent.[7] Methodologically, the present study differs starkly from these papers. Our approach also 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, excluding the possibility of full information).

In recent decades, the high reward to top-managers of large corporations has been a central issue in the policy debate (for a survey of the debate, see Murphy, 2012). Much of that literature either emphasizes the role of stiff competition for scarce talents in the market or the role of top-officers' strong bargaining power within organizations. These factors appear also in this study, in terms of the outside option of managers and both parties' indifference curves in contract space. However, in this paper we give a richer (higher-dimensional) view on optimal contracts to top-managers – in the context of a non-standard principal-agent model.

[6]See also Steiner, Stewart and Matějka (2017) for an extension of the rational inattention model to dynamic decision problems.

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

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. Section 4 studies the robustness of the rational-inattention model by developing a model of explicit signal-extraction with additive normal noise and endogenous signal precision.[8] Section 5 concludes. Mathematical proofs are provided in Appendix A for the rational inattention model, and in Appendix B for the signal-extraction model.

2. A RATIONALLY INATTENTIVE INVESTOR

We begin by studying a risk-neutral and rational investor who considers 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 probability distribution of X is known by the investor. Its finite support is $M = \{x_1, x_2, ..., x_m\}$, with $x_1 < x_2 < ... < x_m$ for some $m > 1$, and the probability for each such potential realization x_i is positive and denoted $\theta(x_i)$ or θ_i The vector $\theta = (\theta_1, ..., \theta_m)$ is 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 easily available information about the project at hand, such as credit ratings and earnings records of the people involved etc. The key assumption is that this information is available to the investor for free. Given her prior, the investor first decides how much time and effort, if any, to spend on further information acquisition about the project before she decides whether or not to invest in the project. The investment decision is thus binary.

If the investor chooses not to invest, she obtains a risk free net-return rate r. Hence, Ir is the her opportunity cost for investing. Alternatively, r can be thought of as 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. This posterior belief will be based upon the information she obtains. 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.

[8] This section builds upon an earlier working paper, see Lindbeck and Weibull (2015).

We will analyze this scenario, with and without agent, by applying the rational-inattention model of Matějka and McKay (2015), a model which, in turn, builds upon pioneering work by Sims (1998, 2003, 2006). Accordingly, we will 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 prior belief, and with information costs represented in terms of entropy reduction.[9] This will lead to investment probabilities that are conditional on the true state of nature. These conditional investment probabilities will depend on the investor's choice of how well-informed she wants to be when making her investment decision. 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 using Theorem 1 and Lemma 2 in Matejka and McKay (2015), one immediately obtains 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

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

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, as shown in Corollary 2 in Matejka and McKay (2015) 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

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

[9] 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 diverse applications to economics and the social sciences, see Snickars and Weibull (1977), Mattsson and Weibull (2002), Gossner, Hernandez and Neyman (2006), and Yang (2015). Cabrales, Gossner and Serrano (2013) analyze orderings of information structures in terms of their informativeness for choice of an investment project from a finite set of alternatives, with emphasis on orderings based on entropy.

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 $f(\hat{q}) = 1$).[10] 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 (before information has been acquired): $\hat{q} = \mathbb{E}\left[\hat{p}_{|X}\right]$.

The investor's conditionally expected profit, given that the investment's true net return rate is $x \in M$, equals the net return xI from investing, times the conditional investment probability in that state of nature, plus the net return from not investing, times the conditional probability for not investing in that state, minus the investor's information costs. This *conditionally expected profit* can be written in the following form

$$\hat{\Pi}(x,c) = \frac{(x-r)\, I\hat{q}e^{xI/c}}{\hat{q}e^{xI/c} + (1-\hat{q})\, e^{rI/c}} + Ir - c \cdot \left[H(\hat{q}) - H\left(\hat{p}_{|X=x}\right)\right] \quad \forall x \in M \qquad (6)$$

Here $c > 0$ is the investor's *unit cost of information*, mentioned above, and H is the entropy function for a binary probability distribution, that is, $H : [0,1] \to \mathbb{R}$ is defined by

$$H(p) = -\, p \ln p - (1-p) \ln(1-p) \qquad (7)$$

(with the convention $0 \ln 0 = 0$).

Entropy represents the uncertainty embodied in a probability distribution. It is minimal and takes the value zero when there is no uncertainty (when $p = 0$ or $p = 1$), and it is maximal when all outcomes are equally likely ($p = 1/2$ in the binary case). In his seminal paper, Shannon (1948), characterizes axiomatically entropy as a quantitative measure of the uncertainty inherent in probability distributions with finite support. More precisely, his Theorem 1 establishes that entropy is the unique measure (up to scaling) that satisfies three desirable qualitative properties. First, the measure should be continuous in the probabilities. Second, if all outcomes have the same probability, then the measure should be larger the more outcomes there are. Third, suppose that the random mechanism generating the outcomes consists of two steps, where first a random draw is made of a subset (or "cell") from a partitioning of the set of outcomes, and then a random draw is made from within the selected subset (or cell). Then the uncertainty measure should be the probabilistically weighted sum of the uncertainty measures of the random draws within each subset.

In (6), $H(\hat{q})$ is the entropy of the *ex ante* investment probability and $H\left(\hat{p}_{|X=x}\right)$ the entropy of the conditional investment probability when $x \in M$. The difference $H(\hat{q}) - H\left(\hat{p}_{|X=x}\right)$ thus represents the entropy reduction when moving from the average investment probability $\hat{q}$ to the adapted investment probability $\hat{p}_{|X=x}$ in the state

[10] 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 $\hat{q} \in (0,1)$.

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 (or, in practice, feasible). The investor has to trade off information costs against information benefits for her subsequent investment decision.

Taking expectations, according to the investor's prior $\theta(\cdot)$, we obtain the following expression for the investor's *ex ante* expected profit:

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

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

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

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 $\hat{q}$ to $\hat{p}_{|X}$.

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 (almost) always invests, $\hat{p}_{|X} = 1$ (a.s.), while in the second case she (almost) 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, which agrees with Lemma 2 in Woodford (2008) and Proposition 1 in Yang (2015), characterizes the three investment conditions in our simple model:

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*

$$-c\ln\mathbb{E}\left[e^{-XI/c}\right] < Ir < c\ln\mathbb{E}\left[e^{XI/c}\right], \quad (10)$$

sour if

$$Ir \geq c\ln\mathbb{E}\left[e^{XI/c}\right], \quad (11)$$

and sweet if

$$Ir \leq -c\ln\mathbb{E}\left[e^{-XI/c}\right]. \quad (12)$$

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:

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 $\upsilon > 0$,

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

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

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

The last equation implies that $\hat{q} \to \Pr[X > r]$ as $\upsilon \to +\infty$.[11] 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 $\upsilon \to +\infty$, the investor acts as if she were perfectly informed about the true state of nature.

We illustrate the above analysis graphically for a double-or-nothing project that either has net return rate 1 or -1. Let θ be the probability for the first outcome. The

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

investor's *ex ante* expected profit is shown as a function of the prior θ for the "good" state of nature is shown in Figure 3 below, drawn for $I = 1$, $r = 0$ and $c = 0.5$.

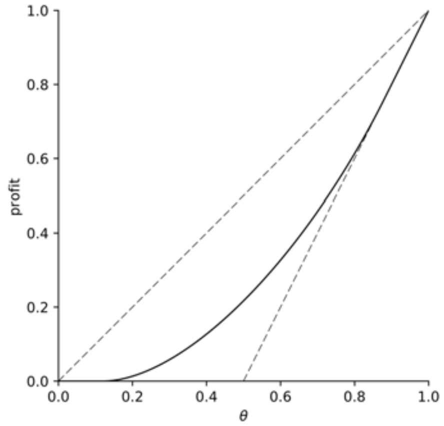

Figure 1: The investor's expected profit in the rational-inattention model.

In this example, the range of priors θ for which the investment conditions are normal is approximately $0.12 < \theta < 0.88$.[12] The upward-bending curve is the investor's expected profit. Not surprisingly, it is increasing in the prior θ for the good state of nature. Its horizontal line segment, for low θ, represent her expected profit (zero) in sour investment conditions ($\theta \lesssim 0.12$). For $\theta \gtrsim 0.88$, 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 (that is, who knows the true state of nature, or, equivalently, has zero information cost, $c = 0$). Such an investor will invest if and only if the state of nature is good (which it is with probability θ). By contrast, the expected profit to a completely uninformed investor (who's only information is the prior, or equivalently has infinite information cost, $c = +\infty$) is zero for all $\theta \leq 1/2$, since it is then optimal for her not to invest. It is $2\theta - 1$ for all $\theta > 1/2$, since it is then optimal for her to invest. This is the dashed steep line.

The diagram thus shows how the *ex ante* expected profit to the rationally inattentive investor lies between the two extremes—of perfect information or no information—in normal investment conditions, and how the investor chooses not to be informed

[12] More exactly, the normality condition writes

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

when the probability for the good state of nature is either low or high. Not surprisingly, the rationally inattentive investor always does worse than the perfectly informed investor, never worse than the uninformed investor, and, for an interval of moderate priors, does better than the latter. For low and high priors she does no better than the uninformed investor. It is in the middle range—under normal investment conditions, and only then—that she does better than the latter. As the unit information cost $c > 0$ diminishes, the profit curve in the diagram moves upwards until it hits the diagonal in the limit as $c \to 0$.

3. Delegation to a rationally inattentive agent

Suppose now that the investor considers the possibility of hiring an agent who has some comparative advantage in acquiring and processing relevant information. We model also this actor as rationally inattentive. Suppose that the principal knows the agent's unit cost for collecting and analyzing information as well as his 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 return. In view of these informational difficulties, the principal delegates the investment decision to the agent, if hired.

The agent is risk neutral or risk-averse, and we take him to be purely self-interested, a *homo oeconomicus*. His effort and investment decision will thus be entirely driven by self interest. The investor may benefit from such an arrangement—compared with acting in autarky—if the agent is talented enough (has a low unit cost of information), makes sufficient effort to gather and process information, and subsequently makes a wise investment decision from the principal's viewpoint.

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 $\bar{u}$.

Again applying Theorem 1, Lemma 2 and Corollary 2 in Matejka and McKay (2015), we obtain similar expressions for the agent, once hired, as were obtained for the investor in autarky. Essentially, one only has to replace the investor's net returns by utilities from remuneration when investing and not investing. The agent's salary will play the same role as the investor's opportunity cost rI.

To be more precise, let $c_a > 0$ be the agent's unit information cost, q^* his *a priori* investment probability, and $p^*_{|X=x}$ his conditional investment probability when $x \in M$. 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

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

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

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

Just as for the investor in autarky, the agent's investment conditions may be "sour", "normal" or "sweet". Necessary and sufficient conditions for the 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 $x \in M$.[13]

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*

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

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 he is not worthwhile to hire). Third, among all contracts, if any, that meet these two requirements, the contract should yield the highest possible expected profit to the principal. Fourth and finally, the principal's maximal expected profit from hiring the agent should exceed the expected profit from not hiring the agent, which could be either to not invest, yielding the net return rate r, or to acquire information herself (at unit information cost c) and making the investment decision single-handedly.

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, in analogy with (8), be written as

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

The first condition mentioned above, 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 words, 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. As will be seen shortly, this requirement implies that the agent's participation constraint will not always be binding, not even when the agent is risk neutral. It may be in the principal's interest to pay the agent enough in salary so that a potential loss of salary deters him from making risky investments or being poorly informed.

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 $\langle y, \boldsymbol{w} \rangle$. This expected profit is the convex combination of two terms, where the first is a probability-weighted sum of the net returns to the principal from investing in each state (net of the payment to the agent). The second term is the principal's interest earnings net of the salary to the agent:

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

The weight on the first term, q^*, is the agent's *ex ante* investment probability—before acquiring information—under contract $\langle y, \boldsymbol{w} \rangle$. 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, $q^* = 0$, 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, $q^* = 1$, then the agent invests "blindly", resulting in $\mathbb{E}[\Pi(y, \boldsymbol{w})] = I \cdot \mathbb{E}[X - \boldsymbol{w}(X)] \leq I \cdot \mathbb{E}[X]$. Clearly, there is no point 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 and provides normal investment conditions to the agent. Formally, this third 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, when acquiring information in autarky has unit information cost c, then it is optimal for her to offer the agent the said contract if and only if

$$\Pi(y^*, \boldsymbol{w}^*) \geq \hat{\Pi}(I, c), \tag{26}$$

where we note that if the principal's unit information cost is prohibitively high ($c \to \infty$), then she would in autarky not acquire any information, and then (26) would become $\Pi(y^*, \boldsymbol{w}^*) \geq I \cdot \max\{r, \mathbb{E}[X]\}$.

3.1. Risk-neutral agent. We here examine in more detail the special case of a risk-neutral agent.[14] The principal's expected profit from any contract $\langle y, \boldsymbol{w} \rangle$ can then be written in the form

$$\Pi(y, \boldsymbol{w}) = \sum_{i=1}^{m} \frac{q^* \theta_i [(x_i - r) I + y - w_i]}{q^* + (1 - q^*) e^{(y - w_i)/c_a}} - y \tag{27}$$

where $\theta_i = \theta(x_i)$ and $w_i = \boldsymbol{w}(x_i)$ for each state of nature i, and, under normal investment conditions for the agent, $q^* \in (0, 1)$ satisfies

$$\sum_{i=1}^{m} \frac{\theta_i}{q^* + (1 - q^*) e^{(y - w_i)/c_a}} = 1. \tag{28}$$

[14]The case of a risk-averse agent with logaritmic utility of income is also analytically tractable. However, for the sake of brevity we do not treat that case.

The agent's behavior, once hired, is thus driven entirely by the *net transfers* after investment from the principal to the agent, the differences $t_i = w_i - y$ (when positive a bonus, when negative a penalty). Because the vector of net transfers, $t = (t_1, ..., t_m)$, determines q^*, the *ex ante* probability that the agent will invest, according to equation (28), and, given q^*, this determines in turn all conditional investment probabilities, according to equation (16), which results in equation (27).

The agent's expected utility takes the form

$$\begin{aligned} U(y, \boldsymbol{w}) \quad &= \quad y \;+\; \sum_{i=1}^{m} \frac{q^* \theta_i (w_i - y)}{q^* + (1 - q^*) e^{(y - w_i)/c_a}} \;-\; c_a \cdot H(q^*) \\ &\quad + c_a \cdot \sum_{i=1}^{m} \theta_i H\left(\frac{q^*}{q^* + (1 - q^*) e^{(y - w_i)/c_a}} \right) \end{aligned} \tag{29}$$

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

Proposition 2. *If a contract $\langle y, \boldsymbol{w} \rangle$ solves (25), then $q^* \in (0, 1)$ is uniquely determined by (28). Moreover,*

$$Ix_i - w_i - c_a \frac{q^*}{1 - q^*} e^{(w_i - y)/c_a} = Ix_j - w_j - c_a \frac{q^*}{1 - q^*} e^{(w_j - y)/c_a} \tag{30}$$

for any states of nature i and j with $w_i, w_j > 0$.

Each side of (30) is strictly decreasing in the payment to the agent after investment, but 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: if $w_i, w_j > 0$ and $x_i > x_j$, then $w_i > w_j$. 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 directly in the definition (27) of the principal's expected profit. Suppose that $w_1 > 0$. Then subtract some small amount $\varepsilon \in (0, w_1)$ from all payments w_i after investment and also from the salary, y. This does not affect the first term in (27), since the agent's behavior, once hired, is driven entirely by the net transfers. However, such a subtraction reduces the second term, the agent's salary, by ε. Hence, a net increase in the principal's expected profit (by $\varepsilon > 0$). In sum:

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

$$w_1 = ... = w_k = 0 < w_{k+1} < ... < w_m. \tag{31}$$

Moreover, irrespective of whether the agent's participation constraint is binding or not, the pay schedule after investment is strictly monotonic in the realized net return of investment, wherever positive. Indeed, it follows from (30) that the pay schedule is strictly concave where positive.[15] Consequently, linear contracts (such as a fixed salary plus stocks) are suboptimal.

Corollary 4. *If a contract $\langle y, \mathbf{w} \rangle$ is optimal, then there exists a strictly increasing and strictly concave function $g : \mathbb{R} \to \mathbb{R}_+$ such that $\mathbf{w}(x) = \max\{0, g(x)\}$ for all $x \in M$, where*

$$g(x) = Ix - c_a \cdot W\left(Ae^{Ix/c_a}\right) - B \tag{32}$$

for constants $A > 0$ and B, where W is the Lambert W function.

The so-called Lambert W function is implicitly defined by $y = W(z)$ where $ye^y = z$, see e.g. Corless *et al.* (1996). The result in Corollary 4 is illustrated in the diagram below (drawn for $I = 1$, $c_a = 0.05$, $A = 0.5$ and $B = 0.015$).

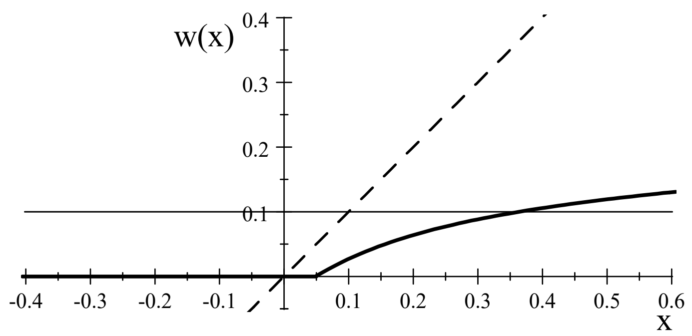

Figure 2: The shape of the optimal contract to a risk-neutral agent.

[15] More precisely, all payments on M can be represented by a continuous, strictly increasging and concave function on the convex hull of M. There are infinitely many other reprentation functions with this wider domain that lack all these properties outside the finite subset M of theír domain.

The kinked solid curve is the after-investment payment function $\boldsymbol{w}$, the thin horizontal line the agent's salary y, and the dashed straight line the net return, Ix, from investment. In this example, the agent thus has to pay a penalty if he invests and the return rate x falls below $\approx 36\%$, and he is paid a bonus if he invests and the net return rate exceeds this level. His bonus increases with the investment's net return, but at a falling rate. The principal thus retains an increasing share. The penalty is maximal—equal to agent's salary—for all net return rates below $\approx 5\%$.

Returning to the general case, we note that optimal pay schemes are "more linear" the lower is the agent's unit information cost. In the limit when the agent is perfectly informed about the state of nature at no cost, the pay scheme becomes affine. The conflict between the need to incentivize (a) information acquisition and (b) judgement at the moment of investment is then mute. To see this, consider Proposition 2, assume that $w_1 < y < w_m$, and let $c_a \to 0$. By (28), $q^* \to \mathbb{P}[w(X) > y] \in (0,1)$. Hence, in the limit $Ix_i - w_i = Ix_j - w_j$ for all $w_i, w_j > 0$. In other words, the principal then pays the agent a constant share of the net return from investment.

Is it worthwhile for the principal to hire an agent? This depends on the agent's comparative advantage in acquiring and processing relevant information about investment projects. In Figure 3 below we compare autarky with delegation in the same numerical specification as in Figure 1 ($I = 1$, $m = 2$, $x_1 = -1$, $x_2 = 1$, $r = 0$, and $c = 0.5$).

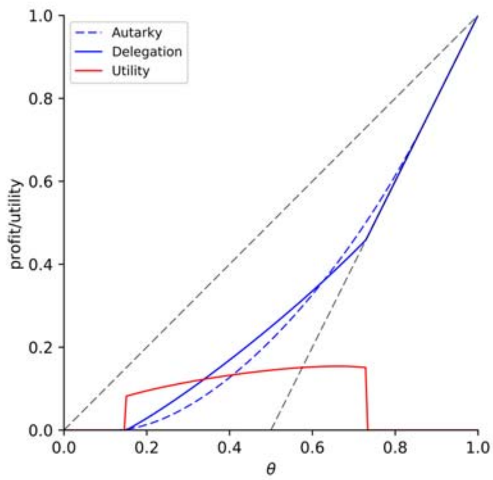

Figure 3: The investor's ex ante expected profit under delegation.

The solid curve is the investor's expected profit under delegation to a risk neutral agent with ten times lower unit information cost than the investor, $c_a = 0.05$. The dashed curve is the investor's expected profit in autarky. We see that delegation results in a higher expected profit to the investor for a wide intermediate range of priors θ. In particular, for priors near one half, when there is most to be gained by way of information acquisition, it is worthwhile for the principal to hire the agent. But will the agent accept the offered contract? The low curve, that jumps up from zero, then increases slightly, and then falls back to zero, is the agent's expected utility under the corresponding globally optimal contract. Hence, the agent's participation constraint will not be binding if his reservation utility $\bar{u}$ is less than approximately 0.08 in this example. We also see that, for the intermediate interval of θ-values at which there are positive gains of trade, welfare, defined as the sum of expected profits and utility, is an increasing function of θ, running not far below the maximum possible (the diagonal), with the share that befalls the principal increasing from zero to about two thirds.

Evidently the gains of trade are (smaller) larger the (higher) lower is the agent's unit cost of information (given the investor's own unit cost of information). Figures 4 and 5 illustrate this.

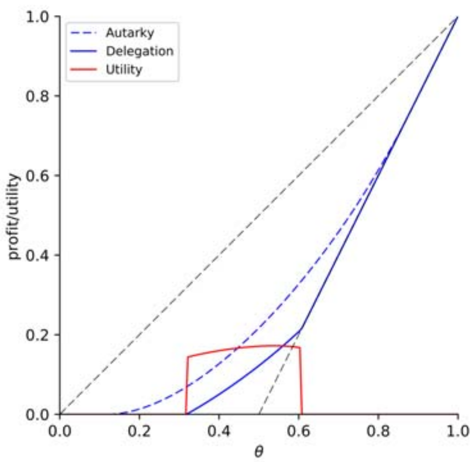

Figure 4: The principal's expected profit from delegation to a less able agent.

In Figure 4, the agent's unit cost of information is one fifth of that of the principal, $c_a = 0.1$. The diagram shows that there exists no contract that makes it worthwhile for the principal to hire such an agent: for all priors the principal is better off in autarky (the dashed curve lies above the solid curve). If the principal does not have the time or possibility to acquire information herself (equivalent to setting $c = +\infty$), then her *ex ante* expected profit would be zero for all $\theta \leq 1/2$ and rise linearly (along the steep dashed straight line) she would be willing to hire even this less able agent, but only for a relatively small range of priors (approximately $0.33 < \theta < 0.60$). Figure 5 shows the opposite situation when the agent is very able, when his unit cost is only $c_a = 0.01$. The diagram shows that such an agent is well worthwhile to hire for a wide range of priors.

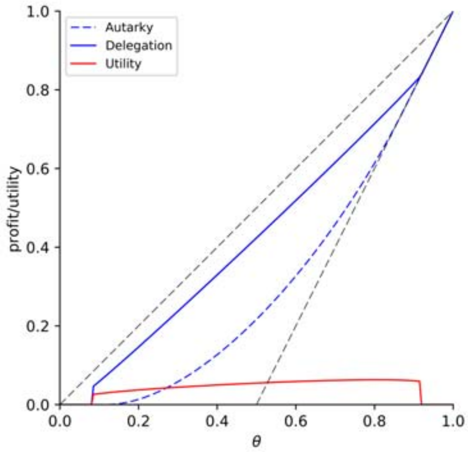

Figure 5: The principal's expected profit under delegation to a more able agent.

We note that the agent's expected utility if hired, as indicated in the above diagrams, is higher the *less* able the agent is. This may appear counter-intuitive. However, from Corollary 3 we know that the agent, when his participation constraint is not binding (as we here assume), will be paid nothing after a failed investment: $w_1 = 0$. In order to obtain information about the project at hand, a more able agent needs to exert less effort and so has to be paid less than a less able agent, who has to work hard to become well-informed. Hence, in the absence of a (binding) participation constraint, a smaller bonus is sufficient for the more able agent, at any given level of the penalty, and the penalty is equal to the salary. The above diagrams show that the principal benefits from the agent's increased ability (lower unit cost of information c_a). The optimal contracts, conditional upon hiring, in these three examples (evaluated at $\theta = 0.5$) are:

TABLE 1

c_a	salary	bonus	penalty	profit	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

Both expected profit and welfare (defined as the sum of utility and profit) increase with increased ability of the agent while the agent's expected utility decreases (in this range). The principal's expected profit in autarky (when her unit information cost is $c = 0.5$) is approximately 0.217. Hence, she will not offer any contract to the low-ability agent.

The present analysis is premised on the assumption that the principal knows the agent's type (preferences, 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. Hence, a high-ability agent with reservation utility not above 0.055 does not necessarily want to appear less able than he is. A low-ability agent might not want to appear as more able that he is either, since then the conditions of the contract might become too tough. 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. Hence, these are the critical reservation utilities for enabling the principal to use screening and self-selection in order to make only high-ability agents "bite". These observations can be generalized and would be worthwhile to study more in depth, but, in the interest of brevity we leave them for future research.

Returning to our complete-information setting, we also note the slight asymmetry, despite the project's symmetry: the bonus rates, defined as $(w_2 - y)/y$, are always larger than the penalty rate (which is always 100% due to the binding limited-liability constraint). This asymmetry is also apparent in Corollary 4, which establishes that optimal contracts are non-linear. In particular, incentive schemes containing a salary and a constant share of the project's net return may not be optimal. The reason is that, although such contracts will perfectly align the agent's incentives with those of the principal (had the principal had access to the agent's private information about the project at hand) at the moment of investment, they 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, so a share of the project return (such as stocks in the company) may be insufficient. As the bonus and penalty are increased in order to enhance the agent's information acquisition, the penalty will soon hit the limited-liability constraint. Indeed, for projects with binary return distributions (as in the above diagrams), the limited liability constraint is binding; $w_1 = 0$. 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 intact, also of the payment after a successful investment. Hence, if the agents is paid a fixed salary and some stocks, then some options (say, to buy future stocks at today's price) has to be added so as to boost the reward after successful investments.

In order to study the role of the participation constraint, one needs to solve the principal's problem when this constraint is binding. Figure 6, below (drawn for the same parameters as in Figure 3, and $\theta = 0.5$), illustrates precisely this.

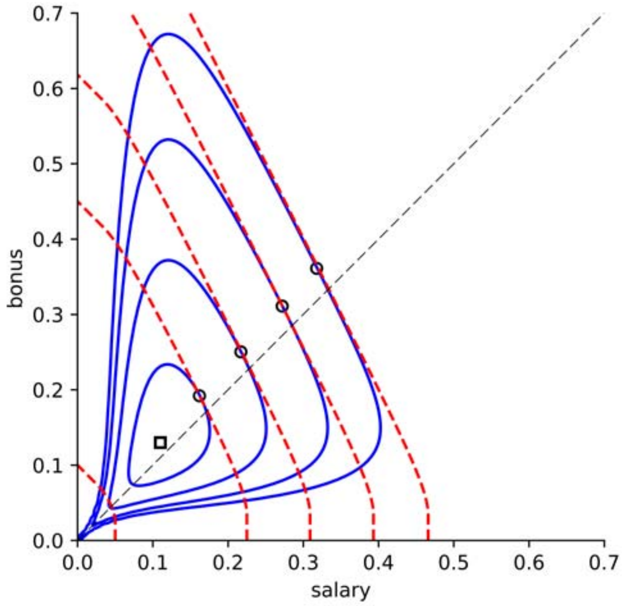

Figure 6: Iso-profit curves and iso-utility curves in contract space.

The diagram shows indifference curves for both parties in contract space, with the salary y (which is equal to the penalty) on the horizontal axis and the bonus, $b = w_2 - y$, on the vertical. The principal's iso-profit curves are the solid curves that form closed loops, while the agent's iso-utility curves are the dashed and negatively sloped curves. The principal's global optimum, in the absence of the agent's participation constraint, is indicated by the small square at approximately, $y^* \approx 0.110$ and $b^* \approx 0.129$, at which points the principal's expected profit is approximately 0.250 and the agent's expected utility approximately 0.144. Hence, if the agent's reservation utility is below this level, then his participation constraint is slack. The largest iso-profit curve in the diagram corresponds to zero expected profit, the profit the principal would earn in autarky if not acquiring information herself. If her unit cost of information acquisition would be $c = 0.5$ (the same as in Figures 1 and 3), then her expected profit in autarky would be approximately 0.217, corresponding to the smallest iso-profit curve in the diagram.

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 being 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 $\bar{u}$. These are the optimal contracts for the principal to offer. 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 ($x_1 = -1$, $r = 0$, $x_2 = 1$, and $\theta = 1/2$). This asymmetry of the optimal contracts, discussed in general above, implies that the agent's incentive, at the moment of investment (after he has acquired his private information) are not perfectly aligned with those of the investor. Under an optimal contract, the agent is thus willing to take a little more risk than the investor, to invest even when his private information is not sufficient to make investment rational for the investor, had the investor had that information. In other words, optimality (for the principal) requires that her desire to align the incentives at the investment decision have to be traded off against her desire to induce the agent to be well informed. In a sense, information is a public good for the two parties, so the agent has a tendency to under-invest in it unless compensated by the principal, who, however, cannot monitor the agent's effort and cannot increase the penalty without raising the salary.

Remark 1. *The non-convexity in the risk neutral agent's preferences as defined over contracts, exhibited in Figures 3-6, and to be seen also in the next section, is not too surprising in the light of previous research. Already Radner and Stiglitz (1984) showed that information, when treated as a good, leads to non-convexities. See also Chade and Schlee (2002) and Weibull, Mattsson and Voorneveld (2007).*

4. A SIGNAL-EXTRACTION MODEL

Are the qualitative results model specific, due to some "hidden" feature of the rational-inattention model framework? In order to investigate this issue we here compare the above results with those of an alternative model, a canonical model of signal-extraction. We elaborate this model only for projects with binary return distributions, risk neutral agents, and investments with binary returns. In the present framework, information takes the form of a noisy signal about the true return rate, and this noise is costly to reduce. We begin by considering an investor in autarky and then turn to delegation.[16]

[16]This comparative exercise differs from that in Fosgerau et al (2017) in that while they consider the choice of a single decision maker who faces finitely many alternatives, we here consider both autarky and agency, in simple binary investment problems, and also analyze implications for expected profits and utilities.

4.1. Investor in autarky. Just as in Sections 2-4, a risk-neutral and rational investor considers an indivisible investment opportunity, or *project*, that requires a lump-sum investment, $I > 0$, and gives a risky return Y. The probability distribution of the random net return $X = Y/I - 1$ is known by the investor and has support $M = \{x_1, x_2\}$. If the investor opts not to invest, then she obtains a sure net return rate $r = 0$. We assume that $x_1 < 0 < x_2$. If she decides to acquire information about the project, then she will receive a noisy *signal* S about its net return rate,

$$S = X + \varepsilon, \tag{33}$$

where the *noise* ε term is statistically independent of X. It is normally distributed with mean zero and variance $1/\tau > 0$. We call $\tau > 0$ the *signal precision.* While in the rational-inattention approach the decision-maker is free to design how to obtain information and what information to obtain, here she receives information as a signal, and can only affect the noisiness of the signal. More specifically, the investor's cost of obtaining any given degree of signal precision $\tau > 0$ is $C(\tau)$, where C is a strictly increasing 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 that the investor has obtained information, the realization of a signal S of precision $\tau > 0$. Her information costs being sunk, it is now optimal for her to invest if and only if the received signal is such that the associated conditionally expected net return rate is nonnegative,

$$\mathbb{E}[X \mid S = s] \geq 0. \tag{34}$$

Since the simple signal structure (33) meets the monotone likelihood ratio property (MLRP), this inequality can be written equivalently as

$$S \geq \hat{s}(\tau) \tag{35}$$

for some $\hat{s}(\tau) \in \mathbb{R}$, the optimal (finite or infinite) *signal threshold.*[17] This threshold defines an optimal *investment strategy*, given any signal precision $\tau > 0$ that the investor's signal may have.

[17] 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 Lemma 1 in Appendix A.

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 so that it maximizes her *ex ante* expected profit, the expected financial profit net of information costs,

$$\tilde{\Pi}(\tau) = V(\tau) - C(\tau), \tag{36}$$

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

$$V(\tau) \;=\; I \cdot \mathbb{E}[X \mid S > \hat{s}(\tau)] \cdot \Pr[S \geq \hat{s}(\tau)] \tag{37}$$

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).

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,

$$V'(\tau) = C'(\tau). \tag{38}$$

We will say that nature is in *the good state* when $X = x_2$ and in *the bad state* when $X = x_1$, and write $\theta = \theta(x_2)$ for the prior probability for the good state, just as in the binary case studied in Section 3 above. As shown in Lemma 1 in Appendix B, the investor's optimal signal threshold is

$$\hat{s}(\tau) \;=\; \frac{x_2 + x_1}{2} - \frac{\ln \hat{\rho}}{(x_2 - x_1)\tau}, \tag{39}$$

where

$$\hat{\rho} = \frac{\theta}{1-\theta} \cdot \frac{x_2}{|x_1|}, \tag{40}$$

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", $(1-\theta)|x_1|$. Equation (39) 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 is not difficult to verify (see Appendix B) that the associated marginal value of information can be written as

$$V'(\tau) = \frac{x_2 - x_1}{2} \cdot \frac{I \cdot \hat{\kappa}}{\sqrt{2\pi\tau}} \cdot \exp\left[-\frac{1}{2\tau}\left(\frac{\ln \hat{\rho}}{x_2 - x_1}\right)^2 - \frac{\tau}{2}\left(\frac{x_2 - x_1}{2}\right)^2\right], \tag{41}$$

where

$$\hat{\kappa} = I\sqrt{\theta(1-\theta)\, x_2 |x_1|}, \tag{42}$$

a quantity we will 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 (41) 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, is largest for projects with unit risk balance. This is 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

$$C(\tau) = \gamma \cdot \int_0^\tau \frac{1}{\sqrt{t}} \exp\left(\alpha t - \frac{\beta}{t}\right) dt, \tag{43}$$

for $\alpha, \beta, \gamma > 0$. The associated marginal information cost, $C'(\tau)$, is positive at all positive signal precisions, running from zero as zero signal precision towards plus infinity as precision tends to plus infinity. Indeed, C is convex if and only if $16\alpha\beta \geq 1$, and it is always convex on the interval $\left[0, \left(1 - \sqrt{1 - 16\alpha\beta}\right)/4\alpha\right]$, see Appendix B. For moderate ranges of signal precision, the graph of these cost functions can be made almost indistinguishable from more conventional cost functions such as $C(\tau) = \tau^2$.[18]

[18] Arguably, this class of cost functions are close to canonical cost functions in economics. An

For cost functions of the form (43), the optimal signal precision can be solved explicitly. Let

$$\hat{K} = \ln \hat{\kappa} + \ln \left(\frac{x_2 - x_1}{2} \right) - \ln \left(\gamma \sqrt{2\pi} \right), \tag{44}$$

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

$$\hat{D} = \hat{K}^2 - \left[\left(\frac{x_2 - x_1}{2} \right)^2 + 2\alpha \right] \cdot \left[\left(\frac{\ln \hat{\rho}}{x_2 - x_1} \right)^2 - 2\beta \right] \tag{45}$$

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

$$\hat{\tau} = \frac{\hat{K} + \sqrt{\hat{D}}}{\left(x_2 - x_1\right)^2 / 4 + 2\alpha}. \tag{46}$$

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 3. *Suppose that the cost function is of the form (43). 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 $\hat{D} \le 0$, 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 (46) if either $\hat{K} \ge 0$ or

$$\hat{K} < 0 \quad \text{and} \quad e^{-(x_2 - x_1)\sqrt{2\beta}} < \hat{\rho} < e^{(x_2 - x_1)\sqrt{2\beta}}. \tag{47}$$

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

alternative to these functions would be to use entropy also here. The noise term in our model is assumed to be normally distributed and the entropy of the normal distribution is an increasing function of its variance. Hence, a decreasing function of signal precision τ. However, to use such a cost function would weaken our robustness results since it would bring the signal extraction model closer to the rational inattention model.

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 3, 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, $\hat{\rho} = \theta / (1 - \theta)$, and the *riskiness is* the geometric mean of the probabilities for the two states of nature, $\hat{\kappa} = \sqrt{\theta (1 - \theta)}$. Moreover, $K \geq 0$ if and only if $\gamma \leq \sqrt{\theta (1 - \theta) / 2\pi}$. This inequality holds if the information-cost parameter γ is small and/or the prior θ is close to one half. In particular, $\hat{K} < 0$ for all $\theta \in [0, 1]$ 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

$$\left(\frac{\ln \theta - \ln (1 - \theta)}{2} \right)^2 < 2\beta. \tag{48}$$

For $\beta = 1/2$, this is identical with the definition of (strictly) normal investment conditions in the rational-inattention model when applied this example (for $c = 1/2$).[19]

The diagram below shows the investor's expected profit (solid curve) as a function of the prior θ for the good state of nature in the present signal-extraction model for $\alpha = 0.1$ and $\beta = 0.5$, and $\gamma = 0.25$. Comparing this with the investor's value function in the rational-inattention model (dashed curve, the same as in Figure 1), we again note model robustness.

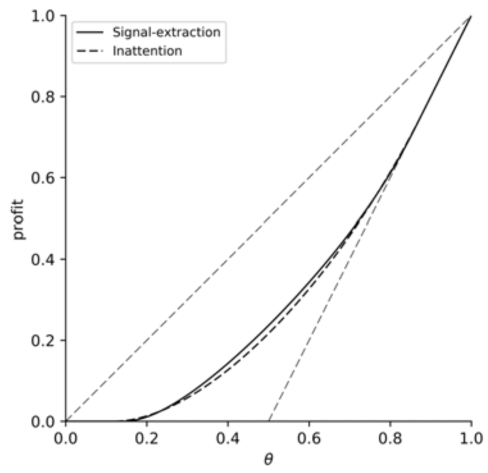

Figure 7: The investor's expected profit in the signal-extraction model.

[19]To see this, note that (48) then is equivalent with $-2 < \ln [\theta / (1 - \theta)] < 2$, or $e^{-2} < \theta / (1 - \theta) < e^2$, and compare with footnote 12.

4.2. 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, 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. It consists of a non-negative *pay schedule,* $\boldsymbol{w}$, which specifies the agent's remuneration conditional on the realized net return of the investment, if made, and a non-negative pay, y, to the agent if the latter decides not to invest. The agent's remuneration is thus a random variable,

$$\tilde{Y} = \begin{cases} \boldsymbol{w}(X) & \text{if investment is made} \\ y & \text{otherwise} \end{cases}. \tag{50}$$

The non-negativity requirement, the *limited-liability constraint*, rules out contracts $\langle y, \boldsymbol{w} \rangle$ by which the agent may (with positive probability) have to make a net payment to the principal. Formally, $\boldsymbol{w}(X) \geq 0$ with probability one and $y \geq 0$.

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

$$\tilde{U} = u\left(\tilde{Y}\right) - C_a(\tau_a), \tag{51}$$

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 $\tau_a \geq 0$. Both functions, u and C_a, are twice differentiable, and we assume that $u', C_a' > 0$, $u'' \leq 0$, and $C_a'' \geq 0$. Zero signal precision is costless, $C_a(0) = 0$. 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.

In the binary case that we here consider, any contract $\langle y, \boldsymbol{w}\rangle$ between the investor-*cum*-principal and agent boils down to only three payments to the agent; his "salary" y if he does not invest, his pay $w_2 = \boldsymbol{w}(x_2)$ if investing in the good state, and his pay $w_1 = \boldsymbol{w}(x_1)$ if investing in the bad state. We focus on contracts that are *strictly monotonic* in the sense that they pay most after a successful investment and least after a failed investment; $w_1 < y < w_2$.

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,

$$\mathbb{E}\left[\tilde{U} \mid S_a = s\right] \geq u(y). \tag{52}$$

Equality in this condition uniquely determines a signal threshold, $s^*(\boldsymbol{w}, \tau_a) \in \mathbb{R}$, 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 (Lemma 3 in Appendix B) to be

$$s^*(\boldsymbol{w}, \tau_a) = \frac{x_2 + x_1}{2} - \frac{\ln \rho(\boldsymbol{w})}{(x_2 - x_1)\tau_a}. \tag{53}$$

where

$$\rho(\boldsymbol{w}) = \frac{\theta}{1-\theta} \cdot \frac{u(w_2) - u(y)}{u(y) - u(w_1)}, \tag{54}$$

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 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:

$$\frac{x_2 - x_1}{2} \cdot \frac{\kappa(y, \boldsymbol{w})}{\sqrt{2\pi\tau_a}} \cdot \exp\left[-\frac{(x_2 - x_1)^2 \tau_a}{8} - \left(\frac{\ln \rho(\boldsymbol{w})}{x_2 - x_1}\right)^2 \frac{1}{2\tau_a}\right] = C_a'(\tau_a), \tag{55}$$

where

$$\kappa(y, \boldsymbol{w}) = \sqrt{\theta(1-\theta)(u(w_2) - u(y))(u(y) - u(w_1))} \tag{56}$$

(see Appendix B). We will call $\kappa(y, \boldsymbol{w})$ the contract's *power*; another useful concept in our analysis It is proportional to the product of the agent's two utility gains from "doing the right thing" in each state of nature. The proportionality factor is larger the more uncertain the project is *ex ante* (before the agent's signal has been observed). 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 (55) is formally identical with the necessary first-order condition for the investor's choice of signal precision in autarky, (41). 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 $\kappa(y, \boldsymbol{w})$.

The left-hand side of (55) is the agent's marginal increase of his expected remuneration utility from a marginal increases in his signal precision. The right-hand side is his marginal disutility of raising his signal precision. One sees that the marginal remuneration utility to the agent is higher the higher is the power, $\kappa(y, \boldsymbol{w})$, of his contract, *ceteris paribus*. One also sees that the agent's marginal remuneration utility is non-monotonic in the carrot-stick ratio, $\rho(\boldsymbol{w})$, of the contract. It is maximal when the carrot-stick ratio is 1. We also note that it is necessary that the contract contains both a bonus and a penalty, that $w_1 < y < w_2$. For otherwise the contract would have zero power, in which case the left-hand side in (55) would vanish, which would induce the agent to acquire no information at all.

Suppose that the agent's disutility or cost of information acquisition, C_a, belongs to the same family of cost functions as that of the investor, that is, is of the form (43), for some $\alpha_a, \beta_a, \gamma_a > 0$. The agent's optimal signal precision under any strictly monotonic contract can then $\boldsymbol{w} \in W$ be found by an application of Proposition 3, with $\hat{K}$ and $\hat{D}$ replaced by

$$K(y, \boldsymbol{w}) = \ln \kappa(\boldsymbol{w}) + \ln\left(\frac{x_2 - x_1}{2}\right) - \ln\left(\gamma_a \sqrt{2\pi}\right) \tag{57}$$

and

$$D(y, \boldsymbol{w}) = [K(y, \boldsymbol{w})]^2 - \left[\left(\frac{x_2 - x_1}{2}\right)^2 + 2\alpha_a\right] \cdot \left[\left(\frac{\ln \rho(\boldsymbol{w})}{x_2 - x_1}\right)^2 - 2\beta_a\right], \tag{58}$$

respectively. For contracts with $D(y, \boldsymbol{w}) > 0$, let

$$\tau_a^*(y, \boldsymbol{w}) = \frac{K(y, \boldsymbol{w}) + \sqrt{D(y, \boldsymbol{w})}}{(x_2 - x_1)^2 / 4 + 2\alpha_a}. \tag{59}$$

For any contract $\langle y, \boldsymbol{w} \rangle$ and signal precision $\tau_a > 0$, write $U(y, \boldsymbol{w}, \tau_a)$ for the agent's expected utility when acquiring signal precision τ_a and employing his optimal investment strategy for this signal precision under contract $\langle y, \boldsymbol{w} \rangle$.

Corollary 5. *Suppose that the agent's cost function is of the form (89). If $D(y, \boldsymbol{w}) > 0$, $\tau_a^*(y, \boldsymbol{w}) > 0$ and $U(y, \boldsymbol{w}, \tau_a^*(y, \boldsymbol{w})) > \bar{u}$, then the agent will choose signal precision $\tau_a^*(y, \boldsymbol{w})$. Otherwise, the agent will choose signal precision zero.*

Clearly it is not in the interest of the principal to offer the agent a contract under which the agent will not acquire any information. Hence, it is necessary that the contract satisfies $D(y, \boldsymbol{w}) > 0$, or, equivalently,

$$\left[\ln\left(\frac{x_2 - x_1}{2\gamma_a\sqrt{2\pi}}\right) + \ln\kappa(\boldsymbol{w})\right]^2 > \left[\left(\frac{x_2 - x_1}{2}\right)^2 + 2\alpha_a\right] \cdot \left[\left(\frac{\ln\rho(\boldsymbol{w})}{x_2 - x_1}\right)^2 - 2\beta_a\right], \tag{60}$$

and also that $K(y, \boldsymbol{w}) + \sqrt{D(y, \boldsymbol{w})} > 0$, or, equivalently, either

$$\ln\kappa(\boldsymbol{w}) \geq \ln\left(\gamma_a\sqrt{2\pi}\right) - \ln\left(\frac{x_2 - x_1}{2}\right) \tag{61}$$

or

$$\ln\kappa(\boldsymbol{w}) < \ln\left(\gamma_a\sqrt{2\pi}\right) - \ln\left(\frac{x_2 - x_1}{2}\right) \quad \text{and} \quad \left(\frac{\ln\rho(\boldsymbol{w})}{x_2 - x_1}\right)^2 < 2\beta_a. \tag{62}$$

These observations suggest a way ahead for an analysis of the delegation problem that the investor faces. First, let W_0 denote the set of feasible contracts $\langle y, \boldsymbol{w} \rangle$ that satisfy (60) and either (61) or (62). This will be called the set of feasible and *potentially profitable contracts* for the principal, the contracts under which the agent will acquire some information. For every such contract $\langle y, \boldsymbol{w} \rangle$ the agent will choose the positive signal precision $\tau_a^*(y, \boldsymbol{w})$ defined in (59) and subsequently use his optimal signal threshold $s^*((y, \boldsymbol{w}), \tau_a)$, defined for all signal precisions $\tau_a > 0$ in (53). Anticipating this, the principal expects the profit

$$\tilde{\Pi}(y, \boldsymbol{w}) = I(x_2 - w_2 + y) \cdot p_G(y, \boldsymbol{w}) + I(x_1 - w_1 + y) \cdot p_B(y, \boldsymbol{w}) - y, \tag{63}$$

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.

As shown in Appendix B, for $\langle y, \boldsymbol{w} \rangle \in W_0$ equation (63) can be written explicitly in terms of the primitives of the model (see Lemma 3). Let $W_1 \subseteq W_0$ be the (potentially empty) set of contracts in W_0 that meet the agent's participation constraint.

If W_1 is empty, then the principal will offer no contract to the agent. If W_1 is nonempty and $\sup_{w \in W_1} \tilde{\Pi}^* (y, \boldsymbol{w}) > 0 = r$, then there are contracts that are profitable to the principal and acceptable by the agent. In other words, then there exist gains of trade between the two parties. If the supremum profit among those contracts is in fact achieved by some contract, then the principal will offer any such contract. If the supremum profit is not achieved by any contract in W_1 (which is *a priori* possible, since W_1 is not a compact set), then for every $\varepsilon > 0$ there will exist a nonempty subset of contracts in W_1 under which the principal's expected profit is within ε from the supremum profit. Hence, there will then exist contracts that are ε-optimal for the principal, for arbitrarily small $\varepsilon > 0$.

Suppose that the agent is risk neutral. Then the agent's signal threshold is a function of the bonus and penalty, and so is the agent's optimal signal precision. Hence, once hired, the agent's behavior is completely determined by the bonus and penalty. 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 result in a gain to the investor, compared with doing without an agent? The diagram below has been created for the same project as used in the rational-attention model, in Figure 6.

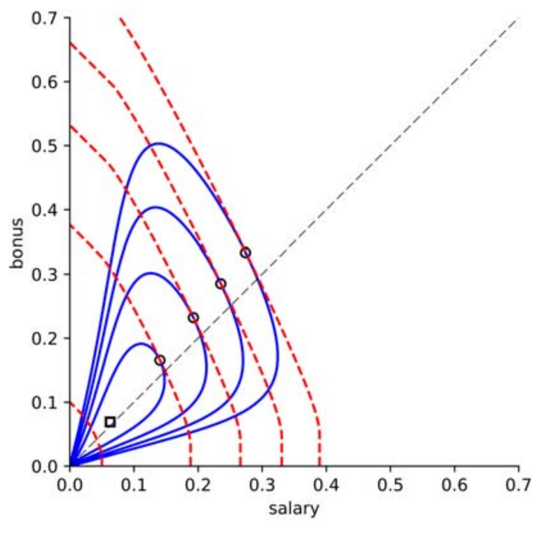

Figure 8: Iso-profit and iso-utility curves in contact space.

The information-cost parameters for the agent are here $\gamma_a = 0.025$, $\alpha_a = 0.1$ and $\beta_a = 0.5$. It turns out that if the principal would have the same parameters, except for a ten times higher cost parameter $\gamma = 0.25$ (the same value as in Figure 7), then there would be no gains of trade between the two parties. Hiring the agent under the globally optimal contract (indicated by the little square in the diagram) would result in expected profit to the principal of approximately 0.183 while the principal's expected profit in autarky would be about 0.236.[20] Hence, for these parameter values, the signal-extraction model delivers less gains of trade than the rational-inattention model; a quantitative difference between the two models. By contrast, the isoquants, solid for the principal, dashed for the agent, are not unlike for those in the rational-inattention model. Yet another qualitative similarity between the two models.

5. DISCUSSION

The rational inattention model was here developed and analyzed for projects with finitely many potential outcomes. However, the analysis also applies to return rates with much more general probability distributions. The present model also allows generalization from one investment project to any finite number of investment projects. Such generalizations are much harder in the signal-extraction model, which we also found harder to work with even in the special case of a single investment project with binary outcomes. We showed that the two modelling approaches gave qualitatively similar results in this special case. The main analytical advantage of the rational-inattention model over the signal-extraction model is that instead of having to nest the agent's optimization problem within the principal's maximization problem, the agent's optimal information-cum-investment strategy is already represented in the principal's goal function, as the agent's *ex ante* expected investment probability, which uniquely determines his conditional investment probabilities in all states of nature.

For each of the two models, we identified three "investment conditions". Rational investors invest "blindly" in "sweet" investment conditions, that is, they do not acquire any additional information but rely entirely on their prior beliefs. Under such conditions, investors only learn by experience, after they have invested and investments start to yield returns. Hence, in an ongoing economy with many investors of the kind studied here, a gradual deterioration in investment conditions (which may be correlated across investors and projects) will then only be observed with a delay. Under "sour" investment conditions, there is also no information acquisition, but now

[20] The globally optimal contract for the principal is $y \approx 0.063$ and $w_2 \approx 0.132$, and hence bonus rate $\approx 110\%$, numbers that are comparable with those for the rational-inattention model, see Table 1.

there is also no experience accumulated, not even with a delay since no investment takes place. Hence, even if investment conditions would improve, this may not be noticed by rational investors. Hence, it may take a long time to get an economy out of sour investment conditions, much longer and slower than the time to get out of sweet investment conditions. Under normal investment conditions, however, investors quickly notice even small changes in the environment since they acquire (private) information before they make their investment decisions, and condition their investment decisions on the obtained information. Hence, in such conditions, but only then, do investors "know their business". These asymmetries between the three conditions, and transitions between them, may play some role for understanding the dynamics of business cycles, a topic for future study.

The present model builds upon many heroic simplifications. A relevant but challenging extension of our model would be to allow for incomplete information about the agent's talent (unit information cost), risk attitude (Bernoulli function of income), and/or outside option. Can a principal then use screening to let agents self-select a contract? This depends crucially upon whether agents know their own type or not. It would be particularly risky if some agents have inflated beliefs about their own talent (that is, underestimate their unit cost of information and/or overestimate the precision of their signal). In such cases, a wise and experienced principal may note a candidate's biased beliefs and may actually gain from hiring an over-confident agent, since such an agent may be willing to accept a tougher contract. Another interesting extension would be to consider agents with career concerns, and/or social or moral preferences such as loyalty with the principal or a wish to "do the right thing". However, in order to analyze any of these richer and more realistic cases, it seems necessary to first understand the simpler case of a single agent whose type is known by both parties. This is precisely the task we have here undertaken.

6. Appendix A: The rational inattention model

We first establish that equation (4) is necessary for optimality. Taking the derivative of the maximand in (17) with respect to q, one obtains

$$\sum_{x\in M} \frac{\theta(x)\cdot e^{xI/c}}{\hat{q}e^{xI/c}+(1-\hat{q})\,e^{rI/c}} = \sum_{x\in M} \frac{\theta(x)\cdot e^{rI/c}}{\hat{q}e^{xI/c}+(1-\hat{q})\,e^{rI/c}}. \tag{64}$$

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

$$(1-\hat{q})\sum_{x\in M} \frac{\theta(x)\cdot e^{xI/c}}{\hat{q}e^{xI/c}+(1-\hat{q})\,e^{rI/c}} = \sum_{x\in M} \theta(x)\cdot\frac{(1-\hat{q})\,e^{rI/c}}{\hat{q}e^{xI/c}+(1-\hat{q})\,e^{rI/c}}. \tag{65}$$

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

$$\frac{(1-\hat{q}) e^{rI/c}}{\hat{q}e^{xI/c} + (1-\hat{q}) e^{rI/c}} = 1 - \hat{p}_{|X=x}. \tag{66}$$

for all $x \in M$. Multiplying both sides of the latter equation by $\theta(x)$ and summing over all $x \in M$, and using the identity $\sum_{x\in M} \theta(x) \hat{p}_{|X=x} = \hat{q} f(\hat{q})$, we obtain that the right-hand side of (65) equals $1 - \hat{q} f(\hat{q})$. Hence, $(1-\hat{q}) f(\hat{q}) = 1 - \hat{q} f(\hat{q})$, or $f(\hat{q}) = 1$.

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

$$f'(q) = -\sum_{x\in M} \frac{\theta(x) e^{xI/c} \left(e^{xI/c} - e^{rI/c}\right)}{\left[qe^{xI/c} + (1-q) e^{rI/c}\right]^2} \quad \forall q \in [0,1] \tag{67}$$

and

$$f''(q) = 2\sum_{x\in M} \frac{\theta(x) e^{xI/c} \left(e^{xI/c} - e^{rI/c}\right)^2}{\left[qe^{xI/c} + (1-q) e^{rI/c}\right]^3} \geq 0 \ \forall q \in [0,1]. \tag{68}$$

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 $\mathbb{E}\left[e^{(r-X)I/c}\right] \leq 1$.

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

$$\ln \mathbb{E}\left[e^{XI/c}\right] + \ln \mathbb{E}\left[e^{-XI/c}\right] > 0. \tag{69}$$

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

$$\ln \mathbb{E}\left[e^{XI/c}\right] + \ln \mathbb{E}\left[e^{-XI/c}\right] > \mathbb{E}\left[\ln e^{XI/c}\right] + \mathbb{E}\left[\ln e^{-XI/c}\right] = I \cdot (\mathbb{E}[X/c] - \mathbb{E}[X/c]) = 0. \tag{70}$$

6.3. 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

$$g'(q) = -\sum_{x \in M} \frac{\theta(x)\, e^{u(\boldsymbol{w}(x))/c_a} \left(e^{u(\boldsymbol{w}(x))/c_a} - e^{u(y)/c_a}\right)}{\left[q e^{u(\boldsymbol{w}(x))/c_a} + (1-q)\, e^{u(y)/c_a}\right]^2} \tag{71}$$

and

$$g''(q) = 2\sum_{x \in M} \frac{\theta(x)\, e^{u(\boldsymbol{w}(x))/c_a} \left(e^{u(\boldsymbol{w}(x))/c_a} - e^{u(y)/c_a}\right)^2}{\left[q e^{u(\boldsymbol{w}(x))/c_a} + (1-q)\, e^{u(y)/c_a}\right]^3} \geq 0 \tag{72}$$

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.

6.4. Proof of Proposition 2. Let $\langle y, \boldsymbol{w} \rangle$ be a contract that solves program (25) for a risk neutral agent. By the Karush-Kuhn-Tucker theorem, there exists a non-negative scalar λ, a Lagrangian associated with the agent's participation constraint, such that the following equation holds for all $k \in M$ with $w_k > 0$:

$$\frac{d\Pi(y, \boldsymbol{w})}{dw_k} + \lambda \cdot \frac{dU(y, \boldsymbol{w})}{dw_k} = 0. \tag{73}$$

Here

$$\frac{d\Pi(y, \boldsymbol{w})}{dw_k} = \frac{\partial \Pi(y, \boldsymbol{w})}{\partial w_k} + \frac{d\Pi(y, \boldsymbol{w})}{dq^*} \cdot \frac{dq^*}{dw_k}, \tag{74}$$

where the first term is the direct effect and the second term the indirect effect via the *ex ante* investment probability q^*. By the envelope theorem (applied to the maximization program in Corollary 1 of Matějka and McKain, 2015), the derivative of the agent's expected utility with respect to payment w_k only contains the direct effect:

$$\frac{dU(y, \boldsymbol{w})}{dw_k} = \frac{\partial V(y, \boldsymbol{w})}{\partial w_k}, \tag{75}$$

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

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

Hence, the necessary condition (73) can be written as

$$(1-\lambda) \cdot \frac{\partial \Pi(y, \boldsymbol{w})}{\partial w_k} + \frac{d\Pi(y, \boldsymbol{w})}{dq^*} \cdot \frac{dq^*}{dw_k} = 0. \tag{76}$$

Let $g(q, \boldsymbol{w})$ denote the continuously differentiable left-hand side of (28), and recall that $\partial g(q, \boldsymbol{w})/\partial q \neq 0$ at $q = q^*$. Hence, by the implicit-function theorem, equation (28), written as $g(q^*, \boldsymbol{w}) = 1$, uniquely defines q^* as a differentiable function of w_k (on an open neighborhood around the initial point $(\boldsymbol{w}, q^*)$), such that

$$\frac{dq^*}{dw_k} = -\frac{\partial g(q^*, \boldsymbol{w})}{\partial w_k} \cdot \left(\frac{\partial g(q^*, \boldsymbol{w})}{\partial q^*}\right)^{-1}. \tag{77}$$

Hence, (76) can be written as

$$(1-\lambda) \cdot \frac{\partial \Pi(y, \boldsymbol{w})}{\partial w_k} - \rho \cdot \frac{\partial g(q^*, w_k)}{\partial w_k} = 0 \tag{78}$$

where

$$\rho = \frac{d\Pi(y, \boldsymbol{w})}{dq^*} \cdot \left(\frac{\partial g(q^*, \boldsymbol{w})}{\partial q^*}\right)^{-1} \tag{79}$$

is the same for all $k \in M$ with $w_k > 0$. Moreover,

$$\begin{aligned} \frac{\partial \Pi(y, \boldsymbol{w})}{\partial w_k} &= \theta_k \cdot \frac{(1-q^*) q^* [I(x_k - r) + y - w_k] e^{(y-w_k)/c_a}}{c_a \cdot [q^* + (1-q^*) e^{(y-w_k)/c_a}]^2} \\ &\quad -\theta_k \cdot \frac{q^* \left[q^* + (1-q^*) e^{(y-w_k)/c_a}\right]}{[q^* + (1-q^*) e^{(y-w_k)/c_a}]^2}, \end{aligned} \tag{80}$$

and

$$\frac{\partial g(q^*, w_k)}{\partial w_k} = \theta_k \cdot \frac{(1-q^*) e^{(y-w_k)/c_a}}{c_a \cdot [q^* + (1-q^*) e^{(y-w_k)/c_a}]^2}, \tag{81}$$

so (78) can be written as

$$\begin{aligned} 0 &= (1-\lambda)(1-q^*) q^* [I(x_k - r) + y - w_k] e^{(y-w_k)/c_a} \\ &\quad - (1-\lambda) c_a q^* \left[q^* + (1-q^*) e^{(y-w_k)/c_a}\right] \\ &\quad -\rho \cdot (1-q^*) e^{(y-w_k)/c_a} \end{aligned}$$

or

$$q^* [I(x_k - r) + y - w_k] - c_a \frac{q^*}{1-q^*} \left[q^* e^{(w_k - y)/c_a} + (1-q^*)\right] = \frac{\rho}{1-\lambda}$$

or

$$Ix_k - w_k - c_a \frac{1}{1-q^*} \left[q^* e^{(w_k-y)/c_a}\right] = Q$$

where Q is the same for all $k \in M$ with $w_k > 0$. This gives (30).

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

$$Ix = w(x) + P \cdot e^{w(x)/c_a} + Q \tag{82}$$

or, equivalently,

$$e^{w(x)/c_a} = \frac{Ix - Q - w(x)}{P} \tag{83}$$

for

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

Equation (83) 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 principal branch W_0 of the (multi-valued) Lambert W function, implicitly defined for all $z > -1/e$ by $y = W_0(z)$ for $ye^y = z$. First, consider the following simple transcendental algebraic equation in $w \in \mathbb{R}$:

$$e^{-w/c} = a(w - x) \tag{84}$$

for $a, c, x \in \mathbb{R}$. 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 $z = e^{-x/c}/(ac)$. Hence, by definition of the Lambert W function,

$$W_0\left(\frac{e^{-x/c}}{ac}\right) = \frac{w - x}{c},$$

or

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

which thus solves (84). It follows that the similar equation $e^{w/c} = a(x - w)$ has solution

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

Accordingly, a solution of (83) is

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

which establishes (32) if A is such that

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

or

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

Hence, $A > 0$ under normal investment conditions, which are necessary for optimality of the contract.

7. APPENDIX B: THE SIGNAL-EXTRACTION MODEL

We begin by establishing a lemma, that states that the optimal signal threshold is always finite (for normally distributed noise) and can be written in the form (39). The lemma also provides formulae for the two conditional investment probabilities. Assume that the noise term ε is normally distributed with mean value zero and variance $\sigma^2 = 1/\tau$. As is well-known, its PDF

$$\phi_\tau(z) = \sqrt{\frac{\tau}{2\pi}} \cdot e^{-\tau z^2/2}$$

meets the MLRP, for every $\tau > 0$, and its CDF is

$$\Phi_\tau(z) = \sqrt{\frac{\tau}{2\pi}} \int_{-\infty}^{z} e^{-\tau y^2/2} dy$$

We will call $\tau > 0$ the signal precision. We write $\phi = \phi_1$ and $\Phi = \Phi_1$ for the PDF and CDF of the standard normal distribution, $\mathcal{N}(0,1)$. Let $x_1 < 0 < x_2$ be the possible values of X and write θ for the probability that $X = x_2$, the "good" outcome.

Lemma 1. *The optimal signal threshold for the investor in autarky is*

$$\hat{s}(\tau) = \frac{x_2 + x_1}{2} - \frac{\ln \hat{\rho}}{(x_2 - x_1)\tau}. \tag{85}$$

The conditional investment probability in the "good" state of nature ($X = x_2$) is

$$p_2(\tau) = \Phi\left(\frac{\ln \hat{\rho}}{(x_2 - x_1)\sqrt{\tau}} + \frac{x_2 - x_1}{2}\sqrt{\tau}\right), \tag{86}$$

and the conditional investment probability in the "bad" state of nature ($X = x_1$) is

$$p_1(\tau) = (1-\theta) \cdot \Phi\left(\frac{\ln \hat{\rho}}{(x_2 - x_1)\sqrt{\tau}} - \frac{x_2 - x_1}{2}\sqrt{\tau}\right). \tag{87}$$

Proof: For any $s \in \mathbb{R}$ and any $m > 1$:

$$\mathbb{E}[X \mid S = s] = \sum_{i=1}^{m} x_i \cdot \mathbb{P}[X = x_i \mid S = s].$$

For any $\delta > 0$, and writing $\theta_i = \mathbb{P}(X = x_i)$:

$$\mathbb{P}[X = x_i \mid S \in (s-\delta, s+\delta)] = \frac{\theta_i \mathbb{P}[S \in (s-\delta, s+\delta) \mid X = x_i]}{\sum_{k=1}^{m} \theta_k \mathbb{P}[\in (\hat{s}-\delta, \hat{s}+\delta) \mid X = x_k]}.$$

Moreover,

$$\mathbb{P}[S \in (s-\delta, s+\delta) \mid X = x_i] = \mathbb{P}[\varepsilon \in (s - x_i - \delta, s - x_i + \delta) \mid X = x_i]$$

For small $\delta > 0$,

$$\mathbb{P}[\varepsilon \in (s - x_i - \delta, s - x_i + \delta) \mid X = x_i] \approx 2\delta \cdot \phi_\tau (s - x_i)$$

Hence, as $\delta \downarrow 0$,

$$\mathbb{P}[X = x_i \mid S \in (s-\delta, s+\delta)] \to \frac{\theta_i \phi_\tau (s - x_i)}{\sum_{k=1}^{m} \theta_k \phi_\tau (s - x_k)}$$

In the case of the normal distribution,

$$\begin{aligned} \frac{\theta_i \phi_\tau (s - x_i)}{\sum_{k=1}^{m} \theta_k \phi_\tau (s - x_k)} &= \frac{\theta_i \exp\left[-\tau (s - x_i)^2 /2\right]}{\sum_{k=1}^{m} \theta_k \exp\left[-\tau (s - x_k)^2 /2\right]} \\ &= \frac{\theta_i \exp[\tau (s - x_i/2) x_i]}{\sum_{k=1}^{m} \theta_k \exp[\tau (s - x_k/2) x_i]} \end{aligned}$$

Hence,

$$\mathbb{E}[X \mid S = s] = \sum_{i=1}^{m} \frac{x_i \theta_i \exp[\tau (s - x_i/2) x_i]}{\sum_{k=1}^{m} \theta_k \exp[\tau (s - x_k/2) x_i]}. \tag{88}$$

For $m = 2$, and writing θ for θ_2:

$$\mathbb{E}[X \mid S = s] = \frac{(1-\theta) x_1 \exp[\tau (s - x_1/2) x_1] + \theta x_2 \exp[\tau (s - x_2/2) x_2]}{(1-\theta) \exp[\tau (s - x_1/2) x_1] + \theta \exp[\tau (s - x_2/2) x_2]}$$

If $0 < \theta < 1$ and $x_1 < 0 < x_2$:

$$\lim_{s \to -\infty} \mathbb{E}[X \mid S = s] = x_1 < 0 < x_2 = \lim_{s \to +\infty} \mathbb{E}[X \mid S = s]$$

Hence, for any $\tau > 0$, the optimal signal threshold for the investor is a real number, $\hat{s}(\tau)$. Moreover, for $s = \hat{s}(\tau)$, $\mathbb{E}[X \mid S = s] = 0$ iff

$$(1-\theta) x_1 \exp[\tau(s - x_1/2) x_1] + \theta x_2 \exp[\tau(s - x_2/2) x_2] = 0$$

or, equivalently,

$$\exp[\tau(s - x_1/2) x_1] = \frac{\theta x_2}{(1-\theta)|x_1|} \exp[\tau(s - x_2/2) x_2]$$

or

$$\exp[\tau(s - x_1/2) x_1 - \tau(s - x_2/2) x_2] = \frac{\theta x_2}{(1-\theta)|x_1|}$$

or

$$\exp\left[-\tau s(x_2 - x_1) + \tau\left(x_2^2 - x_1^2\right)/2\right] = \frac{\theta x_2}{(1-\theta)|x_1|}$$

or, taking the logarithm of both sides,

$$\tau s(x_2 - x_1) = \tau\left(x_2^2 - x_1^2\right)/2 - \ln\hat{\rho}$$

which gives (85).

The conditional investment probabilities, for any investment signal threshold s (optimal or not), is

$$p_X = \mathbb{P}[S > s \mid X] = 1 - \Phi_\tau(s - X).$$

For the optimal signal threshold with respect to normally distributed noise, this gives

$$p_X = \sqrt{\frac{\tau}{2\pi}} \cdot \int_{\hat{s}(\tau)-X}^{+\infty} e^{-\tau y^2/2} dy,$$

which, after some algebraic manipulation results in the claimed conditional probabilities for $X = x_1$ and $X = x_2$ (and exploiting the symmetry of ϕ_1):

$$\begin{aligned} p_2(\tau) &= \theta\sqrt{\frac{\tau}{2\pi}} \cdot \int_{\hat{s}(\tau)-x_2}^{+\infty} e^{-\tau y^2/2} dy = \frac{\theta}{\sqrt{2\pi}} \cdot \int_{(\hat{s}(\tau)-x_2)\sqrt{\tau}}^{+\infty} e^{-z^2/2} dz \\ &= \theta\left(1 - \Phi_1\left([\hat{s}(\tau) - x_2]\sqrt{\tau}\right)\right) = \theta\Phi_1\left([x_2 - \hat{s}(\tau)]\sqrt{\tau}\right) \\ &= \theta\Phi_1\left[\left(x_2 - \frac{x_2 + x_1}{2} + \frac{\ln\hat{\rho}}{(x_2 - x_1)\tau}\right)\sqrt{\tau}\right] \end{aligned}$$

and, likewise,

$$\begin{aligned} p_2(\tau) &= (1-\theta)\Phi_1\left([x_1 - \hat{s}(\tau)]\sqrt{\tau}\right) \\ &= (1-\theta)\Phi_1\left[\left(x_1 - \frac{x_2 + x_1}{2} + \frac{\ln\hat{\rho}}{(x_2 - x_1)\tau}\right)\sqrt{\tau}\right]. \end{aligned}$$

7.1. A family of hyperbolic-exponential functions. Consider functions $C : \mathbb{R}_+ \to \mathbb{R}_+$ of the following form

$$C(\tau) = \int_0^\tau \gamma t^{-d} e^{\alpha t - \beta/t} dt \tag{89}$$

for parameters $\alpha, \beta, \gamma > 0$, and $d \in \mathbb{R}$. We will here henceforth set $d = 1/2$.[21] Clearly these functions are differentiable at positive τ, with derivative

$$C'(\tau) = \frac{\gamma}{\sqrt{\tau}} \cdot e^{\alpha\tau - \beta/\tau}. \tag{90}$$

We note that $C(0) = 0$, $\lim_{\tau\to 0} C'(\tau) = \lim_{x\to+\infty} \gamma x \cdot \exp(-\beta x^2) = 0$, and $\lim_{\tau\to+\infty} C'(\tau) = \lim_{x\to+\infty} (\gamma/x) \cdot \exp(\alpha x^2) = +\infty$. Moreover,

Lemma 2. *For any $\alpha, \beta, \gamma > 0$, and $d = 1/2$, the cost function C in (89) is convex if and only if $16\alpha\beta \geq 1$. Moreover, if $16\alpha\beta < 1$, then it is convex on the intervals $\left[0, \left(1 - \sqrt{1 - 16\alpha\beta}\right) / (4\alpha)\right]$ and $\left[\left(1 + \sqrt{1 - 16\alpha\beta}\right) / (4\alpha), +\infty\right)$.*

Proof:

$$C''(x) = \frac{\gamma}{\sqrt{x}} \left(\alpha + \frac{\beta}{x^2}\right) e^{\alpha x - \beta/x} - \frac{1}{2} \frac{\gamma}{x\sqrt{x}} e^{\alpha x - \beta/x}$$

Hence, $C''(x) \geq 0$ if

$$\alpha + \frac{\beta}{x^2} \geq \frac{1}{2x}$$

or $\alpha x^2 + \beta \geq x/2$. This inequality evidently holds for all x near zero, since $\beta > 0$. Moreover, it holds for all $x \in \mathbb{R}$ if $16\alpha\beta > 1$. To see this, note that the equation

$$\alpha x^2 + \beta = \frac{x}{2}$$

then has no real root, while if $16\alpha\beta = 1$ it has exactly one root, namely, $x = (4\alpha)^{-1}$, and if $16\alpha\beta < 1$ it has two roots,

$$x = \frac{1 \pm \sqrt{1 - 16\alpha\beta}}{4\alpha}$$

If $16\alpha\beta = 1$, then $C''(x) \geq 0$ for all x, with equality iff $x = (4\alpha)^{-1}$, while if $16\alpha\beta < 1$, then $C''(x) \geq 0$ for all

$$x \leq \frac{1 - \sqrt{1 - 16\alpha\beta}}{4\alpha}$$

[21] The more general cases are also of interest, but will not be discussed, since they are not needed in the present analysis.

and all

$$x \geq \frac{1+\sqrt{1-16\alpha\beta}}{4\alpha}.$$

7.2. Proof of Proposition 3:. For the investor in autarky we have

$$\begin{aligned} V(\tau) &= \theta I x_2 \cdot \Phi\left(\frac{\ln \hat{\rho}}{(x_2-x_1)\sqrt{\tau}} + \frac{x_2-x_1}{2}\sqrt{\tau}\right) + \\ &\quad + (1-\theta) I x_1 \cdot \Phi\left(\frac{\ln \hat{\rho}}{(x_2-x_1)\sqrt{\tau}} - \frac{x_2-x_1}{2}\sqrt{\tau}\right) \end{aligned} \tag{91}$$

Clearly V is differentiable in τ for any $\tau > 0$, and

$$\begin{aligned} V'(\tau) &= \theta I x_2 \cdot \frac{\partial}{\partial \tau}\Phi\left(\frac{\ln \hat{\rho}}{(x_2-x_1)\sqrt{\tau}} + \frac{x_2-x_1}{2}\sqrt{\tau}\right) \\ &\quad + (1-\theta) I x_1 \cdot \frac{\partial}{\partial \tau}\Phi\left(\frac{\ln \hat{\rho}}{(x_2-x_1)\sqrt{\tau}} - \frac{x_2-x_1}{2}\sqrt{\tau}\right) \end{aligned} \tag{92}$$

where

$$\frac{\partial}{\partial \tau}\Phi\left(\frac{\ln \hat{\rho}}{(x_2-x_1)\sqrt{\tau}} + \frac{x_2-x_1}{2}\sqrt{\tau}\right) =$$

$$\begin{aligned} &= \frac{\partial}{\partial \tau}\left[\frac{\ln \hat{\rho}}{(x_2-x_1)\sqrt{\tau}} + \frac{x_2-x_1}{2}\sqrt{\tau}\right] \cdot \phi\left(\frac{\ln \hat{\rho}}{(x_2-x_1)\sqrt{\tau}} + \frac{x_2-x_1}{2}\sqrt{\tau}\right) \\ &= \left[\frac{x_2-x_1}{4\sqrt{\tau}} - \frac{\ln \hat{\rho}}{2(x_2-x_1)\tau\sqrt{\tau}}\right] \cdot \phi\left(\frac{\ln \hat{\rho}}{(x_2-x_1)\sqrt{\tau}} + \frac{x_2-x_1}{2}\sqrt{\tau}\right), \end{aligned} \tag{93}$$

and

$$\frac{\partial}{\partial \tau}\Phi\left(\frac{\ln \hat{\rho}}{(x_2-x_1)\sqrt{\tau}} - \frac{x_2-x_1}{2}\sqrt{\tau}\right) =$$

$$\begin{aligned} &= \frac{\partial}{\partial \tau}\left[\frac{\ln \hat{\rho}}{(x_2-x_1)\sqrt{\tau}} - \frac{x_2-x_1}{2}\sqrt{\tau}\right] \cdot \phi_1\left(\frac{\ln \hat{\rho}}{(x_2-x_1)\sqrt{\tau}} - \frac{x_2-x_1}{2}\sqrt{\tau}\right) \\ &= -\left[\frac{x_2-x_1}{4\sqrt{\tau}} + \frac{\ln \hat{\rho}}{2(x_2-x_1)\tau\sqrt{\tau}}\right] \cdot \phi_1\left(\frac{\ln \hat{\rho}}{(x_2-x_1)\sqrt{\tau}} - \frac{x_2-x_1}{2}\sqrt{\tau}\right) \end{aligned} \tag{94}$$

Hence,

$$V'(\tau)/I =$$

$$= \theta x_2 \cdot \left[\frac{x_2 - x_1}{4\sqrt{\tau}} - \frac{\ln\hat{\rho}}{2(x_2 - x_1)\tau\sqrt{\tau}}\right] \cdot \phi\left(\frac{\ln\hat{\rho}}{(x_2 - x_1)\sqrt{\tau}} + \frac{x_2 - x_1}{2}\sqrt{\tau}\right) \tag{95}$$
$$-(1-\theta)x_1 \cdot \left[\frac{x_2 - x_1}{4\sqrt{\tau}} + \frac{\ln\hat{\rho}}{2(x_2 - x_1)\tau\sqrt{\tau}}\right] \cdot \phi_1\left(\frac{\ln\hat{\rho}}{(x_2 - x_1)\sqrt{\tau}} - \frac{x_2 - x_1}{2}\sqrt{\tau}\right),$$

or, since $\exp\left[-(a+b)^2/2\right] = \exp\left[-(a^2+b^2)/2\right]\cdot\exp(-ab)$ and $\exp\left[-(a-b)^2/2\right] = \exp\left[-(a^2+b^2)/2\right]\cdot\exp(ab)$:

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

$$= \hat{\rho}\cdot\left[\frac{x_2 - x_1}{4\sqrt{\tau}} - \frac{\ln\hat{\rho}}{2(x_2 - x_1)\tau\sqrt{\tau}}\right]\cdot\exp\left(-\frac{\ln\hat{\rho}}{2}\right) \tag{96}$$
$$+\left[\frac{x_2 - x_1}{4\sqrt{\tau}} + \frac{\ln\hat{\rho}}{2(x_2 - x_1)\tau\sqrt{\tau}}\right]\cdot\exp\left(\frac{\ln\hat{\rho}}{2}\right)$$
$$= \hat{\rho}\cdot\left[\frac{x_2 - x_1}{4\sqrt{\tau}} - \frac{\ln\hat{\rho}}{2(x_2 - x_1)\tau\sqrt{\tau}}\right]\cdot\frac{1}{\sqrt{\hat{\rho}}}$$
$$+\left[\frac{x_2 - x_1}{4\sqrt{\tau}} + \frac{\ln\hat{\rho}}{2(x_2 - x_1)\tau\sqrt{\tau}}\right]\cdot\sqrt{\hat{\rho}}$$
$$= \frac{x_2 - x_1}{2}\cdot\sqrt{\frac{\hat{\rho}}{\tau}}.$$

or

$$V'(\tau) = \frac{x_2 - x_1}{2}\cdot I(1-\theta)|x_1|\cdot\sqrt{\frac{\hat{\rho}}{2\pi\tau}}\cdot\exp\left[-\frac{1}{2}\left(\frac{\ln\hat{\rho}}{(x_2-x_1)\sqrt{\tau}}\right)^2 - \frac{1}{2}\left(\frac{x_2-x_1}{2}\sqrt{\tau}\right)^2\right] \tag{97}$$

This establishes the necessary FOC, which can be written in the form

$$\frac{I(x_2 - x_1)\hat{\kappa}}{2\gamma\sqrt{2\pi}} = \exp\left(\left[\frac{1}{2}\left(\frac{\ln\hat{\rho}}{x_2 - x_1}\right)^2 - \beta\right]\cdot\frac{1}{\tau} + \left[\frac{1}{2}\left(\frac{x_2 - x_1}{2}\right)^2 + \alpha\right]\cdot\tau\right) \tag{98}$$

or, taking the logarithm of both sides,

$$\left[\frac{1}{2}\left(\frac{x_2 - x_1}{2}\right)^2 + \alpha\right]\cdot\tau - \ln\left[\frac{I(x_2 - x_1)\hat{\kappa}}{2\gamma\sqrt{2\pi}}\right] + \left[\frac{1}{2}\left(\frac{\ln\hat{\rho}}{x_2 - x_1}\right)^2 - \beta\right]\cdot\frac{1}{\tau} = 0 \tag{99}$$

Multiplication of both sides by τ gives

$$A\tau^2 - K\tau + B = 0 \tag{100}$$

where

$$A = \frac{1}{2}\left(\frac{x_2 - x_1}{2}\right)^2 + \alpha > 0 \quad \text{and} \quad B = \frac{1}{2}\left(\frac{\ln \hat{\rho}}{x_2 - x_1}\right)^2 - \beta, \tag{101}$$

and $K = \hat{K}$ as defined in the statement of the proposition. Equation (100) has one root if $K^2 = 4AB$, no root if $K^2 < 4AB$, and two roots if $K^2 > 4AB$. In the last case they are

$$\tau_1 = \frac{1}{2A}\left(K - \sqrt{K^2 - 4AB}\right) \qquad \text{and} \qquad \tau_2 = \frac{1}{2A}\left(K + \sqrt{K^2 - 4AB}\right). \tag{102}$$

The first root is a local minimum and the second root a local maximum point. Writing out τ_2 explicitly yields (46).

To verify global optimality, let $\Pi(\tau) = V(\tau) - C(\tau)$. For any $\tau > 0$:

$$\begin{aligned} \Pi'(\tau) &= \frac{I(x_2 - x_1)\hat{\kappa}}{2\sqrt{2\pi\tau}} \cdot \exp\left[-\frac{1}{2\tau}\left(\frac{\ln \hat{\rho}}{x_2 - x_1}\right)^2 - \frac{\tau}{2}\left(\frac{x_2 - x_1}{2}\right)^2\right] \\ &\quad - \frac{\gamma}{\sqrt{\tau}} \cdot \exp\left(\alpha\tau - \frac{\beta}{\tau}\right) \end{aligned}$$

From the above we have that if $D < 0$ then $\Pi'(\tau) < 0$ for all $\tau > 0$. Hence $\hat{\tau} = 0$. If $D = 0$, then $\Pi'(\tau) \leq 0$ for all $\tau > 0$ with equality when $\tau = K/2A$. Again $\hat{\tau} = 0$. If If $D > 0$, then $\Pi'(\tau) < 0$ for all $\tau < \tau_1$ and for all $\tau > \tau_2$ (with equality at each of these two points), while $\Pi'(\tau) > 0$ for all $\tau \in (\tau_1, \tau_2)$. Hence, if $\tau_2 \leq 0$, then $\hat{\tau} = 0$, while if $\tau_2 > 0$, then $\hat{\tau} \in \{0, \tau_2\}$, depending on which of $\Pi(0)$ and $\Pi(\tau_2)$ is biggest, where $\Pi(0) = I \cdot \max\{\mathbb{E}[X], 0\}$, and

$$\begin{aligned} \Pi(\tau_2) &= \theta I x_2 \cdot \Phi\left(\frac{\ln \hat{\rho}}{(x_2 - x_1)\sqrt{\tau_2}} + \frac{x_2 - x_1}{2}\sqrt{\tau_2}\right) + \\ &\quad + (1 - \theta) I x_1 \cdot \Phi\left(\frac{\ln \hat{\rho}}{(x_2 - x_1)\sqrt{\tau_2}} - \frac{x_2 - x_1}{2}\sqrt{\tau_2}\right) \\ &\quad - \int_0^{\tau_2} \frac{\gamma}{\sqrt{t}} \exp\left(\alpha t - \frac{\beta}{t}\right) dt. \end{aligned}$$

It is evident from the above (and the continuity of Π at zero) that a sufficient condition for $\hat{\tau} = \tau_2$ is that $\tau_1 \leq 0 < \tau_2$. We note that $\tau_1 \leq 0$ if and only if $K \leq 0$.

7.3. The agent's optimal signal threshold. We here derive expressions for the agent's optimal signal threshold and investment probabilities, by similar arguments as those given in the proof of Lemma 1. Again Φ denotes the CDF of the standard normal distribution.

Lemma 3. *The agent's optimal signal threshold under any strictly monotonic contract* $\boldsymbol{w} \in W$ *is*

$$s^*(\boldsymbol{w},\tau) \;=\; \frac{x_2+x_1}{2} - \frac{\ln\rho(\boldsymbol{w})}{(x_2-x_1)\,\tau}\;. \tag{103}$$

When using this investment strategy, the probability that the state of nature is good and that he will investment is

$$p_G(\boldsymbol{w},\tau) \;=\; \theta\cdot\Phi\left(\frac{\ln\rho(\boldsymbol{w})}{(x_2-x_1)\sqrt{\tau}} + \frac{x_2-x_1}{2}\sqrt{\tau}\right), \tag{104}$$

and the probability that the state will be bad and that he will invest is

$$p_B(\boldsymbol{w},\tau) \;=\; (1-\theta)\cdot\Phi\left(\frac{\ln\rho(\boldsymbol{w})}{(x_2-x_1)\sqrt{\tau}} - \frac{x_2-x_1}{2}\sqrt{\tau}\right). \tag{105}$$

Proof: Investment is optimal for the agent if his signal $s\in\mathbb{R}$ satisfies

$$\theta e^{-\tau(s-x_2)^2/2}u(w_2) + (1-\theta)\,e^{-\tau(s-x_1)^2/2}u(w_1) \geq$$
$$\geq \left[\theta e^{-\tau(s-x_2)^2/2} + (1-\theta)\,e^{-\tau(s-x_1)^2/2}\right]u(y)$$

or

$$\theta e^{-\tau(s-x_2)^2/2}\left[u(w_2)-u(y)\right] \geq (1-\theta)\,e^{-\tau(s-x_1)^2/2}\left[u(y)-u(w_1)\right]$$

or

$$\rho(\boldsymbol{w}) \geq e^{-\tau(s-x_1)^2/2+\tau(s-x_2)^2/2} = e^{\frac{1}{2}\tau(x_2-x_1)(x_1-2s+x_2)}$$

or

$$\ln\rho(\boldsymbol{w}) \geq \frac{1}{2}\tau(x_2-x_1)(x_2+x_1) - \tau(x_2-x_1)\,s$$

or

$$\tau(x_2-x_1)\,s \geq -\ln\rho(\boldsymbol{w}) + \frac{1}{2}\tau(x_2-x_1)(x_1+x_2)$$

or

$$s \geq \frac{x_2+x_1}{2} - \frac{\ln\rho(\boldsymbol{w})}{(x_2-x_1)\,\tau}.$$

Using this result one obtains, with ϕ_τ and Φ_τ denoting the PDF and CDF of the normal distribution with mean value zero and variance $\sigma^2 = 1/\tau$:

$$\begin{aligned} p_G(\boldsymbol{w},\tau) &= \theta \int_{s^*(\boldsymbol{w},\tau)}^{+\infty} \phi_\tau(s-x_2)\,ds = \theta \int_{s^*(\boldsymbol{w},\tau)-x_2}^{+\infty} \phi_\tau(x)\,dx \\ &= \theta\sqrt{\frac{\tau}{2\pi}} \cdot \int_{s^*(\boldsymbol{w},\tau)-x_2}^{+\infty} e^{-\tau x^2/2}dx = \frac{\theta}{\sqrt{2\pi}} \cdot \int_{(s^*(\boldsymbol{w},\tau)-x_2)\sqrt{\tau}}^{+\infty} e^{-z^2/2}dz \\ &= \theta\left(1-\Phi_1\left[\left(s^*(\boldsymbol{w},\tau)-x_2\right)\sqrt{\tau}\right]\right) = \theta\Phi_1\left[\left(x_2-s^*(\boldsymbol{w},\tau)\right)\sqrt{\tau}\right] \\ &= \theta\Phi_1\left[\left(x_2-\frac{x_2+x_1}{2}+\frac{\ln\rho(\boldsymbol{w})}{(x_2-x_1)\tau}\right)\sqrt{\tau}\right] \\ &= \theta\Phi_1\left[\left(\frac{x_2-x_1}{2}+\frac{\ln\rho(\boldsymbol{w})}{(x_2-x_1)\tau}\right)\sqrt{\tau}\right] \end{aligned}$$

and

$$\begin{aligned} p_B(\boldsymbol{w},\tau) &= (1-\theta)\int_{s^*(\boldsymbol{w},\tau)}^{+\infty} \phi_\tau(s-x_1)\,ds = (1-\theta)\,\Phi_1\left[\left(x_1-s^*(\boldsymbol{w},\tau)\right)\sqrt{\tau}\right] \\ &= (1-\theta)\,\Phi_1\left[\left(x_1-\frac{x_2+x_1}{2}+\frac{\ln\rho(\boldsymbol{w})}{(x_2-x_1)\tau}\right)\sqrt{\tau}\right], \end{aligned}$$

leading to the claimed expressions.

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 (2016): "Rational inattention, optimal consideration sets and stochastic choice", mimeo., New York University, Columbia University, and University of Michigan.

[3] Chade, H., and E. Schlee (2002): "Another look at the Radner-Stiglitz nonconcavity in the value of information", *Journal of Economic Theory* 107, 421-452.

[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., and A. de Palma (2016): "Generalized entropy models", MPRA Paper No. 70249.

[10] 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.

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

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

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

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

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

[16] 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.

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

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

[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 go got there", in G. Constantinides, M.Harris and , R. Stulz (eds), *Handbook of the Economics of Finance*. Amsterdam: North-Holland.

[22] Myerson, R., and J. Weibull (2015): "Tenable strategy blocks and settled equilibria", *Econometrica* 83, 943-976.

[23] Radner R., and J. Stiglitz (1984): "A nonconcavity in the value of information", pp. 33-52, in M. Boyer and R.E. Kihlstrom (eds.), *Bayesian Models in Economic Theory.* North-Holland.

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

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

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

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

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

[29] Snickars, F., and J. Weibull (1977): "A minimum information principle", *Regional Science and Urban Economics* 7, 137–168.

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

[31] 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.

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

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

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