Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 162 (2016) 55–92

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Delegated information acquisition with moral hazard ☆

Hector Chade, Natalia Kovrijnykh *

Department of Economics, Arizona State University, United States

Received 6 January 2015; final version received 18 November 2015; accepted 10 December 2015

Available online 17 December 2015

Abstract

We analyze a principal–agent problem with moral hazard where a principal searches for an opportunity of uncertain return, and hires an agent to evaluate available options. The agent's effort affects the informativeness of a signal about an option's return. Based on the information provided by the agent, the principal decides whether to exercise the option at hand. We derive properties of the optimal contract in both static and dynamic versions of the problem. We show that there are intermediate values of the prior probability that the option is of high quality at which positive effort cannot be sustained. We also show that if the prior is below a threshold, then the agent is rewarded for delivering 'bad news' about the option's quality. We derive distortions (relative to the first best) on the implemented effort level and optimal stopping decision. For some parameter values, it is optimal for the principal to commit to an ex-post suboptimal stochastic decision to exercise an option.

JEL classification: C73; D82; D83; D86

Keywords: Moral hazard; Principal–agent model; Optimal contracts; Information acquisition; Search; Lotteries

☆ We are grateful to Larbi Alaoui, Ying Chen, Peter DeMarzo, Amanda Friedenberg, Marco Ottaviani, Kevin Reffett, Richard Rogerson, Edward Schlee, Jeroen Swinkels, Balázs Szentes, and William Zame for helpful comments and suggestions. We also benefited from comments by seminar participants at Arizona State University, University of British Columbia, CEA-Universidad de Chile, IESE, UCLA, Universitat Pompeu Fabra, UCSD, Toulouse School of Economics, Econometric Society World Congress 2010, 2010 Society for Economic Dynamics meetings, 2010 Southwest Economic Theory Conference, and 2011 University of Hong Kong Summer Microeconomics Seminars. This paper previously circulated under the title "Information Acquisition, Moral Hazard, and Rewarding for Bad News."

* Corresponding author.
E-mail addresses: hector.chade@asu.edu (H. Chade), natalia.kovrijnykh@asu.edu (N. Kovrijnykh).

http://dx.doi.org/10.1016/j.jet.2015.12.003

1. Introduction

Many situations of economic interest feature a client (principal) who searches for an opportunity/option whose return is uncertain, and hires an expert (agent) to evaluate potential options. The agent's effort, which he chooses privately, affects the quality of information he receives about the options. Based on the information the agent provides, the principal compensates him and decides whether to exercise an option or continue the search.

Interactions of this sort abound in real-world applications. For instance, consider an individual seeking to purchase a house, who hires a home inspector to provide her with information about houses that become available. Or take an uninformed investor who hires an expert to evaluate potential investment opportunities. Although other features may be present in these applications, information acquisition certainly plays an important role.

The purpose of this paper is to analyze how to structure incentives in these environments, where the quality of information is endogenous and hidden, and to shed light on the main distortions that this moral-hazard problem impinges on information acquisition and on the principal's decision to exercise an option.

To focus on the main trade-offs involved in this problem, we build a tractable principal–agent model with risk-neutral parties, moral hazard, and limited liability. The agent's role is to acquire information about the unknown quality of an available opportunity, which can be high or low. Information consists of a binary signal whose informativeness depends on the agent's effort. We assume that the signal realization is publicly observable, and thus the agent cannot misreport it. The principal conditions her decision to accept or reject an opportunity on the information acquired by the agent. We also assume that contracts can only be conditioned on the signal and the principal's decision to exercise an option. That is, although the quality of the option may potentially be observable ex post (after the option is exercised), it is not contractible.

The above assumptions are not far-fetched, for instance, in the context of the home-inspector application. Regarding quality being non-contractible, that a house is of bad quality might be revealed several years after the purchase, making it hard to incorporate that event into the contract. For example, a house may have a problem with the roof, which might not be detected during an inspection and will cause leakages, but only during a very long and heavy rain that seldom occurs. As for the signal being publicly observable, many tests—as well as their results—that home inspectors perform on a house are verifiable and are difficult to misreport. Also, even when misreporting is possible, it does not seem to be a particular concern in this business (at least if the inspector is independently hired by the potential home buyer as opposed to being recommended by the real-estate agent). At the same time, overlooking an important problem due to a careless inspection is certainly a big issue, and we focus on it in this paper.

We first derive the optimal contract in the static case, with only one option to evaluate. Then we tackle the dynamic case, where there is an option in every period that is evaluated before the principal decides to stop the search. As a benchmark, and also to highlight the drastic effect that moral hazard and information acquisition have on the optimal contract, we show that without informational problems the solution is trivial. Indeed, in the first-best case, depending on the prior belief about the option quality, the principal either does not hire the agent and always/never buys, or she implements a constant level of effort (signal precision) over time and exercises the first option whose signal realization is high.

Results are radically different under moral hazard. We show that there is a threshold value of the prior belief that an option is of high quality such that the agent is rewarded for a high signal realization if and only if the prior is above this threshold. This cut-off value is higher when the relative effect of effort on the signal distribution in the low- vs. high-quality state is higher. Moreover, at the threshold, it is infinitely costly to implement any positive effort, and hence the principal implements zero effort in a neighborhood of the threshold. Thus, the principal does not use the agent's services, even though information might be very valuable.

When the prior belief is below the threshold, the agent is rewarded for delivering bad news about the option's quality. We call this feature *rewarding for bad news*.[1] We argue that it arises naturally when the agent's effort generates information about a characteristic that is payoff-relevant for the principal.

In the static setting, rewarding for bad news means that the agent gets paid when the principal chooses not to exercise an option. Interestingly, this last feature disappears in the dynamic case, because the principal prefers to postpone payments until she stops the search. Thus, the agent *only* gets paid at the end, that is, when the principal exercises an option. The size of that payment depends on the history observed up to that point. Hence, rewarding for bad (good) news in the dynamic case means that the agent's expected future payments increase (decrease) upon a low signal realization.

The implemented effort and the principal's decision to exercise an option are both distorted under moral hazard compared to the first best. In the static case, effort is distorted downwards, so the decision to exercise an option is based on less precise information than in the first-best case. Additional distortions arise in the dynamic case. For instance, in the first-best case effort is constant over time, while in the second-best case it can fluctuate. The reason is that the agent's expected payments optimally change over time depending on the history, which in turn makes the implemented effort vary over time.

Another distortion in the dynamic case is that, unlike in the first best, the principal's decision to exercise an option might involve randomization. The principal might find it optimal to commit to a lottery over exercising an option and not exercising it, even though ex post she is not indifferent between the two alternatives. Specifically, we show that for some parameter values the principal will not exercise an option with some probability, even though she receives favorable news about it from the agent, and exercising it would be ex-post profitable. This occurs because of the asymmetry embedded in the model, namely, that the search stops if the principal exercises an option and it continues if she does not. Finally, we also provide a fairly complete description of the optimal contract when the principal can only commit to a deterministic purchase decision.

Related Literature. The paper is related to the large literature on repeated moral hazard (e.g., Rogerson, 1985b, Spear and Srivastava, 1987, and Sannikov, 2008). What distinguishes our model is that effort affects the informativeness of a signal about the option quality and not the distribution of output, and that the principal uses information provided by the agent to make a decision. Our paper is also related to Toxvaerd (2006), Mason and Valimaki (2015), and Lewis and Ottaviani (2008). In these papers, effort affects the probability of completing a project, or the arrival of opportunities that lead the agent to stop the search. That is, the principal delegates the search to the agent. In our model, the principal conducts the search, and delegates the acquisition of information to the agent.

[1] In our model, the signal provides information about both the effort chosen by the agent and the quality of the option. As a signal of effort, a low realization can be good or bad news depending on the prior. But as a signal of quality, a low realization is *always* bad news (the option's expected return is lower). It is in this sense that bad news should be interpreted. Moreover, the principal always receives a lower payoff if a low signal is realized (even after paying the agent), which reinforces the bad news interpretation.

There is a vast literature on decisions with information acquisition—see, e.g., Chade and Schlee (2002) and the references cited therein—where, before making a decision, an individual acquires a signal, whose cost is exogenous and is increasing in the signal's informativeness. In our model, cost is endogenous, determined by the moral hazard problem.

Among papers that analyze *private* information acquisition, the following are the most related ones. Szalay (2009) and Krahmer and Strausz (2011) study static procurement problems with adverse selection where the agent privately acquires information. Both papers show that information acquisition leads to distortions in effort and production. Unlike them, we derive the rewarding-for-bad-news property and, more importantly, provide a detailed analysis of the dynamic case. Also, the last section of Inderst and Ottaviani (2012) analyzes a static contracting problem where an agent exerts hidden effort that affects the informativeness of a signal about the suitability of a match between a firm's product and a buyer's valuation for it. Unlike in our model, it is the buyer and not the firm (the principal) who uses the information provided by the agent before purchasing, and the authors focus on different distortions than we do. Finally, Kashyap and Kovrijnykh (forthcoming) analyze a model of credit ratings, where rating precision depends on the hidden effort exerted by the credit rating agency. Their model is similar to one of the extensions that we present in the Online Appendix where the principal can observe an additional signal, but they focus on completely different questions than we do.

Another related literature is that on delegated expertise, such as Gromb and Martimort (2007), Szalay (2005), and Eso and Szentes (2007) (see also the references cited therein). Gromb and Martimort (2007) study (static) incentive contracts for experts who can collude among themselves or with the principal. Their one-agent-one-signal case has similarities with our model, except that effort is binary, and the agent can misreport his signal. Szalay (2005) analyzes a model where the agent has private information about the signal, and the agent has an intrinsic interest in making an accurate decision. He shows that the optimal contract rewards for 'unexpected news,' while in our setting, when the information structure is, e.g., symmetric, the agent is instead compensated for confirming the principal's prior. Eso and Szentes (2007) study a static problem between a client and an expert consultant. The expert provides the client with information that reduces the noise in her estimate of the value of an investment opportunity. The client has private information and the expert designs the contract. In our case, the principal is the 'client' and designs the contract. Also, she does not have private information, and moral hazard is on the expert's side instead.

Finally, a prediction similar in flavor to our rewarding-for-bad-news result appears in Levitt and Snyder (1997). They analyze a model where the agent has a private signal of the likelihood of the project's success, and early access to that information is valuable to the principal because she can terminate the project. They show that the agent is rewarded for coming forward with bad news. Compared to our result, theirs emerges from adverse selection and relies on the benefits from early information revelation.

The outline of the paper is as follows. Section 2 describes the model. Section 3 derives main properties of the optimal contract in the static case. Section 4 focuses on the dynamic case, where we separately consider the cases of commitment to deterministic and random purchase decisions. Section 5 concludes. Proofs are in Appendix A, and the Online Appendix contains some extensions.

2. The model

Our dynamic principal–agent model, which subsumes the static version as a special case, has the following characteristics.

THE ENVIRONMENT. In each period, there is an option. The quality q of an option is unobservable, and it can be low or high, i.e., $q \in \{\ell, h\}$. If exercised, a high-quality option yields a return $y_h > 0$ to the principal, while the return of a low-quality option is $y_\ell < 0$. Once an option is exercised, the relationship ends.[2] The prior belief that the option is of high quality is $0 \leq \gamma \leq 1$ (the same in every period), which can also be interpreted as the fraction of high-quality options in the pool.

The role of the agent in this relationship is that of an expert who provides the principal with information about the quality of the option at hand in each period. More precisely, the agent observes a signal $\theta \in \{\theta_\ell, \theta_h\}$ that is correlated with the quality of that option. How informative the signal is about quality depends on the level of effort e that the agent exerts, where $0 \leq e \leq \bar{e} \leq \infty$. To capture this dependence, we focus on a class of information structures described by the following conditional probabilities:

$$\Pr\{\theta = \theta_h | q = h, e\} = \alpha + \beta_h \eta(e) \quad \text{and} \quad \Pr\{\theta = \theta_h | q = \ell, e\} = \alpha - \beta_\ell \eta(e), \tag{1}$$

where $0 < \alpha < 1$, $\beta_i \geq 0$, $i = \ell, h$, $\beta_\ell + \beta_h > 0$, and the function η satisfies $\eta(0) = 0$, $\eta'(e) > 0$, and $\eta''(e) \leq 0$ for all e. Also, to ensure that the expressions in (1) are between zero and one, we require $\lim_{e \to \bar{e}} \eta(e) \leq \min\{(1-\alpha)/\beta_h, \alpha/\beta_\ell\}$.

Note that if effort is zero, the conditional distribution of the signal is the same regardless of an option's quality, and therefore the signal is uninformative. Conditional on the option being of a certain quality, the probability of observing a signal consistent with that quality is increasing in the agent's effort. That is, higher effort makes the signal more informative.[3]

The assumed class of information structures nests the extreme cases $\beta_h = 0$ or $\beta_\ell = 0$, and the symmetric case with $\beta_h = \beta_\ell$. When $\beta_h = 0$, the agent's effort only affects the distribution of the signal if the option's quality is low, so the agent's effort matters only in detecting bad options. The situation is reversed if instead $\beta_\ell = 0$. And when $\beta_h = \beta_\ell$, the agent's effort increases the likelihood of observing a signal consistent with the unknown option quality by the same amount in both states.

Exerting effort e is a costly activity for the agent, which entails a disutility given by $\psi(e)$. The function ψ satisfies $\psi(0) = 0$, $\psi'(e) > 0$, $\psi''(e) > 0$ for all $e > 0$, and $\lim_{e \to \bar{e}} \psi(e) = +\infty$. We also assume that ψ'/η' satisfies $\psi'(0)/\eta'(0) = 0$, $(\psi'/\eta')'(0) = 0$, and ψ'/η' is convex in e. The first assumption ensures that the first-order condition of the agent's effort problem captures its solution even if it is zero, and together with the second one it implies that the principal induces positive effort if she buys only after observing a high signal realization. The third condition makes the principal's problem strictly concave in e in the static case.[4]

[2] For convenience, we use the terms 'exercising an option,' 'buying,' and 'purchasing' interchangeably.

[3] An increase in effort makes observing the signal a more informative experiment in Blackwell's sense. See Blackwell and Girshick (1954), Athey and Levin (2001), Jewitt (2007), and Quah and Strulovici (2009).

[4] Convexity of ψ'/η' ensures that the principal's marginal cost of implementing effort under moral hazard is increasing in e. Technically, since the principal's problem includes the first-order condition with respect to effort from the agent's problem as an incentive constraint, we impose (sufficient) conditions not only on the second but also on the third derivatives to guarantee that local second-order conditions are satisfied globally. When $\eta(e)$ is linear in e, the convexity of ψ'/η' reduces to the convexity of the marginal disutility of effort ψ', a common assumption in principal–agent problems—see, e.g., Jewitt et al. (2008).

The signal realization is publicly observable, thereby ruling out the possibility that the agent may lie about it. (We discuss the relaxation of this assumption in the case when the principal observes an additional signal in the Online Appendix.)

Both the principal and the agent are risk neutral and each maximizes the expected sum of discounted payoffs using a discount factor $0 \leq \delta < 1$. The principal's per-period payoff is equal to the expected payoff from her buying decision (equal to zero if she decides not to buy) minus the expected wages paid to the agent. In turn, the agent's per-period payoff is equal to the expected wage received minus the disutility of effort.

If the agent does not work for the principal, he can instead enjoy a reservation utility (in expected discounted value terms) equal to zero.[5] To avoid a trivial solution, we assume that the agent is protected by limited liability, and thus wages are restricted to be nonnegative.

CONTRACTS. At the beginning of the first period, the parties sign a long-term contract that covers the duration of the relationship. In each period and for every possible history, it specifies the recommended effort to be exerted by the agent, the principal's decision to exercise an option, and a wage paid to the agent (all possibly random) until the principal stops the search, thereby ending the relationship. Both the signal realization and the decision to exercise the option are contractible events. Finally, we assume full commitment to long-term contracts by both parties. (It will become clear below that only the principal's commitment assumption bites in the model because the agent's reservation utility is zero.)

TIMING. In each period while the game continues, the timing is as follows. First the agent exerts effort, which generates a signal, whose realization is publicly observed. Then the principal buys or does not buy as prescribed by the contract.[6] After that the agent receives the wage in accordance with the realized signal and the purchase decision. Finally, if the principal exercises the option, the game ends; otherwise the search continues in the next period. She can also stop the search without buying by committing to never buy and to pay nothing to the agent in all subsequent periods.

Formally, let θ_t be the signal realization in period t, and let $\mathcal{I}_t$ denote an indicator function that describes whether the principal exercises the option ($\mathcal{I}_t = 1$) or not ($\mathcal{I}_t = 0$) in period t. A non-terminal public history in period t is $h^t = (\theta_0, \mathcal{I}_0, \theta_1, \mathcal{I}_1, \ldots, \theta_t, \mathcal{I}_t)$, such that $\mathcal{I}_\tau = 0$ for all $\tau = 0, 1, \ldots, t$.[7] Let H^t be the set of all non-terminal histories in period t. Then, a long-term contract is a triple of sequences of recommended effort, probability of buying and wage functions $\{e_t(h^{t-1}), \sigma_t(h^{t-1}, \theta_t), w_t(h^{t-1}, \theta_t, \mathcal{I}_t)\}_t$ for all $t \geq 0$, all $h^{t-1} \in H^{t-1}$, and all possible realizations of signal θ_t and purchase decision $\mathcal{I}_t$. Finally, a contract that also allows for random wages and recommended effort levels is a probability distribution over these long-term contracts. As usual, a contract is feasible if it satisfies the agent's participation constraint and the agent follows the recommended effort levels at each non-terminal history (i.e., the recommended effort satisfies his incentive constraints). A contract is optimal if it is feasible and maximizes the principal's expected discounted profits.

5 We make this assumption to simplify the presentation of the results. With some minor modifications, the analysis can accommodate any nonnegative reservation utility.

6 If the contract entails a stochastic purchase decision, then the outcome of the lottery is first observed. Similarly, if the contract involves randomization in payments (and recommended effort), then the outcome of a lottery is observed before the agent makes his effort choice.

7 Strictly speaking, the definition of a non-terminal history should also include the indicator function of the principal's decision whether or not to stop the search without exercising an option. But since stopping the search can be thought of as not paying to the agent and not buying in all subsequent periods, we exclude this additional decision for notational simplicity.

Notice also that we focus on contracts that condition payments on the signal realization and purchase decision, but not on the ex-post realization of the state. In addition, we assume that the signal realization is public. These assumptions, which are important features of our model, can be defended on several grounds. First, as we mentioned in the introduction, they are plausible for the described applications (e.g., home inspectors are not paid conditioned on whether a house turns out ex post to be of low quality, and the results of many of their tests are easy to observe). Second, they lead to novel predictions about the structure of incentives and distortions compared to other studies on delegated expertise (e.g., compared to the models of Szalay, 2005, where the signal is private, and with Gromb and Martimort, 2007, who condition on the ex-post realization of the state). For further comparison, in the Online Appendix we analyze a variation of the model that allows for conditioning contracts on an additional signal (and also potentially allowing the agent's signal to be his private information), which collapses to the ex-post realization of the state if the signal is perfectly informative.

3. The static case

We first analyze the one-period version of the model. To simplify the notation, we define $\pi(e) \equiv \gamma(\alpha + \beta_h \eta(e)) + (1-\gamma)(\alpha - \beta_\ell \eta(e))$ as the unconditional probability that a signal is high when the agent's effort level is e. Its derivative, which plays an important role below, is thus given by $\pi'(e) = \eta'(e)(\gamma\beta_h - (1-\gamma)\beta_\ell)$. Furthermore, if the principal exercises the option after observing a certain signal realization, her expected returns conditional on a given level of effort and on observing a high a low signal realization are given by

$$E_h(e) \equiv E[y|\theta_h, e] = \frac{\gamma(\alpha + \beta_h \eta(e))y_h + (1-\gamma)(\alpha - \beta_\ell \eta(e))y_\ell}{\pi(e)},$$

$$E_\ell(e) \equiv E[y|\theta_\ell, e] = \frac{\gamma(1-\alpha - \beta_h \eta(e))y_h + (1-\gamma)(1-\alpha + \beta_\ell \eta(e))y_\ell}{1-\pi(e)}.$$

Also, we define $u_h(e) \equiv \pi(e)E_h(e)$ and $u_\ell(e) \equiv (1-\pi(e))E_\ell(e)$.

3.1. Observable effort

Let us consider the first-best case in which the agent's effort is observable and there is no limited liability. Since the solution for this case is equivalent to the one in which the principal acquires information herself, we analyze the latter problem. Let σ_i denote the probability that the principal exercises the option after signal θ_i. The principal solves

$$\max_{e\geq 0, \sigma_h \in [0,1], \sigma_\ell \in [0,1]} -\psi(e) + \pi(e)\sigma_h E_h(e) + (1-\pi(e))\sigma_\ell E_\ell(e).$$

It immediately follows that it is optimal for the principal to buy after a particular signal realization if the expected quality of the option given the signal is positive, and not to buy otherwise. Also, since $E_h(e) \geq E_\ell(e)$, with strict inequality if and only if $e > 0$, it follows that $\sigma_h \geq \sigma_\ell$. It also follows that it only pays off to exert effort if the principal intends to exercise the option after observing a high signal and not to exercise it after observing a low signal. Indeed, if the principal's decision is independent of the signal realization, then information is wasted, and hence putting effort into acquiring it is not optimal. Thus the problem reduces to three relevant alternatives available to the principal, namely, exert no effort and do not buy (which yields a payoff of zero), exert no effort and buy (with the expected payoff $\gamma y_h + (1-\gamma)y_\ell$), and exert an optimal level of effort and buy only if the signal realization is high (yielding $\max_e -\psi(e) + \pi(e)E_h(e)$).

The optimal level of effort under the third alternative uniquely solves $\psi'(e) = \eta'(e)(\gamma\beta_h y_h - (1-\gamma)\beta_\ell y_\ell)$. We will denote by e^* the principal's optimal effort choice in the first-best case.

Intuition suggests that the principal's optimal decision is to never buy if γ is small enough and to always buy if γ is high enough. For intermediate values of γ, she acquires information and buys only if the signal is high. The following result formalizes this intuition.

Lemma 1 *(Static first best). There exist thresholds* $\underline{\gamma}^*$ *and* $\bar{\gamma}^*$, *satisfying* $0 < \underline{\gamma}^* < -y_\ell/(y_h - y_\ell) < \bar{\gamma}^* < 1$, *such that the principal:*

(*i*) *Exerts no effort, i.e.,* $e^* = 0$, *and does not exercise the option if* $\gamma \in [0, \underline{\gamma}^*]$*;*

(*ii*) *Exerts no effort, i.e.,* $e^* = 0$, *and exercises the option if* $\gamma \in [\bar{\gamma}^*, 1]$*;*

(*iii*) *Exerts a positive level of effort, i.e.,* $e^* > 0$, *and exercises the option only when the signal realization is high if* $\gamma \in (\underline{\gamma}^*, \bar{\gamma}^*)$.

3.2. Unobservable effort

We now turn to the case with unobservable effort—i.e., moral hazard—and limited liability. The principal designs a compensation scheme $(w_{h1}, w_{h0}, w_{\ell 1}, w_{\ell 0})$, where w_{i1} is the payment from the principal to the agent if the signal is θ_i, $i \in \{h, \ell\}$, and the principal exercises the option, and w_{i0} is the corresponding payment if the principal does not exercise the option after signal θ_i. The optimal contracting problem is

$$\begin{aligned}
\max_{e,\{w_{i1},w_{i0},\sigma_i\}_{i\in\{h,\ell\}}} \quad & \pi(e)[\sigma_h E_h(e) - \sigma_h w_{h1} - (1-\sigma_h)w_{h0}] \\
& +(1-\pi(e))[\sigma_\ell E_\ell(e) - \sigma_\ell w_{\ell 1} - (1-\sigma_\ell)w_{\ell 0}] \\
\text{s.t.} \quad & -\psi(e) + \pi(e)[\sigma_h w_{h1} + (1-\sigma_h)w_{h0}] + (1-\pi(e))[\sigma_\ell w_{\ell 1} + (1-\sigma_\ell)w_{\ell 0}] \geq 0, \\
& \psi'(e) = \pi'(e)[\sigma_h w_{h1} + (1-\sigma_h)w_{h0} - \sigma_\ell w_{\ell 1} - (1-\sigma_\ell)w_{\ell 0}], \\
& e \geq 0,\ w_{i1} \geq 0,\ w_{i0} \geq 0,\ \sigma_i \in [0,1] \text{ for } i \in \{h,\ell\}.
\end{aligned}$$

The first constraint is the agent's participation constraint. The second one summarizes the incentive constraints using the first-order condition of the agent's effort choice problem. The last constraints are due to the nonnegativity of effort, the presence of limited liability, and the restriction that the probability of purchase lies between zero and one.

Notice that by defining $w_i \equiv \sigma_i w_{i1} + (1-\sigma_i)w_{i0}$, i.e., the expected payment after signal θ_i, $i \in \{h, \ell\}$, one can equivalently rewrite the above problem as

$$\max_{e,\{w_i,\sigma_i\}_{i\in\{h,\ell\}}} \quad \pi(e)[\sigma_h E_h(e) - w_h] + (1-\pi(e))[\sigma_\ell E_\ell(e) - w_\ell]$$

$$\text{s.t.} \quad -\psi(e) + \pi(e)w_h + (1-\pi(e))w_\ell \geq 0, \tag{2}$$

$$\psi'(e) = \pi'(e)[w_h - w_\ell], \tag{3}$$

$$e \geq 0,\ w_i \geq 0,\ \sigma_i \in [0,1] \text{ for } i \in \{h,\ell\}. \tag{4}$$

Just as in the first-best case, it is optimal to set $\sigma_i = 1$ whenever $E_i(e) > 0$ and $\sigma_i = 0$ otherwise. That is, the principal never benefits from randomizing over the purchase decision after a particular signal. Interestingly, we will show later that this result no longer holds in the dynamic version of the model.

Hence, we have again reduced the problem to one where there are only three relevant cases: (*i*) the principal does not hire the agent and does not exercise the option (i.e., $e = 0$ and

$\sigma_h = \sigma_\ell = 0$), (*ii*) she does not hire the agent and exercises the option ($e = 0$ and $\sigma_h = \sigma_\ell = 1$), and (*iii*) she implements positive effort and exercises the option only after the high signal ($e > 0$, $\sigma_h = 1$, and $\sigma_\ell = 0$). In terms of the original contract, in the first case, $w_{h0} = w_{\ell 0} = 0$, while the choices of w_{h1} and $w_{\ell 1}$ are irrelevant, so without loss of generality these payments can be set to zero. Similarly, in the second case $w_{h1} = w_{\ell 1} = 0$, while the choices of w_{h0} and $w_{\ell 0}$ are irrelevant. In the last case $w_h = w_{h1}$, $w_\ell = w_{\ell 0}$, and the values of w_{h0} and $w_{\ell 1}$ are irrelevant.

Let us take a closer look at the first-order condition of the agent's problem, (3), which, using the definition of $\pi(e)$, can be rewritten as

$$\psi'(e) = \eta'(e)(\gamma\beta_h - (1-\gamma)\beta_\ell)[w_h - w_\ell]. \tag{5}$$

Let $\hat{\gamma} \equiv \beta_\ell/(\beta_\ell + \beta_h)$ and notice that $\hat{\gamma} \in [0, 1]$, taking the extreme values 0 and 1 when $\beta_\ell = 0$ and $\beta_h = 0$, respectively (i.e., when the agent's effort affects the signal distribution in one of the two states only), and the value $1/2$ when $\beta_h = \beta_\ell$ (the symmetric information structure). Also, $\pi'(e) \geq 0$ if and only if $\gamma \geq \hat{\gamma}$. Thus, the information structure neatly partitions the solution of the principal's problem into *two distinct cases*: (*i*) $\gamma > \hat{\gamma}$, in which case $w_h > w_\ell$ in order to induce any positive level of effort from the agent, and (*ii*) $\gamma < \hat{\gamma}$, in which case $w_\ell > w_h$. In both instances the agent's problem is strictly concave in effort and thus the first-order approach (replacing the agent's optimization problem by its first-order condition) is valid (Rogerson, 1985a).

Notice that a result that follows from (5) is that as γ approaches $\hat{\gamma}$, it becomes *infinitely costly* to induce the agent to choose any positive level of effort. Hence, no effort is implemented for values of γ close to $\hat{\gamma}$. Intuitively, the reason is that at this threshold, effort has no impact on the unconditional distribution of the signal. Thus, it is impossible to statistically distinguish whether or not the agent has exerted effort.

The existence of a value γ such that *no* positive effort level is implementable depends on the class of information structures that we have assumed. To see this, let $\alpha + \eta_h(e)$ be the probability of the high signal in the high state and $\alpha - \eta_\ell(e)$ the corresponding one in the low state. Then $\pi(e) = \gamma[\alpha + \eta_h(e)] + (1-\gamma)[\alpha - \eta_\ell(e)]$ and thus $\pi'(e) = \gamma\eta_h'(e) - (1-\gamma)\eta_\ell'(e)$. It is immediate that the existence of a value for the prior γ such that $\pi'(e) = 0$ for *all* levels of effort requires that $\eta_h'(e)/\eta_\ell'(e)$ be a constant, which reduces to the class of information structures we consider.[8] Perhaps more interestingly, notice that even if the ratio is not a constant, for each effort level there is a level of the prior γ that makes $\pi'(e) = 0$. In other words, for any level of the prior γ we can pin down which effort levels are *not* implementable.

Proposition 1 summarizes main properties of the optimal contract. Besides $\hat{\gamma}$, a threshold that plays an important role below is $\tilde{\gamma}$, defined by $\tilde{\gamma} y_h + (1-\tilde{\gamma})y_\ell = 0$. That is, $\tilde{\gamma}$ is the value of the prior belief at which the principal is just indifferent between buying and not buying without information, and thus information is most valuable at this point.

[8] This argument extends beyond the binary-signal case. Let $F(\cdot|h, e)$ and $F(\cdot|\ell, e)$ be the conditional c.d.f. of the signal θ under the high and low states, respectively. Then $F(\theta|e) = \gamma F(\theta|h, e) + (1-\gamma)F(\theta|\ell, e)$ and thus $F_e(\theta|e) = \gamma F_e(\theta|h, e) + (1-\gamma)F_e(\theta|\ell, e)$. Necessary conditions for $F_e(\theta|e)$ to be equal to zero are then (*i*) $F_e(\theta|h, e)F_e(\theta|\ell, e) \leq 0$ for all θ and e (which is easy to justify from the Blackwell's informativeness order as a function of e), and (*ii*) whenever (*i*) holds with strict inequality on a set of θ of positive measure, then $F_e(\theta|h, e)/F_e(\theta|\ell, e) = c$, where c is a negative constant.

Proposition 1 *(Optimal static contract).*

(i) If $\gamma > \hat{\gamma}$, then $w_h \geq w_\ell = 0$, and if $\gamma < \hat{\gamma}$, then $w_\ell \geq w_h = 0$. In both cases, the inequality is strict if the implemented level of effort is positive.
(ii) The optimal effort level e is lower than the first-best level e^*, and strictly so if $e^* > 0$.
(iii) The optimal effort and buying decisions satisfy the following properties:
(a) There is a threshold $\underline{\gamma} > \underline{\gamma}^*$ such that, for all values of the prior $\gamma \in [0, \underline{\gamma}]$, the principal does not hire the agent and never buys, i.e., $e = 0$ and $\sigma_h = \sigma_\ell = 0$.
(b) There is a threshold $\bar{\gamma} < \bar{\gamma}^*$ such that, for all values of the prior $\gamma \in [\bar{\gamma}, 1]$, the principal does not hire the agent and always buys, i.e., $e = 0$ and $\sigma_h = \sigma_\ell = 1$.
(c) Unless $\hat{\gamma} = \tilde{\gamma}$, there exists an interval around $\hat{\gamma}$ such that for any value of the prior γ in this interval the principal does not hire the agent ($e = 0$), and either always buys ($\sigma_h = \sigma_\ell = 1$) if $\gamma > \tilde{\gamma}$, or never buys ($\sigma_h = \sigma_\ell = 0$) if $\gamma \leq \tilde{\gamma}$.
(d) If $\hat{\gamma} \neq \tilde{\gamma}$, then there exists an interval around $\tilde{\gamma}$, such that for values of the prior in this interval the optimal level of effort is positive at the optimum.

As in any moral-hazard problem, compensation is determined by the relative likelihood of observing a high vs. a low signal under the implemented effort. The appropriate likelihood ratios involve the unconditional probabilities of the signal realizations. Hence

$$w_h \geq w_\ell \Leftrightarrow \pi'(e)/\pi(e) \geq -\pi'(e)/(1-\pi(e)) \Leftrightarrow \pi'(e) \geq 0 \Leftrightarrow \gamma \geq \hat{\gamma} \Leftrightarrow \gamma\beta_h \geq (1-\gamma)\beta_\ell$$

Notice that when $\gamma < \hat{\gamma}$, the optimal compensation scheme is such that the agent gets paid *only* under the *low* signal realization. We call this feature of the optimal compensation scheme *rewarding for bad news*. As the statement above reveals, $\gamma < \hat{\gamma} \Leftrightarrow \gamma\beta_h < (1-\gamma)\beta_\ell$. Thus, the agent is rewarded for bad news if and only if the impact of his effort on the distribution of the signal is sufficiently stronger in the low vs. high state. The intuition underlying this property is easily explained in the extreme case where $\beta_h = 0$, and thus $\pi(e) = \alpha - (1-\gamma)\beta_\ell\eta(e)$. Since the agent's effort decreases the unconditional probability of a high signal, it makes observing a low signal relatively more likely when the agent chooses the effort level that the principal wants to implement, than when he chooses a lower level of effort. Hence, his wage should be higher when the principal does not buy.

If $\gamma > \hat{\gamma}$, then the agent is rewarded when the principal exercises an option; i.e., the agent is rewarded for good news. This is a typical result in standard principal–agent models with moral hazard, where effort stochastically increases the distribution of 'output' instead of affecting the informativeness of the signal as in our case. Within our framework, this is akin to assuming that effort affects the distribution of available options (i.e., higher effort makes a draw of a high-quality option more likely). Formally, let $\gamma(e)$ be the probability that an option's quality is high given effort level e, with $\gamma'(e) > 0$ for all e. Suppose that the agent observes a signal of quality $\theta \in \{\theta_\ell, \theta_h\}$, which is informative (i.e., the probability of observing a high signal when the quality is high is larger than when it is low), and is independent of effort. Then it is immediate that $\pi'(e) > 0$, and the optimal contract *always* exhibits $w_h \geq w_\ell = 0$, i.e., the agent gets paid only when the principal buys.[9]

[9] In some applications (e.g., when the agent is a real-estate agent or a headhunter) effort can affect both the *selection* of an option (the expected quality), and its *evaluation* (the informativeness of the signal). Our theory predicts that when the agent's effort affects both margins, incentives for the two margins are aligned only when the quality of the pool of options is sufficiently improved, i.e., when γ exceeds $\hat{\gamma}$.

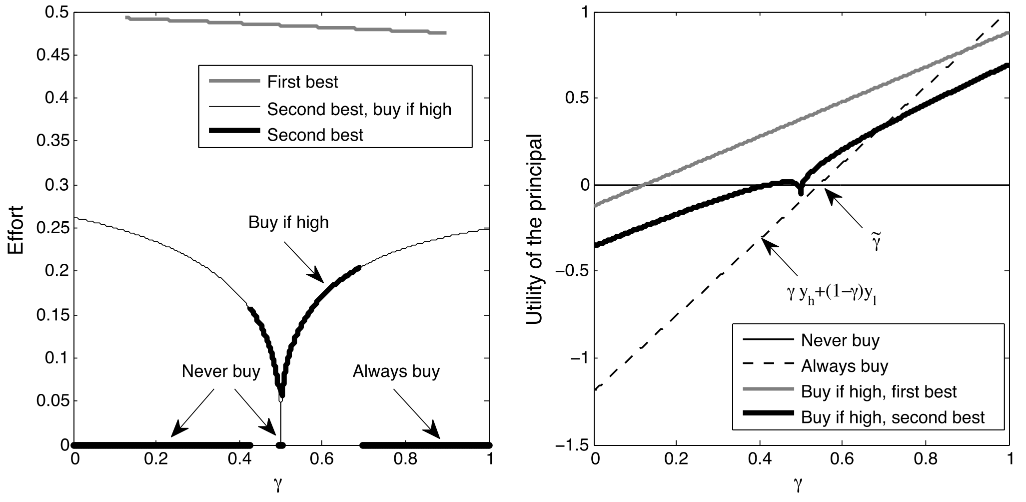

Fig. 1. **Effort and profits.** The left panel depicts the first- and second-best effort. The right panel shows the principal's profits, i.e., upper envelope of the expected payoff of the three alternatives, in the first- and second-best cases as a function of γ. Parameter values and functions: $\alpha = 1/2$, $\beta_h = \beta_\ell = 1$, $\eta(e) = e$, $\psi = 4e^5$, $y_h = 1$, and $y_\ell = -1/2$. In this case, $\hat{\gamma} = .5$ and $\tilde{\gamma} = .545$.

Regarding the implemented effort, moral hazard and limited liability distort it downward from its first-best level. This is a standard result in moral-hazard problems in which contracts are conditioned on a binary outcome (e.g., high or low signal realization).

More interesting is the distortion in the principal's decision to buy. First, as in the first-best case, she never exercises an option when γ is close to zero, and always exercises it when γ is close to one. We show that the intervals of γ at both ends are larger than in the first-best case, due to the lower implemented effort. Second, as γ approaches $\hat{\gamma}$, the endogenous cost of inducing positive effort diverges to infinity and precludes any incentive provision. Overall, the principal implements a positive level of effort (i.e., induces nontrivial information acquisition) less often than in the first-best case.

This begs the question of whether there are values of γ where the principal implements positive effort at all. A natural candidate is a neighborhood of $\tilde{\gamma}$, since at that point the principal is indifferent between buying and not buying without information, and thus information is highly valuable at this point. In the knife-edge case when $\gamma = \hat{\gamma} = \tilde{\gamma}$, the principal will not implement effort at $\gamma = \tilde{\gamma}$. But so long as $\hat{\gamma} \neq \tilde{\gamma}$, the optimal effort is positive at $\gamma = \tilde{\gamma}$ (and in a neighborhood of it).[10] Still, when $\tilde{\gamma}$ is close enough to $\hat{\gamma}$, there are priors between $\tilde{\gamma}$ and $\hat{\gamma}$ where information is very valuable, yet it is not acquired.

Fig. 1 illustrates the optimal effort and decision to buy as functions of γ in the first- and second-best cases. Effort in the latter is lower than in the former, and it is positive for intermediate values of γ outside an interval around $\hat{\gamma} = 1/2$. Also, the principal buys after the high signal realization on a smaller set of values of γ in the second best.

[10] The reason is that the marginal cost of inducing a small amount of effort is zero (since $\psi'(0)/\eta'(0) = (\psi'/\eta')'(0) = 0$), while its marginal benefit is positive. Note that $\hat{\gamma} \neq \tilde{\gamma}$ is sufficient but not necessary for effort to be positive for some priors.

3.2.1. Comparison with the standard principal–agent model

An insight that percolates throughout the paper is that sometimes the principal rewards the agent when she does not buy, which happens when the signal realization is low. A simple explanation is that the signal plays a dual role (see Grossman and Hart, 1983, p. 9). First, it affects the principal's decision to buy, and in this role she prefers a higher signal realization. Second, it provides information about the agent's effort, and in this role she might prefer a higher or a lower signal depending on which one is more informative about effort. Thus, when the principal's preferences regarding these roles are not aligned, rewarding for bad news ensues. This paper provides a natural class of principal–agent problems with moral hazard where a conflict between these roles emerges *endogenously*.

To support this claim, we illustrate the difficulties in obtaining a similar result in a standard principal–agent problem with moral hazard and two outcomes, success and failure. Let success yield revenue $q > 0$ while failure, for simplicity, yields 0. Also, let $\pi(e)$ be the probability of success. The rest is as in our model.

If one *exogenously* assumes that $\pi'(e) < 0$, then clearly the optimal contract that implements any positive level of effort pays a higher wage if failure is observed, and thus exhibits rewarding for bad news. But in this case the optimal level of effort the principal implements is trivially zero, since expected revenue $\pi(e)q$ decreases with effort.

One could assume that success yields a random output whose mean $\bar{q}(e)$ is increasing in e, and that contracts can only be conditioned on success or failure but not on the observed output. But since $\pi'(e) < 0$, one needs another exogenous assumption to ensure that $\pi(e)\bar{q}(e)$ does *not* decrease in e, otherwise the optimal choice would again be $e = 0$.

Similar remarks apply to our no-effort-around-$\hat{\gamma}$ result, where as γ approaches $\hat{\gamma}$, the likelihood ratio $\pi'(e)/\pi(e)$ goes to zero for all e, and hence effort implementation becomes prohibitively costly for the principal (i.e., the role of the signal as information about effort disappears). This has no natural counterpart in the standard principal–agent model, since π is a primitive of the problem instead of being 'micro-founded' via the prior and the signal distribution, as in our model. And although one can construct instances of the standard model where the principal prefers not to hire the agent due to moral hazard costs, they would require assumptions on the primitives (such as little variation of the likelihood ratio), rather than occurring endogenously as in our model. Moreover, in our model effort is zero only for some (endogenously determined) set of values of γ. Since in the standard model π is a primitive, there is no analogous parameter depending on which there could be no incentive provision.

4. The dynamic case

We now turn to the dynamic version of the model; that is, $\delta > 0$ henceforth.

4.1. Observable effort

As in the static case, we start with the first-best case in which the principal acquires information herself. The problem becomes a simple optimal stopping exercise. Since the environment is stationary, so is the principal's strategy. Moreover, as in the static case, there is no benefit for the principal from using randomization. Thus, the problem once again reduces to the analysis of the three relevant scenarios: (a) the principal does not exert effort and never buys, which yields zero profits; (b) she does not exert effort and buys the first option, which yields $\gamma y_h + (1-\gamma)y_\ell$; and ($c$) she exerts a constant positive level of effort in each period and buys the first time a signal realization is high.

Let S^* be the expected discounted profits for the principal under strategy (c). Then S^* satisfies the recursion $S^* = \max_e -\psi(e) + \pi(e)E_h(e) + (1-\pi(e))\delta S^*$. The first-order condition to this problem is $\psi'(e) = \eta'(e)(\gamma\beta_h y_h - (1-\gamma)\beta_\ell y_\ell - (\gamma\beta_h - (1-\gamma)\beta_\ell)\delta S^*)$. It is easy to verify that the level of effort that solves this equation is positive.

Let e^{**} be the optimal level of effort in the dynamic first-best case. The following result describes main properties of the solution to the principal's problem.

Lemma 2 *(Dynamic first best). There exist thresholds $\underline{\gamma}^{**}$ and $\bar{\gamma}^{**}$, satisfying $0 < \underline{\gamma}^{**} < -y_\ell/(y_h - y_\ell) < \bar{\gamma}^{**} < 1$, so that the principal:*

(i) Exerts no effort, i.e., $e^{**} = 0$, and does not exercise the option if $\gamma \in [0, \underline{\gamma}^{**}]$;

(ii) Exerts no effort, i.e., $e^{**} = 0$, and exercises the option if $\gamma \in [\bar{\gamma}^{**}, 1]$;

(iii) Exerts a positive level of effort, i.e., $e^{**} > 0$, and exercises the option if and only if the signal realization is high, if $\gamma \in (\underline{\gamma}^{**}, \bar{\gamma}^{**})$;

(iv) If $\gamma > \hat{\gamma}$, then $e^{**} \le e^*$ (with strict inequality if $e^* > 0$), $\underline{\gamma}^{**} > \underline{\gamma}^*$, and $\bar{\gamma}^{**} < \bar{\gamma}^*$. If $\gamma < \hat{\gamma}$, then $e^{**} \ge e^*$ (with strict inequality if $e^{**} > 0$), $\underline{\gamma}^{**} < \underline{\gamma}^*$, and $\bar{\gamma}^{**} > \bar{\gamma}^*$. If $\gamma = \hat{\gamma}$, then $e^{**} = e^*$, $\underline{\gamma}^{**} = \underline{\gamma}^*$, and $\bar{\gamma}^{**} = \bar{\gamma}^*$.

We stress the simple nature of the problem's solution in the first-best case, derived from the extremely basic search problem of the principal. Effort is constant and it is positive if and only if the principal plans to buy under the high but not under the low signal. This simple solution is all the more striking when compared to the richness of the dynamics and distortions that we obtain below, which are due exclusively to the moral-hazard problem that delegated information acquisition introduces.

4.2. Unobservable effort

We now proceed to analyze the principal's optimal contracting problem in the second-best case, i.e., when effort is unobservable and there is limited liability.

4.2.1. Recursive formulation of the contracting problem

Following a standard argument, one can rewrite the sequence problem corresponding to the optimal contract in recursive form, with the *promised value* to the agent as a state variable, and continuation values as control variables.[11] Crucially, in any period the promised value summarizes the history of play. As a result, one can recover the solution in sequence form from the recursive one, and thus such formulation is without loss of generality.

To this end, let $U(V)$ be the principal's value function when the promised value to the agent is V. Also, let w_{i1} denote the immediate payment from the principal to the agent if the principal exercises an option after signal θ_i (in which case the game ends). Finally, let w_{i0} and V_i be the agent's immediate payment and his continuation value if the principal does not exercise the option after signal θ_i.

[11] The classic references are Abreu et al. (1986), Abreu et al. (1990), and Spear and Srivastava (1987).

The maximization problem of the principal is as follows:[12]

$$U(V) = \max_{e,\{w_{i1},w_{i0},V_i,\sigma_i\}_{i\in\{h,\ell\}}} E\{\pi(e)[\sigma_h(E_h(e)-w_{h1})+(1-\sigma_h)(-w_{h0}+\delta U(V_h))] \\ +(1-\pi(e))[\sigma_\ell(E_\ell(e)-w_{\ell 1})+(1-\sigma_\ell)(-w_{\ell 0}+\delta U(V_\ell))]\} \quad (6)$$

$$\text{s.t. } E\{-\psi(e)+\pi(e)[\sigma_h w_{h1}+(1-\sigma_h)(w_{h0}+\delta V_h)] \\ +(1-\pi(e))[\sigma_\ell w_{\ell 1}+(1-\sigma_\ell)(w_{\ell 0}+\delta V_\ell)]\} = V, \quad (7)$$

$$\psi'(e) = \pi'(e)[\sigma_h w_{h1}+(1-\sigma_h)(w_{h0}+\delta V_h)-\sigma_\ell w_{\ell 1}-(1-\sigma_\ell)(w_{\ell 0}+\delta V_\ell)] \\ \text{for each realization } (w_{h1}, w_{\ell 1}, w_{h0}, w_{\ell 0}, V_h, V_\ell, \sigma_h, \sigma_\ell), \quad (8)$$

$$e \geq 0,\ w_{i1} \geq 0,\ w_{i0} \geq 0,\ V_i \geq 0,\ \sigma_i \in \Sigma \text{ for } i \in \{h, \ell\}, \quad (9)$$

where Σ is a subset of [0, 1] that includes 0 and 1. In what follows, besides the general case $\Sigma = [0, 1]$, we will also consider the case where the principal can only commit to purchase or not with certainty, i.e., $\Sigma = \{0, 1\}$. We will see that this case is tightly related to the case where the principal *cannot* commit to the purchase decision at all.

Constraint (7) is the *promise-keeping* constraint: the principal indeed has to deliver the promised value V to the agent.[13] Constraint (8) is the incentive constraint, which ensures that the agent optimally chooses the effort implemented by the principal. Finally, constraints (9) ensure that wages are nonnegative (limited liability) and similarly for the agent's promised values.[14]

A technical hurdle in solving this problem is that, although U is concave due to the use of lotteries, it need not be differentiable everywhere. Its presence in the objective function makes the problem nonsmooth, and we tackle it in Appendix A using superdifferentials.

4.2.2. *General properties of the optimal contract*

As in the static case, it follows from the incentive constraint that the cost of implementing any positive effort level goes to infinity as γ approaches $\hat{\gamma}$, thereby precluding any incentive provision. Also, it is easy to verify that $e = 0$ and $\sigma_h = \sigma_\ell = 0$ if γ is close to zero, and $e = 0$ and $\sigma_h = \sigma_\ell = 1$ if γ is close to one.

Before we derive the main properties of the optimal contract, we mention a useful result (see Lemma 7 in Appendix A), namely, that the value function U has slope greater than minus one. Indeed, suppose that the promised value to the agent increases by one unit. The principal can always deliver the extra value by increasing all current wages by one unit and continuation values by $1/\delta$, so that the marginal cost of an increase in the promised value to the principal is exactly one. By optimizing, she can potentially do better.

A straightforward implication of this result is that the principal weakly prefers to set $w_{h0} = w_{\ell 0} = 0$, as the marginal cost of compensating the agent with current wages is one, while the marginal cost of compensating him using continuation values (given by the absolute value of

[12] Since the principal may benefit from the use of stochastic contracts, the choice of wages, continuation values, and hence the induced level of effort, can be random variables. The expectation in the objective function and the promise-keeping constraints is taken with respect to these random variables. Wlog, we consider random variables with finite support as they are enough for our purposes.

[13] Note that, as is standard in models with moral hazard, the promise-keeping constraint has to be imposed with equality in all periods except the first period, when it is imposed with a weak inequality.

[14] The agent can always guarantee for himself zero discounted utility by exerting no effort in every period. Alternatively, nonnegativity of continuation values can be interpreted as the agent's participation constraints if we assumed that he can walk away from the contract and collect his zero reservation utility.

the slope of $U(V_i)$, $i \in \{h, \ell\}$) is less than or equal to one. Moreover, whenever the inequality is strict, setting $w_{i0} > 0$, $i \in \{h, \ell\}$, is not optimal so long as $\sigma_i < 1$, and thus the principal *strictly prefers* to postpone payments. (If $\sigma_i = 1$, then the choice of w_{i0} is irrelevant.) Similarly, whenever $\sigma_i < 1$, it is optimal to set $w_{i1} = 0$, $i \in \{h, \ell\}$.

This observation reveals an important difference between the static and dynamic cases. In the static case, the agent is compensated in the event of purchase if and only if $\gamma > \hat{\gamma}$. By contrast, in the dynamic case he gets paid at the *end* of the relationship (i.e., when the principal buys) for *all* values of γ. To provide the agent with incentives to exert effort, the level of the payment depends on the history, summarized by the promised value V.

Clearly, when the slope of U is minus one, the principal is indifferent between making payments now or in the future. We will see that the first best obtains in this case.

Define $V^* = \inf\{V|U'(V) = -1\}$. Using the above result, we will simplify the notation by setting w_{h0} and $w_{\ell 0}$ to zero and use w_h and w_ℓ instead of w_{h1} and $w_{\ell 1}$, respectively. Thus, in particular, the agent's incentive constraint becomes $\psi'(e) = \pi'(e)[\sigma_h w_h + (1-\sigma_h)\delta V_h - \sigma_\ell w_\ell - (1-\sigma_\ell)\delta V_\ell]$. Notice that the value $\sigma_i w_i + (1-\sigma_i)\delta V_i$, $i \in \{h, \ell\}$, is the expected payoff to the agent after signal i. The next proposition establishes some properties of the optimal contract.

Proposition 2 *(Optimal dynamic contract).*

(i) *Suppose that* $V \in (0, V^*)$.
 (a) *If* $\gamma > \hat{\gamma}$, *then* $\sigma_h w_h + (1-\sigma_h)\delta V_h \geq \sigma_\ell w_\ell + (1-\sigma_\ell)\delta V_\ell$, *with strict inequality if* $e > 0$. *Moreover, if* $e > 0$, *then* $V_h > V$ *(and* $w_h = 0$ *is optimal) whenever* $\sigma_h < 1$, *and* $V_\ell < V$ *(and* $w_\ell = 0$ *is optimal) whenever* $\sigma_\ell < 1$.
 (b) *If* $\gamma < \hat{\gamma}$, *then* $\sigma_h w_h + (1-\sigma_h)\delta V_h \leq \sigma_\ell w_\ell + (1-\sigma_\ell)\delta V_\ell$, *with strict inequality if* $e > 0$. *Moreover, if* $e > 0$, *then* $w_h = 0$ *is always optimal,* $V_h < V$ *whenever* $\sigma_h < 1$, *and* $V_\ell > V$ *(and* $w_\ell = 0$ *is optimal) whenever* $\sigma_\ell < 1$.

(ii) *If* $V = 0$, *then* $e = 0$, $\sigma_h w_h + (1-\sigma_h)\delta V_h = 0$, *and* $\sigma_\ell w_\ell + (1-\sigma_\ell)\delta V_\ell = 0$. *Moreover,* $\sigma_h = \sigma_\ell = 1$ *(and* $w_h = w_\ell = 0$*) if* $\gamma > \tilde{\gamma}$, *while* $\sigma_h = \sigma_\ell = 0$ *(and* $V_h = V_\ell = 0$*) if* $\gamma < \tilde{\gamma}$. *If* $\gamma = \tilde{\gamma}$, *then any values of* σ_h *and* σ_ℓ *are optimal.*

(iii) *If* $V^* > 0$, *then for all* $V \geq V^*$ *the first-best effort level* e^{**} *is implemented in every period.*[15]

As in the static case, when $\gamma < \hat{\gamma}$, the agent is rewarded for bad news (part (i)-(b)). A new feature of the dynamic case is that he is not compensated immediately after a low signal realization, but via an increase in his continuation value (i.e., future payments). Similarly, when $\gamma > \hat{\gamma}$, the agent is punished for bad news by a decrease in his continuation value (part (i)-(a)). In particular, if effort is positive and $\sigma_h = 1$ and $\sigma_\ell = 0$, then $w_h = 0$ and $V_\ell > V$ if $\gamma < \hat{\gamma}$, while $w_h > \delta V_\ell$ and $V_\ell < V$ if $\gamma > \hat{\gamma}$. As mentioned, the first case ($\gamma < \hat{\gamma}$) is specific to our model where the agent's effort affects the signal informativeness.

Part (ii) reveals that $V = 0$ is an absorbing state, since (when the principal continues) all continuation values are set to zero. The agent exerts no effort and does not get paid.

Finally, part (iii) shows that when the agent's promised value is sufficiently high, the principal implements the first-best level of effort from that period onwards. The basic reason is that in this case limited liability is no longer binding. Since the first best is achieved, she only buys after the high signal ($\sigma_h = 1$ and $\sigma_\ell = 0$). The timing of payments (i.e., how the principal uses w_h vs. V_ℓ)

[15] The threshold V^* is positive whenever the principal implements $e > 0$ at the optimum for some V.

from this point onward is indeterminate, as the principal is indifferent between paying now and in the future (for $U'(V) = -1$ for $V \geq V^*$).

Notice that V^* equals zero when the principal implements $e = 0$ for *all* $V \geq 0$. This occurs if $\gamma \in [0, \underline{\gamma}^{**}] \cup [\bar{\gamma}^{**}, 1]$ or is close enough to $\hat{\gamma}$. This reveals an interesting point. Implementation of the first best for high enough V is standard in models with risk neutrality. A novel insight here is that around $\hat{\gamma}$, not only the first best is not achieved, but also positive effort cannot be implemented, regardless of how high the promised value V is.

When $\gamma < \hat{\gamma}$, the optimal contract prescribes no payments until the implementation of the first-best effort (after which point the timing of payments is indeterminate). One way to implement the first-best effort when $\gamma < \hat{\gamma}$ is with a fixed per-period payment independent of the signal realization. This generates incentives to deliver bad news, because after bad news the search continues, which allows the agent to remain employed by the principal.

This is in line with what we observe with home inspectors. Indeed, in reality home inspectors are paid a given amount per inspection, independently of the report. It is important to stress, though, that this contract would *not* implement any positive effort level if $\gamma > \hat{\gamma}$, which is arguably the less relevant case in this application. For while the inspector's effort might be important for finding out good news (e.g., the house has been recently insulated), it is perhaps mostly crucial for finding out bad news (e.g., there is a crack in the roof).

Although in our model γ is fixed, it would not be hard to allow it to vary over time.[16] Then whether the optimal contract rewards the agent for good or bad news will also vary over time, depending on whether the realization of γ falls above or below $\hat{\gamma}$. Also, when the realization of γ is close enough to $\hat{\gamma}$, the agent's services will not be used in that period.

In the context of the home-inspector example, the changing γ reflects the fact that the potential buyer might be considering houses of different 'perceived' quality.[17] This extension suggests an interpretation of our result of no effort around $\hat{\gamma}$. When a house comes along with the prior close to the critical point where the inspector's effort does not change how likely he is to find out good vs. bad news about the house, the potential buyer will not use the inspector's services.

4.2.3. *The* $\Sigma = \{0, 1\}$ *case*

Proposition 2 applies to *any* set Σ that contains 0 and 1, thus potentially allowing the principal to commit to lotteries over the purchase decision. In this section we consider a special case, namely, $\Sigma = \{0, 1\}$, where the principal can only commit to buy or not to buy with probability one after each signal realization.[18] Besides providing insights about the optimal contract in these 'corners,' which are also part of the general set $\Sigma = [0, 1]$, we show that this case is outcome-equivalent to the case in which the principal cannot commit at all to the purchase decision. We assume throughout, however, that the principal can commit to lotteries over contracts, as we did in the previous section.[19]

Given this restriction on Σ, there are four relevant cases to consider (plus convex combinations thereof, due to the use of random contracts). Indeed, (a) the principal can commit not to

[16] In the recursive formulation, the realization of γ would enter as another state variable, and the choices of the continuation values will be contingent on the next period's realization of γ.

[17] Notice that γ reflects not just the quality of the house, but rather the quality relative to the price as well as how well the house matches the potential buyer's criteria.

[18] This could happen because, say, the outcome of the randomizing device that determines σ_i, $i \in \{\ell, h\}$, is not observable to a third party and hence cannot be contracted upon.

[19] Allowing for such lotteries is crucial for our analysis as it ensures that the value function is concave.

buy today after either signal realization (and continue optimally from tomorrow onward), a strategy whose expected discounted profit to the principal when the agent's promised utility is V is denoted by $U_{00}(V)$; (b) she can commit to buy today after either signal realization, with the expected discounted profit denoted by $U_{11}(V)$; (c) she can commit to buy today after a high signal realization and not to buy after a low one, with expected discounted payoff denoted by $U_{10}(V)$; and (d) she can commit to buy today after a low signal realization and not to buy after a high one, with expected discounted payoff denoted by $U_{01}(V)$.

Notice that under alternative (b), the principal buys under both signal realizations, and thus the game ends with probability one. It follows that she will never implement positive effort in this case, and the value of this alternative to the principal is simply $U_{11}(V) = \gamma y_h + (1-\gamma)y_\ell - V$.

Regarding alternative (c), the main properties of the contract are the ones described in Proposition 2 once we specialize to this corner. In addition, we can show that whenever the principal finds it optimal to use this strategy, she implements positive effort. Intuitively, if she acquired no information, either (a) or (b) would dominate this strategy.

Intuition suggests that alternative (d) must be dominated by the other ones. After all, it does not seem to use information in an efficient way. This intuition might be misleading, however, since committing to an ex-ante inefficient decision may pay off to the principal if it leads to a higher continuation value or if it relaxes the constraint set. We show that alternative (d) is indeed dominated and thus irrelevant for the problem at hand. The proof is long and somewhat involved, but its structure is straightforward: we first prove that under alternative (d) the effort implemented by the contract is zero; given this, it follows at once that either alternative (a) or (b) yields a higher payoff. We thus have the following claim.

Claim 1. *In the optimal long-term contract, the principal never commits to buy after a low signal realization and not to buy after a high signal realization.*

Another insight concerns alternative (a). We show that even when the principal chooses not to buy after either signal realization, she sometimes finds it optimal to implement positive effort. This is a priori counterintuitive, especially if one thinks about the first-best case, where the principal would never acquire costly information if she plans not to buy under any signal realization. But under moral hazard there is a subtle countervailing effect: by suitably structuring continuation values for the agent, the expected future surplus from the relationship can go up, and this provides enough incentives for the principal to implement positive effort. We pin down exactly where this occurs when U is sufficiently smooth:[20]

Claim 2. *Suppose that $V > 0$ and the principal optimally commits not to buy after either signal realization. If U is decreasing, then the principal implements zero effort at V. And if U is strictly increasing and twice continuously differentiable at V, then the effort level implemented at V is positive.*[21]

[20] This assumption is only used when applying L'Hopital's Rule at one point in the proof.

[21] The value function U can increase in V over some region because the promise-keeping constraint is imposed with equality. In this region, the optimal contract is not "renegotiation-proof," i.e., the corresponding payoff profile does not lie on the Pareto frontier, so both parties benefit from increasing V. If (7) were imposed with a weak inequality, the principal could assign high continuation values without wasting effort (i.e., with $e = 0$). But to satisfy (7) with equality, higher continuation values require positive effort.

This result is not related, for example, to testing the accuracy of the agent. Indeed, since in our model there is no learning about the pool of the options, effort in any given period serves only two purposes: it improves information about the current option, and allows the principal to compensate the agent with higher continuation values in order to improve information in future periods. When the principal does not buy after either signal, she does not use the first purpose of effort, but may want to use the second one.

All told, when the principal can commit to corners only, the values U_{00}, U_{11}, and U_{10} (plus convex combinations of these alternatives) are the relevant ones to consider. This begs the following questions: Which of the three alternatives is used at a given value V? When does the principal use convex combinations of the alternatives (lotteries)? Part (*ii*) of Proposition 2 tells us that either U_{11} or U_{00} is optimal at $V = 0$ depending on whether γ is above or below $\tilde{\gamma}$. Also, part (*iii*) of Proposition 2 implies that $U = U_{10}$ for V above V^* (when $V^* > 0$). In fact, one can show that $U(V) = U_{10}(V)$ for $V \geq \bar{V}$, where $\bar{V} < V^*$. In particular, when U has an upward sloping portion, $\bar{V}$ is the promised value at which the slope of U is zero. (See Claim 4 in Appendix A.) The principal will use convex combinations of U_{10} and either U_{11} or U_{00} (on a set of promised values) below $\bar{V}$.

It is straightforward to show that the $\Sigma = \{0, 1\}$ case is outcome-equivalent (in terms of on-the-equilibrium-path payments and effort) to the case when the principal cannot commit at all to the purchase decision. For example, although payments can be enforced in court, there may not be a mechanism in place to enforce the purchase decision, since ex post this decision only affects the principal and no other parties. In the latter case, the purchase decision must be ex-post optimal after each signal realization, and thus the corresponding ex-post optimality constraints must be added to the problem.

Since the principal can still commit to lotteries over contracts, it is easy to see that these constraints can be made slack by suitably choosing off-the-equilibrium-path payments. To illustrate, in the U_{10} problem the ex-post optimality constraints are $E_h(e) - w_{h1} \geq -w_{h0} + \delta U(V_h)$ and $-w_{\ell 0} + \delta U(V_\ell) \geq E_\ell(e) - w_{\ell 1}$. Since w_{h0} and V_h do not enter into payoffs of either the principal or the agent (as $1 - \sigma_h = 0$), the principal can set w_{h0} and/or V_h arbitrarily high so that the payoff from not buying after the high signal, $-w_{h0} + \delta U(V_h)$, is arbitrarily low. Similarly with $w_{\ell 1}$.[22] Moreover, the proof of domination of U_{01} does not rely on the assumption of commitment, and hence that result holds in the no-commitment case as well. Since the same alternatives (*a*)–(*c*) are relevant with and without commitment, and since ex-post constraints can be freely satisfied, the two cases are outcome-equivalent.

The case of $\Sigma = \{0, 1\}$ may be also relevant in applications. For instance, in the home-inspector example, it might be hard for a potential home buyer to commit not to buy a house when it is ex-post optimal to do so.

4.2.4. *The* $\Sigma = [0, 1]$ *case and the optimality of lotteries over purchase*

Besides the $\Sigma = \{0, 1\}$ case, an interesting issue to explore is whether the principal could benefit if she were able to commit to 'interior' purchase lotteries. This is a difficult question to answer since purchase lotteries interact with lotteries over contracts, and it is unclear a priori

[22] The only subtlety is when the principal sets things up so as to be indifferent ex post between buying and not buying, in which case she could randomize. However, it is easy to show that in this case the principal can do just as well by following a deterministic purchase decision.

whether the former are redundant given the latter.[23] We show, however, that there is a set of parameter values such that they are indeed used in equilibrium.

We first state an intermediate result, which asserts that there is a threshold effort level such that $\sigma_h = 1$ is optimal if and only if the implemented effort is above that threshold.

Lemma 3. *Assume $\gamma > \hat{\gamma}$, and let $\tilde{e}$ be the unique solution to $E_h(\tilde{e}) = \delta S^*$.[24] Let V be such that the principal implements positive level of effort. Then $\sigma_h = 1$ is optimal if and only if the implemented effort level e is above $\tilde{e}$. Moreover, if $e \in (0, \tilde{e})$, then σ_h arbitrarily close to one cannot be optimal.*

We are now ready to state the main result of this section.

Proposition 3. *Assume $\gamma > \tilde{\gamma} \geq \hat{\gamma}$, $(\psi'/\eta')''(0) > 0$, and that the principal implements positive effort at least for some V. Then there is an interval of small values of $V > 0$ such that the optimal contract involves an interior purchase lottery with $\sigma_h \in (0, 1)$.*

The proof is long and technical, but it is easy to explain its main steps and the use of the assumptions made in the statement of the proposition. Suppose that the principal only uses contracts with deterministic payments (and hence implements a deterministic effort level) in the current period. The promised value V restricts the set of effort levels that can be implemented. In particular, when V is small enough, the implemented effort must be below $\tilde{e}$, and hence by Lemma 3 we have $\sigma_h < 1$. We show further that when $\gamma > \tilde{\gamma}$ so that $E_h(0) = E_\ell(0) > 0$, $\sigma_h = 0$ is not optimal for sufficiently small levels of V. Therefore, σ_h must be interior.

The proof is not complete though, for we still need to rule out a possibility that the interior purchase lottery is actually *not* used in equilibrium. In particular, it could be dominated by a lottery over contracts (i.e., lottery over payments) each of which involves a deterministic purchase decision. We show that imposing the additional assumption $(\psi'/\pi')''(0) > 0$ (together with $\gamma > \tilde{\gamma}$) rules this out.

Intuitively, why is it optimal for the principal to sometimes commit to a random purchase decision? The underlying reason is that it is cheaper for the principal to pay to the agent in the future rather than today (formally, the marginal cost of current payments is equal to one, and it exceeds the marginal cost of future payments, given by $|U'|$). Recall that if $\gamma > \hat{\gamma}$ and the principal implements positive effort and buys with certainty after the high signal realization, then she optimally sets $w_h > 0$ (see part (a) of Proposition 2). Suppose instead that the principal continues the search after the high signal realization with some probability. Since it is optimal to postpone payments, she will now make zero immediate payment. And by choosing $V_h = w_h/[\delta(1 - \sigma_h)]$, where w_h is the immediate payment in the old contract, she can implement the same effort as before. Now, whether the principal can actually increase profits by using an interior lottery is a bit more subtle. For if σ_h is chosen to be arbitrarily close to one, then $V_h \geq V^*$ and hence $U'(V_h) = -1$; as a result, the principal saves nothing by postponing the payment in this case (see the second part of Lemma 3). But if she chooses σ_h so that $U'(V_h) > -1$, then

[23] One may also wonder whether lotteries over contracts might be redundant given the purchase lotteries. If the signal were continuous, then an argument like the one in Spear and Srivastava (1987) would show that the value function is concave, so lotteries over contracts would be redundant. With a discrete signal, however, these lotteries are needed to make the value function concave.

[24] It is easy to check that $\tilde{e} < e^{**}$, where e^{**} is the dynamic first-best effort level.

she can implement the same effort level at a lower cost. When this effect just offsets the loss of surplus due to the ex-post suboptimal purchase decision, the optimality of $\sigma_h \in (0, 1)$ ensues.

The above intuition suggests that the optimality of interior purchase lotteries crucially depends on the asymmetry embedded in our model: once the principal buys, the game is over. It is easy to show that if instead the game always continued (i.e., if the principal could exercise multiple options over time), then interior purchase lotteries would not be optimal. More precisely, after any signal realization, the principal would set the same continuation values after buying and not buying. She no longer reduces cost by not buying, and hence there is no benefit in committing to an ex-post suboptimal purchase decision.

Clearly, unless the principal is just indifferent between buying and not buying after signal θ_i, a lottery with an interior σ_i is ex-post suboptimal. We now explore whether ex post the principal prefers to buy (i.e., $E_i(e) - w_i \geq \delta U(V_i)$) or not to buy (i.e., $E_i(e) - w_i \leq \delta U(V_i)$). Claim 3 shows that either situation may occur depending on the continuation promised value to the agent.[25] Before stating the result, it will be useful to present the first-order condition with respect to σ_i, $i \in \{h, \ell\}$, which for an interior solution is given by (see Appendix A.4 for the derivations)

$$E_i(e) - \delta U(V_i) + \delta U'(V_i)V_i = 0, \tag{10}$$

or equivalently

$$E_i(e) - \delta S(V_i) + \delta S'(V_i)V_i = 0, \tag{11}$$

where $S(V) = U(V) + V$ is the social surplus when the promised value to the agent is V.[26]

Claim 3. *Suppose $\sigma_i \in (0, 1)$ is optimal for some $i \in \{h, \ell\}$. Then it is ex-post optimal for the principal to buy if and only if U is strictly decreasing at V_i. However, from the social point of view it is always ex-post optimal not to buy.*

The first part of the result immediately follows from (10) (recall that $w_i = 0$ when $\sigma_i < 1$), and the second part from (11). Notice that from the social point of view, whenever an interior purchase lottery is used, it is optimal not to buy ex post (i.e., $E_i(e) \leq \delta S(V_i)$, with strict inequality unless $V_i \geq V^*$ for then $S'(V_i) = 0$). The reason is that an increase in the probability of buying, σ_i, is accompanied by an increase in the corresponding continuation value to the agent, V_i (as a way to satisfy the constraints while keeping everything else unchanged), which in turn increases the continuation social surplus, $S(V_i)$. The principal takes this effect into account ex ante when committing to a lottery, but not ex post.

We have provided conditions under which interior lotteries are used in equilibrium. It is very hard and beyond the scope of this section to provide a full-blown characterization of which values of σ_i's will be optimal for a given value V. But in some cases we can say more about the use of these lotteries. For instance, one can show that if $\gamma \leq \tilde{\gamma}$, then $\sigma_\ell = 0$ for all V. Furthermore, if $\gamma < \hat{\gamma}$ and $e > 0$, then $\sigma_h \geq \sigma_\ell$, and both σ_i's cannot be interior. (See Claims 5 and 6 in Appendix A.) In fact, in the $\gamma < \hat{\gamma}$ case our numerical computations always delivered corner solutions for σ_h and σ_ℓ.

[25] In our numerical computations, both situations can occur in equilibrium.

[26] Here for simplicity we write the equations assuming that the value function is differentiable at V_i. The derivations in Appendix A do not use this assumption.

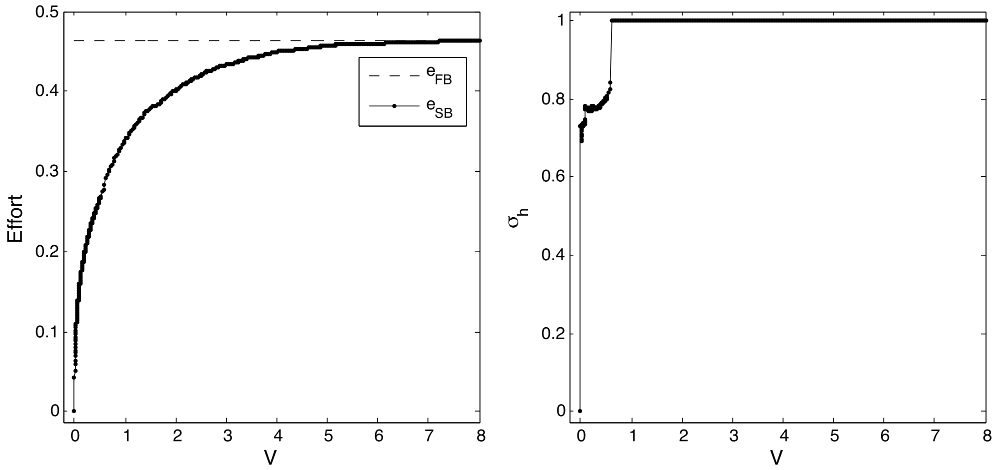

Fig. 2. **Optimal dynamic contract.** The left panel depicts the optimal effort. The right panel shows the optimal probability of purchase after the high signal. Parameter values and functions: $\alpha = 1/2$, $\beta_h = \beta_\ell = 1$, $\eta(e) = e$, $\psi = 3e^4$, $y_h = 1$, $y_\ell = -2.5$, $\delta = .9$, and $\gamma = 0.7$, which is above $\hat{\gamma} = 0.5$ and below $\tilde{\gamma} \approx 0.714$.

Fig. 2 illustrates the optimal choices of σ_i's using numerical computations. The left and right panels depict the optimal choices of effort and σ_h, respectively, as functions of V. The parameter values are listed in the figure's caption. While we only show the numerical results for specific parameter values, the figures are representative in the sense that the curves look similar for many other choices of parameters.

Since $\gamma < \tilde{\gamma}$, it follows that $\sigma_h = \sigma_\ell = 0$ at $V = 0$ (see part (*ii*) of Proposition 2), and $\sigma_\ell = 0$ is optimal for all V (see Claim 5 in Appendix A). Moreover, σ_h is interior up to a promised value $V \approx 0.6$ for which $e(V) = \tilde{e}$ (see Lemma 3), at which point σ_h discretely jumps to one. Finally, the first-best effort is implemented above $V^* \approx 7.2$ (see part (*iii*) of Proposition 2). It is worth pointing out that two of the assumptions of Proposition 3 are not met in this numerical example, namely, the assumptions that $\gamma > \tilde{\gamma}$ and $(\psi'/\pi')''(0) > 0$. Notice, however, that σ_h is still interior for small enough promised values, thereby illustrating that the conditions of Proposition 3 are merely sufficient.[27]

4.2.5. *Optimal effort*

Finally, we conclude the analysis of the properties of the optimal contract with a comparison of the constrained optimal level of effort with the first-best effort. In the static case we showed that the optimal effort was always below its first-best level. This property partially extends to the dynamic case as well.

Proposition 4 *(Optimal effort). If $\gamma < \hat{\gamma}$, then $e(V) \le e^{**}$ for all V.*

[27] We are not sure if the interior portion of the σ_h curve appears jagged due to a numerical error, or it is in fact non-monotone. Notice though that in some cases σ_h can indeed be non-monotone in V. For example, if $\gamma > \hat{\gamma}$ and $\gamma > \tilde{\gamma}$, then $\sigma_h = 1$ at $V = 0$, $\sigma_h < 1$ whenever $e \in (0, \tilde{e})$ (for $V > 0$ small enough) and $\sigma_h = 1$ when $e > \tilde{e}$ (for V large enough).

To understand the intuition behind this result and why it is only stated for $\gamma < \hat{\gamma}$, consider the marginal cost and marginal benefit of effort in the first- and second-best cases. Regardless of γ, the marginal cost of effort is higher in the second-best case due to moral hazard. Also, when $\gamma < \hat{\gamma}$, higher effort makes continuing the search more likely. Since future social surplus in the second best is lower than in the first best, the social benefit of continuing the search is also lower, and hence the marginal benefit of effort is lower. And since the marginal cost is higher and the marginal benefit is lower, lower effort obtains. When $\gamma > \hat{\gamma}$, however, higher effort makes ongoing search less likely, but stopping yields a lower loss of future social surplus relative to the first best. That is, there are two effect that go in opposite directions, making the comparison of effort levels ambiguous.[28]

This result is noteworthy given (to the best of our knowledge) the lack of formal results on the comparison between first- and second-best effort levels in dynamic moral hazard problems, be it with a risk-averse agent or with a risk-neutral one and limited liability.[29] Moreover, we obtain an unambiguous answer exactly when $\gamma < \hat{\gamma}$, which is the novel case that pertains to our model, where effort affects the informativeness of the signal.

5. Concluding remarks

We study a problem where a principal searches for an opportunity of uncertain return, and hires an agent to acquire information about potential options. Information precision depends on the agent's unobservable effort. Based on the information the agent provides, the principal decides whether to stop and exercise the option, or to continue the search.

We provide a complete analysis of the optimal contract both in the static and dynamic cases. Among the main properties of the contract, we highlight the following ones. First, effort provision and the optimal buying decision are distorted compared to the first best. This is due to the cost of providing incentives for effort. In particular, there is a region of the prior belief that the option quality is high, where positive effort cannot be sustained. The reason is that effort has no (or little) impact on the unconditional distribution of the signal, and hence on the likelihood ratio, thereby making incentive provision infinitely costly. Second, for prior beliefs just below that region, the optimal contract calls for rewarding for bad news, i.e., the agent receives a higher compensation in the event in which the information revealed induces the principal to pass on the option at hand. This property can be traced back to the role of the agent's effort in affecting the *informativeness* of the signal about the option quality, and has no counterpart in standard moral-hazard models where an agent's effort affects the distribution of output. Third, we find that the principal may find it optimal to commit to a random (and thus ex-post suboptimal) purchase decision. This result is due to the asymmetry that the purchase decision entails: the game ends if the principal buys, but continues otherwise, and postponing payments benefits the principal.

We assume that the fraction of high quality options is known, so that the parties do not learn about the pool of options over time. We also assume that there is no recall of passed options.

---

[28] However, in all our numerical computations the effort was always below the first best.

[29] In the static case, it is well known that with a risk-neutral agent, limited liability, and two outcomes, the first-best effort level is higher than the second-best one. With a risk-averse agent, however, it is easy to construct examples with at least three effort levels where the opposite is true, even with binary outcomes. (See MasColell et al., 1995 for one example; for another one, let $e \in [0, 1]$, assume the agent's utility is $\log w - 0.5e^2$, his reservation utility is 2, the high outcome is 14 and the low one is 4. Then second-best effort is 1 while the first-best level is about 0.9.) Since in our dynamic case continuation adds a concave term to the objective function, the resulting problem bears some resemblance with the static case with a risk-averse agent. Thus, it is intuitive to surmise that a similar ambiguity can arise.

Relaxing either assumption present nontrivial challenges, as deviations by the agent can lead to divergence of the agent's and principal's beliefs. These seem to be interesting problems for future research.

Appendix A. Omitted proofs

A.1. Static first best: proof of Lemma 1

The principal solves $\max\{0, \gamma y_h + (1-\gamma)y_\ell, \max_e u_h(e) - \psi(e)\}$. Notice that $\gamma y_h + (1-\gamma)y_\ell \geq 0$ if and only if $\gamma \geq \tilde{\gamma}$. At $\gamma = \tilde{\gamma}$ the principal is indifferent between exerting no effort and not exercising the option, and exerting no effort and exercising it. But at that value of γ she strictly prefers to exert positive effort and exercise the option if the signal realization is high. To see this, notice that at $\gamma = \tilde{\gamma}$ the first-order condition becomes $\psi'(e) = -\eta'(e)(\beta_h + \beta_\ell)y_\ell y_h/(y_h - y_\ell)$, which has a unique solution $e > 0$. Since the principal can always obtain zero profits by choosing $e = 0$, and her problem is strictly concave in effort, it must be the case that $u_h(e) - \psi(e) > 0$ at the optimal level of effort when $\gamma = \tilde{\gamma}$.

At $\gamma = 1$ ($\gamma = 0$) the principal's payoff if she exerts no effort and buys (does not buy) is y_h (zero), which is strictly higher than the other two alternatives. Hence, at $\gamma = 0$ the principal does not exert effort and does not buy; at $\gamma = \tilde{\gamma}$ she exerts a positive level of effort and buys if the signal is high; and at $\gamma = 1$ she exerts no effort and buys.

We now show that $\max_e u_h(e) - \psi(e)$ is strictly increasing in γ. By the Envelope Theorem, $\partial(\max_e u_h(e) - \psi(e))/\partial\gamma = (\alpha + \beta_h\eta(e))y_h - (\alpha - \beta_\ell\eta(e))y_\ell > 0$. Thus, $\max_e u_h(e) - \psi(e)$ is strictly increasing in γ. It is also convex in γ, a well-known property of the value of information. Thus, it must single-cross zero at $\underline{\gamma}^* \in (0, \tilde{\gamma})$ and $\gamma y_h + (1-\gamma)y_\ell$ at $\bar{\gamma}^* \in (\tilde{\gamma}, 1)$, which proves the interval structure and effort/buying decisions stated in parts (i)–(iii) of the lemma. □

A.2. Optimal static contract: proof of Proposition 1

Let λ and μ be the Lagrange multipliers associated with constraints (2) and (3). The first-order conditions with respect to w_h, w_ℓ, and e are

$$\lambda - 1 + \frac{\pi'(e)}{\pi(e)}\mu \leq 0, \qquad w_h \geq 0, \tag{A.1}$$

$$\lambda - 1 - \frac{\pi'(e)}{1-\pi(e)}\mu \leq 0, \qquad w_\ell \geq 0, \tag{A.2}$$

$$u_h'(e) - \psi'(e) + \mu\left(\pi''(e)\frac{\psi'(e)}{\pi'(e)} - \psi''(e)\right) \leq 0, \qquad e \geq 0, \tag{A.3}$$

all with complementary slackness, where we substituted constraint (3) into (A.3).

The following results will be used in the proof of the proposition:

Lemma 4. *At the optimum, $w_\ell = 0$ if $\gamma > \hat{\gamma}$, and $w_h = 0$ if $\gamma < \hat{\gamma}$.*

Proof. We only prove the first case (the other case is analogous). If $\gamma > \hat{\gamma}$, then $w_h \geq w_\ell$, and $w_\ell \geq 0$ is the only relevant limited-liability constraint. Suppose $w_\ell > 0$. Then $w_h > 0$, and (A.1) and (A.2) hold with equalities, implying $\mu = 0$ and $\lambda = 1$. Thus (2) and (3) yield $w_\ell = \psi(e) - \pi(e)\psi'(e)/\pi'(e)$, where the right-hand side equals zero at $e = 0$ and is strictly decreasing in e. Hence $w_\ell \leq 0$, a contradiction. □

Lemma 5. *At the optimum, constraint (2) does not bind unless $e = 0$.*

Proof. Let $\gamma > \hat{\gamma}$ (the other case is analogous), and suppose that (2) binds and $e > 0$. Then (3) and Lemma 4 yield $w_h = \psi'(e)/\pi'(e)$, and (2) becomes $-\psi(e) + \pi(e)\psi'(e)/\pi'(e) = 0$. But the left-hand side is positive if $e > 0$, a contradiction. □

Lemma 6. *Suppose the principal buys only after the high signal realization. Then for any $\gamma \neq \hat{\gamma}$, the principal implements positive effort. Moreover, the effort goes to zero as γ approaches $\hat{\gamma}$.*

Proof. Let $\gamma < \hat{\gamma}$ (the other case is analogous). Using the first-order conditions, we obtain $\mu = -(1-\pi(e))/\pi'(e)$, which strictly increases in γ. We can rewrite (A.3) as follows: $\gamma\beta_h y_h - (1-\gamma)\beta_\ell y_\ell - (\psi'/\eta')(e) + [(1-\pi(e))/\pi'(e)](\psi'/\eta')'(e) \leq 0$. Multiplying both sides by $\gamma\beta_h - (1-\gamma)\beta_\ell (< 0$ as $\gamma < \hat{\gamma})$ and rearranging terms, $(\gamma\beta_h - (1-\gamma)\beta_\ell)[\gamma\beta_h y_h - (1-\gamma)\beta_\ell y_\ell] \geq (\gamma\beta_h - (1-\gamma)\beta_\ell)(\psi'/\eta')(e) - [(1-\pi(e))/\eta'(e)](\psi'/\eta')'(e)$. Given our assumptions, the right-hand side is non-positive for any e and equals zero at $e = 0$. At the same time, the left-hand side is strictly negative for any $\gamma < \hat{\gamma}$. Thus if $\gamma \neq \hat{\gamma}$, then $e = 0$ cannot be a solution, and the principal must implement positive effort if she buys only after the high signal realization. But as γ approaches $\hat{\gamma}$, the left-hand side goes to $0 \times (y_h - y_\ell)\beta_\ell\beta_h/(\beta_\ell + \beta_h) = 0$, and hence effort must go to zero. □

We are now ready to prove the proposition:

Proof of Proposition 1. (i) Let $\gamma > \hat{\gamma}$ (the other case is analogous). Since $\pi'(e) > 0$, constraint (3) and Lemma 4 yield $w_h \geq w_\ell = 0$, with strict inequality if $e > 0$.

(ii) If $e = 0$, then it is trivially lower than e^* and strictly so if $e^* > 0$. Suppose $e > 0$ and let $\gamma > \hat{\gamma}$ (the other case is analogous). From Lemma 5, $\lambda = 0$ and hence $\mu = \pi(e)/\pi'(e) > 0$. Then (A.3) becomes $u_h'(e) - \psi'(e) = -(\pi(e)/\pi'(e))(\pi''(e)\psi'(e)/\pi'(e) - \psi''(e)) > 0$. Since e^* solves $u_h'(e^*) - \psi'(e^*) = 0$, it follows that $e < e^*$.

(iii)-(a) At $\gamma = 0$ it is optimal for the principal not to hire the agent and not to exercise the option (the other two alternatives yield a negative payoff as $y_\ell < 0$ and $\alpha > 0$). By continuity, this holds in a neighborhood of $\gamma = 0$. Let $\underline{\gamma}$ be the largest value such that the principal does not hire the agent and does not exercise the option for all $\gamma \in [0, \underline{\gamma}]$. To show that $\underline{\gamma} > \underline{\gamma}^*$, notice that at $\underline{\gamma}^*$, $u_h(e^*) - \psi(e^*) = 0$. Since $0 < e < e^*$, $u_h(e) - \psi(e) < 0$, thus proving that $\underline{\gamma} > \underline{\gamma}^*$.

(iii)-(b) At $\gamma = 1$ it is optimal for the principal not to hire the agent and to exercise the option (it yields y_h, strictly higher than the payoff of the other two alternatives as $\alpha < 1$). By an analogous argument as in (iii)-(a) it follows that $\bar{\gamma} < \bar{\gamma}^*$.

(iii)-(c) Suppose $\gamma = \hat{\gamma}$. Since the cost of implementing any positive level of effort is infinite, $e = 0$. Hence, if the principal exercises the option only if the signal realization is high, then her payoff is $\alpha(\hat{\gamma} y_h + (1-\hat{\gamma}) y_\ell)$. Notice, however, that in this case the principal prefers (strictly unless $\hat{\gamma} = \tilde{\gamma}$) not to hire the agent, and to exercise the option if and only if $\hat{\gamma} y_h + (1-\hat{\gamma}) y_\ell \geq 0$ (i.e., $\gamma = \hat{\gamma} \geq \tilde{\gamma}$).

We now show that the same holds in a neighborhood of $\hat{\gamma}$. Let $\gamma < \hat{\gamma}$ (the other case is analogous). The principal's problem in this case can be written as $v(\gamma) \equiv \max_{e \in [0, e^*]} u_h(e) + (1-\pi(e))\psi'(e)/[\eta'(e)(\gamma\beta_h - (1-\gamma)\beta_\ell)]$, where by part (ii) we restrict attention to $e \leq e^*$. Take $\varepsilon > 0$ small, and consider $\gamma \in [0, \hat{\gamma} - \varepsilon]$. The Theorem of the Maximum implies that v is continuous on $[0, \hat{\gamma} - \varepsilon]$. We will use this fact to show that there is a left neighborhood of $\hat{\gamma}$ such that the principal chooses not to hire the agent.

There are two cases to consider, $\hat{\gamma} y_h + (1-\hat{\gamma}) y_\ell < 0$ ($\hat{\gamma} < \tilde{\gamma}$) and $\hat{\gamma} y_h + (1-\hat{\gamma}) y_\ell > 0$ ($\hat{\gamma} > \tilde{\gamma}$). If $\hat{\gamma} y_h + (1-\hat{\gamma}) y_\ell < 0$, then the principal's profits at $\gamma = \hat{\gamma}$ if she only buys after a high signal are $\alpha(\hat{\gamma} y_h + (1-\hat{\gamma}) y_\ell) < 0$. From Lemma 6, the optimal effort level under this policy goes to zero as γ goes to $\hat{\gamma}$. Thus, if ε is small enough, then $u_h(e)$ evaluated at the optimal effort level will be negative, and hence $v(\gamma) < 0$. Thus, not hiring the agent and not buying dominates hiring him and buying only if the signal is high. Hence, there is a left neighborhood of $\hat{\gamma}$ such that the principal prefers not to hire the agent. If $\hat{\gamma} y_h + (1-\hat{\gamma}) y_\ell > 0$, then the principal's profits at $\gamma = \hat{\gamma}$ if she buys after a high signal are $\alpha(\hat{\gamma} y_h + (1-\hat{\gamma}) y_\ell) > 0$. But as $\alpha(\hat{\gamma} y_h + (1-\hat{\gamma}) y_\ell) < \hat{\gamma} y_h + (1-\hat{\gamma}) y_\ell$, she strictly prefers to buy without hiring the agent. Proceeding as when $\hat{\gamma} y_h + (1-\hat{\gamma}) y_\ell < 0$, one can show that if ε is small enough then $v(\gamma) < \gamma y_h + (1-\gamma) y_\ell$. Thus, there is a left neighborhood of $\hat{\gamma}$ where the principal strictly prefers to buy without hiring the agent.

(*iii*)-(*d*) Suppose $\hat{\gamma} \neq \tilde{\gamma}$. At $\gamma = \tilde{\gamma}$ the principal can obtain zero expected profits if she buys only after the high signal realization and implements $e = 0$, for it yields $\alpha(\gamma y_h + (1-\gamma) y_\ell) = 0$. Also, by Lemma 6, the principal implements a positive level of effort for all $\gamma \neq \hat{\gamma}$ when she buys only after the high signal realization. By revealed preference, she obtains positive expected profits by doing so. Thus, the optimal effort level at $\gamma = \tilde{\gamma}$ is positive. Since effort is continuously differentiable in γ at $\tilde{\gamma}$, it follows that it is positive for values of γ in a neighborhood of $\tilde{\gamma}$. □

A.3. *Dynamic first best: proof of Lemma 2*

We only sketch the proof since it is very similar to its static counterpart. The proofs of (*i*)–(*iii*) are analogous to those of Lemma 1. The proof of the effort comparison in (*iv*) follows from the first-order condition in the text, as $(\gamma \beta_h - (1-\gamma)\beta_\ell)\delta S^*$ is positive, negative, or zero, respectively, if γ is greater, less, or equal to $\hat{\gamma}$. Finally, the proof of the threshold comparison in (*iv*) is analogous to those of parts (*iii*)-(*a*) and (*iii*)-(*b*) of Proposition 1. □

A.4. *Optimal dynamic contract: proof of Proposition 2*

Assuming that the principal can use lotteries, her problem can be rewritten as

$$
\begin{aligned}
U(V) = \max_{\{e_j, w_{i1j}, w_{i0j}, V_{ij}, \sigma_{ij},, V_j, s_j\}_{i \in \{h,\ell\}, j \in \{1,2\}}} & \sum_{j=1}^{2} s_j \{ \pi(e_j)[\sigma_{hj}(E_h(e_j) - w_{h1j}) \\
& + (1-\sigma_{hj})(-w_{h0j} + \delta U(V_{hj}))] + (1-\pi(e_j))[\sigma_{\ell j}(E_\ell(e_j) - w_{\ell 1 j}) \\
& + (1-\sigma_{\ell j})(-w_{\ell 0 j} + \delta U(V_{\ell j}))]\}
\end{aligned}
$$

$$
\begin{aligned}
\text{s.t. } & \sum_{j=1}^{2} s_j V_j = V, \\
& s_j \{-\psi(e_j) + \pi(e_j)[\sigma_{hj} w_{h1j} + (1-\sigma_{hj})(w_{h0j} + \delta V_{hj})] \\
& \qquad + (1-\pi(e_j))[\sigma_{\ell j} w_{\ell 1 j} + (1-\sigma_{\ell j})(w_{\ell 0 j} + \delta V_{\ell j})]\} = V_j \text{ for } j \in \{1,2\}, \\
& \psi'(e_j) = \pi'(e_j)[\sigma_{hj} w_{h1j} + (1-\sigma_{hj})(w_{h0j} + \delta V_{jh}) - \sigma_{\ell j} w_{\ell 1 j} \\
& \qquad \quad - (1-\sigma_{\ell j})(w_{\ell 0 j} + \delta V_{\ell j})] \quad \text{for } j \in \{1,2\}, \\
& e_j \geq 0,\ w_{i1j} \geq 0,\ w_{i0j} \geq 0,\ V_{ij} \geq 0,\ \sigma_{ij} \in \Sigma \text{ for } i \in \{h,\ell\}, j \in \{1,2\}, \\
& s_j \in [0,1], s_1 + s_2 = 1.
\end{aligned}
$$

Due to the possibility of using lotteries, the value function is concave.

We will proceed by deriving properties of the optimal contract for the case in which the principal does not use lotteries over contracts (payments), i.e., when $V_1 = V_2 = V$. In particular, we will take first-order conditions with respect to $e \equiv e_1 = e_2$, $w_{ik} \equiv w_{ik1} = w_{ik2}$, $V_i \equiv V_{i1} = V_{i2}$, $\sigma_i \equiv \sigma_{i1} = \sigma_{i2}$, $i \in \{h, \ell\}$, $k \in \{0, 1\}$. At promised values V for which lotteries are optimal (i.e., $V_1 < V < V_2$ wlog, and $s_j \in (0, 1)$), for each realization $j \in \{1, 2\}$ the corresponding contract will have the derived properties given that the promised value to the agent is equal to V_j.

Let λ, μ, $\sigma_i \xi_{ik}$, and $(1-\sigma_i)\delta\nu_i$ be the Lagrange multipliers on constraints (7), (8), $w_{ik} \geq 0$, and $V_i \geq 0$, $i \in \{h, \ell\}$, $k \in \{0, 1\}$, respectively. Denote by $U'_+(V)$, $U'_-(V)$, and $\partial U(V)$ the right derivative, left derivative, and superdifferential, respectively, of the concave function U at V. The first-order conditions with respect to w_{h1}, w_{h0}, $w_{\ell 1}$, and $w_{\ell 0}$ are:

$$\sigma_h[(\lambda-1)\pi(e) + \mu\pi'(e) + \xi_{h1}] = 0, \tag{A.4}$$

$$(1-\sigma_h)[(\lambda-1)\pi(e) + \mu\pi'(e) + \xi_{h0}] = 0, \tag{A.5}$$

$$\sigma_\ell[(\lambda-1)(1-\pi(e)) - \mu\pi'(e) + \xi_{\ell 1}] = 0, \tag{A.6}$$

$$(1-\sigma_\ell)[(\lambda-1)(1-\pi(e)) - \mu\pi'(e) + \xi_{\ell 0}] = 0, \tag{A.7}$$

and, after some straightforward substitution, the first-order condition with respect to e is

$$-\psi'(e) + \sigma_h u_h'(e) + \sigma_\ell u_\ell'(e) + \pi'(e)\delta[(1-\sigma_h)(U(V_h)+V_h)) - (1-\sigma_\ell)(U(V_\ell)+V_\ell))] + \mu\left(\pi''(e)\frac{\psi'(e)}{\pi'(e)} - \psi''(e)\right) \leq 0, \quad e \geq 0, \tag{A.8}$$

all with complementary slackness. Using standard rules of superdifferential calculus,[30] we obtain the following first-order conditions with respect to V_h and V_ℓ:

$$0 \in (1-\sigma_h)[\pi(e)(\partial U(V_h) + \lambda) + \mu\pi'(e) + \nu_h], \tag{A.9}$$

$$0 \in (1-\sigma_\ell)[(1-\pi(e))(\partial U(V_\ell) + \lambda) - \mu\pi'(e) + \nu_\ell]. \tag{A.10}$$

The envelope condition is[31]

$$-\lambda \in \partial U(V). \tag{A.11}$$

Finally, for completeness we will also derive the first-order condition with respect to σ_i, $i \in \{h, \ell\}$, which in the case when σ_i is interior, can be written as

$$\pi(e)[E_i - \delta U(V_i)] + \left[(\lambda-1)\pi(e) + \mu\pi'(e) + \xi_{i1}\right] w_{i1} - \left[\lambda\pi(e) + \mu\pi'(e) + \nu_i\right]\delta V_i = 0.$$

Using (A.4), the second term is zero. Hence the above equation becomes $\pi(e)[E_i - \delta U(V_i)] - \left[\lambda\pi(e) + \mu\pi'(e) + \nu_i\right]\delta V_i = 0$. Using the first-order condition with respect to V_i, this can further be rewritten as $0 \in E_i - \delta U(V_i) + \delta\partial U(V_i)V_i$. Alternatively, in terms of the social surplus $S(V) = U(V) + V$, we have $\partial U(V_i) = \partial S(V_i) - 1$ and the first-order condition is

$$0 \in E_i - \delta S(V_i) + \delta\partial S(V_i)V_i. \tag{A.12}$$

As we have already argued in the main text, the slope of the value function U is weakly greater than minus one. Formally, in terms of superdifferentials, we have:

30 See, for instance, Borwein and Lewis (2006, Chapter 6, p. 134).

31 See, for instance, Aubin (1979, Chapter 5, pp. 130–140).

Lemma 7. *For all $V \geq 0$ and all $x \in \partial U(V)$, $x \geq -1$. That is, the slope of the value function is always greater than minus one.*

Proof. Summing up (A.4)–(A.7), obtain $\lambda - 1 \leq 0$. If U is differentiable at V so that $\partial U(V) = \{U'(V)\}$, then from (A.11), $U'(V) = -\lambda \geq -1$. Since U is concave, it is differentiable everywhere but on a countable set. It follows that $U'_+(V) \geq -1$ for all values of V (which in turns implies the statement of the lemma since $\partial U(V) = [U'_+(V), U'_-(V)]$ for any $V > 0$, and $\partial U(0) = \{U'_+(0)\}$ if $U'_+(0)$ is finite). To see this, suppose that $U'_+(V') < -1$ for some V'. Take $V'' > V'$ such that $U'(V'')$ exists. By the previous argument, $U'_+(V'') = U'(V'') \geq -1$, which contradicts concavity of U, since a concave function has a decreasing right derivative. □

As explained in the main text, Lemma 7 allows us to set $w_{h0} = w_{\ell 0} = 0$, and hence we denote $w_h \equiv w_{h1}$ and $w_\ell \equiv w_{\ell 1}$.

We now turn to the proof of the proposition:

Proof of Proposition 2. (i)-(a) The incentive constraint (8) implies that $\sigma_h w_h + (1-\sigma_h)\delta V_h \geq \sigma_\ell w_\ell + (1-\sigma_\ell)\delta V_\ell$, with strict inequality if $e > 0$.

Consider the case when $e > 0$. Suppose first that $\sigma_h < 1$, and that $V_h \leq V$. Then from (A.4), (A.9), Lemma 7, and the fact that $V_h \leq V < V^*$, it follows that $w_h = 0$. Then $V > (1-\sigma_h)\delta V_h > \sigma_\ell w_\ell + (1-\sigma_\ell)\delta V_\ell$ and thus $V = -\psi(e) + \pi(e)(1-\sigma_h)\delta V_h + (1-\pi(e))[\sigma_\ell w_\ell + (1-\sigma_\ell)\delta V_\ell] < V$, a contradiction. Hence, $V_h > V$.

Notice that $\mu \geq 0$. Indeed, $V_h > V$ implies $V_h > 0$ so that $\nu_h = 0$, and also $U'_-(V_h) \leq U'_+(V)$. Suppose that $\mu < 0$. Then from (A.9) and (A.11), $-\lambda \in [\partial U(V_h) + \mu\pi'(e)/\pi(e)] \cap \partial U(V) = \emptyset$ since $\pi'(e) > 0$, a contradiction. Furthermore, $\mu \geq 0$ also holds when $\sigma_h = 1$ (and $e > 0$). Indeed, in this case $w_h > \sigma_\ell w_\ell + (1-\sigma_\ell)\delta V_\ell \geq 0$. Then (A.4), $\xi_{h1} = 0$, $\lambda \leq 1$ and $\pi' > 0$ imply $\mu \geq 0$. Moreover, the inequality is strict if $V < V^*$ so that $\lambda < 1$.

Now suppose that $\sigma_\ell < 1$ (and $e > 0$). We want to show that in this case $V_\ell < V$. Suppose that $V_\ell \geq V$. If $\mu > 0$, then using (A.10), (A.11), and $\nu_\ell = 0$ (since $V_\ell \geq V > 0$), we have a contradiction. As we have shown above, if $\sigma_h = 1$ and $V < V^*$, then $\mu > 0$ and hence $V_\ell < V$ follows. The only potentially problematic case is when $\sigma_h < 1$ and $\mu = 0$ (even though $V < V^*$). We are going to rule out this possibility in what follows.

First notice that combining (A.9)–(A.11) we have $0 \in \pi(e)\partial U(V_h) + (1-\pi(e))\partial U(V_\ell) - \partial U(V)$, where we used that $\nu_h = 0$ since $V_h > V > 0$, and $\nu_\ell = 0$ since $V_\ell > V$ by the premise. Since U is concave, if both V_h and V_ℓ exceed V, then the slope of U must the same at V, V_h, and V_ℓ. In other words, the functions $U(V)$ and $S(V) = U(V) + V$ must be linear on $[V, \max\{V_\ell, V_h\}]$. Let x denote the derivative of S on this interval. Suppose first that both σ_h and σ_ℓ are interior. Using the first-order condition (A.12), we have $E_i(e) - \delta S(V_i) + \delta x V_i = 0$, $i \in \{h, \ell\}$. Differentiating the last two terms with respect to V, we have $-x + x = 0$, implying $E_h(e) = E_\ell(e)$ and hence the implemented effort must be zero. This contradicts the premise $e > 0$. Notice that this argument also goes through if σ_ℓ is interior, but $\sigma_h = 0$ as then $E_h(e) - \delta S(V_h) + \delta x V_h \leq E_\ell(e) - \delta S(V_\ell) + \delta x V_\ell = 0$, so that $E_h(e) \leq E_\ell(e)$, contradicting $e > 0$.

The only remaining case that we need to consider is $\sigma_h = \sigma_\ell = 0$. We provide a detailed treatment of the principal's problem in this case (that we refer to as U_{00}) in Appendix A.6. When the slope of U is the same at V_h and V_ℓ, the first-order condition with respect to effort, the left-hand side of which is given by (A.20), is $\delta\pi'(e)[U(V_h) - U(V_\ell)] = 0$ implying $U(V_h) = U(V_\ell)$. If $V_h = V_\ell$, then using the incentive constraint we reach a contradiction with $e > 0$. The only way in which V_h and V_ℓ are not equal to each other and yet $U(V_h) = U(V_\ell)$ is if U has a zero slope on an interval containing V_h and V_ℓ.

Next we will show that if $\sigma_h = \sigma_\ell = 0$, the function U cannot have a zero slope on an interval. Take $\hat{V} \equiv \sup\{V|U'(V) = 0\}$. Rewriting equation (A.19) from Appendix A.6 evaluated at $V = \hat{V}$, we have

$$\pi\delta U(V_h) + (1-\pi)\delta U(V_l) \le \delta U(\pi V_h + (1-\pi)V_l) = \delta U\left(\frac{\hat{V}+\psi}{\delta}\right) < \delta U\left(\frac{\hat{V}}{\delta}\right),$$

implying that $e = 0$ and $V_h = V_\ell = V/\delta$ are optimal. Then $U(\hat{V}) = \delta U(\hat{V}/\delta)$ and $0 = U'(\hat{V}) = U'(\hat{V}/\delta)$, where the first equality follows from the definition of $\hat{V}$. (If U is non-differentiable at those points, we instead have $0 \in \partial U(\hat{V})$ and $0 \in \partial U(\hat{V}/\delta)$.) But since $\hat{V}/\delta > \hat{V}$, the slope of U at $\hat{V}/\delta$ cannot be zero, as $\hat{V}$ is the largest point with that property, a contradiction. This completes the proof of $V_\ell < V$ when $e > 0$ and $\sigma_\ell < 1$.

Finally, the fact that setting $w_h = 0$ dominates $w_h > 0$ (strictly so if $V_h < V^*$) whenever $\sigma_h < 1$ follows from (A.4), (A.9) and Lemma 7, and similarly for w_ℓ.

(i)-(b) The proof is similar to that of part (i)-(a). The incentive constraint (8) implies that $\sigma_h w_h + (1-\sigma_h)\delta V_h \le \sigma_\ell w_\ell + (1-\sigma_\ell)\delta V_\ell$, with strict inequality if $e > 0$.

Consider the case when $e > 0$. Suppose first that $\sigma_\ell < 1$, and suppose that $V_\ell \le V$. Then from (A.6), (A.10), Lemma 7, and the fact that $V_\ell \le V < V^*$, it follows that $w_\ell = 0$. Then $V > (1-\sigma_\ell)\delta V_\ell > \sigma_h w_h + (1-\sigma_h)\delta V_h$ and thus $V = -\psi(e) + \pi(e)[\sigma_h w_h + (1-\sigma_h)\delta V_h] + (1-\pi(e))(1-\sigma_\ell)\delta V_\ell < V$, a contradiction. Thus $V_\ell > V$.

Notice that $\mu \ge 0$. Indeed, $V_\ell > V$ implies $V_\ell > 0$ so that $\nu_\ell = 0$, and also $U'_-(V_\ell) \le U'_+(V)$. Suppose that $\mu < 0$. Then from (A.10) and (A.11), $-\lambda \in [\partial U(V_\ell) - \mu\pi'(e)/(1-\pi(e))] \cap \partial U(V) = \emptyset$ since $\pi'(e) < 0$, a contradiction. Furthermore, $\mu \ge 0$ also holds when $\sigma_\ell = 1$ (and $e > 0$). Indeed, in this case $w_\ell > \sigma_h w_h + (1-\sigma_h)\delta V_h \ge 0$. Then (A.6), $\xi_{\ell 1} = 0$, $\lambda \le 1$ and $\pi' < 0$ imply $\mu \ge 0$. Moreover, the inequality is strict if $V < V^*$ so that $\lambda < 1$. The proof of $V_h < V$ is analogous to the proof of $V_\ell < V$ in (i)-(a) part, and is omitted.

Since $\lambda < 1$, $\pi' < 0$, and $\mu \ge 0$, (A.4) implies that $\xi_{h1} > 0$ and thus $w_h = 0$ is optimal. (If $\sigma_h = 0$, then the choice of w_h is irrelevant.)

(ii) From (7) and (8), we can express $\sigma_h w_h + (1-\sigma_h)\delta V_h = V + \psi(e) + (1-\pi(e))\psi'(e)/\pi'(e)$ and $\sigma_\ell w_\ell + (1-\sigma_\ell)\delta V_\ell = V + \psi(e) - \pi(e)\psi'(e)/\pi'(e)$. Non-negativity constraints $w_h \ge 0$ and $V_h \ge 0$ imply $\sigma_h w_h + (1-\sigma_h)\delta V_h = V + \psi(e) + (1-\pi(e))\psi'(e)/\pi'(e) \ge 0$. Similarly, $w_\ell \ge 0$ and $V_\ell \ge 0$ imply $V + \psi(e) - \pi(e)\psi'(e)/\pi'(e) \ge 0$.

Suppose that $\gamma > \hat{\gamma}$. If $V = 0$, then the only value of e that satisfies $V + [\psi(e) - \pi(e)\psi'(e)/\pi'(e)] \ge 0$ is $e = 0$. To see this, notice that at $e = 0$, the term in the square brackets vanishes, and its derivative with respect to e is negative (so this term is negative for all $e > 0$). Similarly, when $\gamma < \hat{\gamma}$ and $V = 0$, the only value of e that satisfies $V + [\psi(e) + (1-\pi(e))\psi'(e)/\pi'(e)] \ge 0$ is $e = 0$. Hence, if $V = 0$ then $e = 0$ and thus $\sigma_h w_h + (1-\sigma_h)\delta V_h = \sigma_\ell w_\ell + (1-\sigma_\ell)\delta V_\ell = 0$.

(iii) Let $V \ge V^* > 0$. Assume first that $V > V^*$. Then $U'(V)$ exists and equals -1. As we have shown above, $\mu \ge 0$. From (A.11), $\lambda = 1$. Let $\gamma > \hat{\gamma}$. Summing up (A.4) and (A.5) yields $\mu\pi'(e) \le 0$. Thus $\pi' > 0$ and $\mu \ge 0$ imply $\mu = 0$. Similarly, suppose that $\gamma < \hat{\gamma}$. Summing up (A.6) and (A.7) yields $-\mu\pi'(e) \le 0$, so that $\pi' < 0$ and $\mu \ge 0$ imply $\mu = 0$. Also, (A.4)–(A.7) holding at equalities imply that the limited-liability constraints do not bind. Furthermore, from (A.9) and (A.10) with $\lambda = -1$ and $\mu = 0$, $V_h \ge V^* > 0$ and $V_\ell \ge V^* > 0$ are optimal choices, implying that $\nu_h = \nu_\ell = 0$. Therefore the Lagrange multipliers on constraints (8) and (9) are all zero, and the problem of the principal is effectively unconstrained (it is only subject to the promise-keeping constraint). Moreover, since $V_h \geq V^*$ and $V_\ell \geq V^*$, the principal solves the unconstrained problem in all subsequent periods as well. Thus the first best is achieved.

Assume now that $V = V^*$. At that point $U'(V^*)$ need not exist and λ is not necessarily equal to -1. But the argument above shows that there exists a solution to the first-order conditions that achieves the first best. Since for any V profits are higher in the first-best case, the constructed solution is optimal. □

A.5. The $\Sigma = \{0, 1\}$ case: proof of Claim 1

$U_{01}(V)$ is the value of the following problem:

$$U_{01}(V) = \max_{w_{l1}, V_h, e} \pi(e)\delta U(V_h) + u_l(e) - (1 - \pi(e))w_{l1} \tag{A.13}$$
$$\text{s.t.} \quad -\psi(e) + \pi(e)\delta V_h + (1 - \pi(e))w_{l1} = V,$$
$$\psi'(e) = \pi'(e)\left(\delta V_h - w_{l1}\right),$$
$$V_h \geq 0, \quad w_{l1} \geq 0, \quad e \geq 0.$$

As usual, we denote by λ and μ the multipliers associated with the promise-keeping and the incentive constraints, respectively. We will show that U_{01} is dominated in the principal's problem, and thus can be omitted without loss of generality. We proceed in two steps.

Step 1: Optimal effort is zero in problem (A.13). Consider the first-order condition with respect to effort in problem (A.13):

$$\pi'(e)\delta U(V_h) + u_l'(e) + \pi'(e)w_{l1} + \mu\left(\pi''(e)\left(\delta V_h - w_{l1}\right) - \psi''(e)\right) \leq 0, \tag{A.14}$$

$e \geq 0$, with complementary slackness. Adding and subtracting $\pi'(e)\delta V_h$, using the incentive constraint to replace $\delta V_h - w_{l1}$ and the expression for the derivative $(\psi'/\pi')'$ yield

$$\pi'(e)\delta\left(U(V_h) + V_h\right) + u_l'(e) - \psi'(e) - \mu\pi'(e)\left(\frac{\psi'}{\pi'}\right)'(e) \leq 0, \tag{A.15}$$

$e \geq 0$, with complementary slackness. If $\pi' < 0$, we immediately obtain that optimal effort is zero since all the terms are nonpositive. Suppose that $\pi' > 0$ (i.e., $\gamma\beta_h - (1 - \gamma)\beta_l > 0$). Since the last two terms are nonpositive, to prove that optimal effort is zero it suffices that

$$\pi'(e)\delta\left(U(V_h) + V_h\right) + u_l'(e) \leq 0,$$

which can be written as

$$\eta'(e)\left[(\gamma\beta_h - (1 - \gamma)\beta_l)\delta\left(U(V_h) + V_h\right) - (\gamma\beta_h y_h - (1 - \gamma)\beta_l y_l)\right] \leq 0.$$

Let S^* be the first-best surplus. Since $U(V_h) + V_h \leq S^*$, it suffices to show that

$$\gamma\beta_h y_h - (1 - \gamma)\beta_l y_l \geq ((\gamma\beta_h - (1 - \gamma)\beta_l)\,\delta S^*. \tag{A.16}$$

Recall that the optimal strategy in the first-best case is stationary: the principal always chooses the same level of effort. Thus, the following is an upper bound for S^*:

$$S^* \leq \max\left\{0, \gamma y_h + (1 - \gamma)y_l, \frac{u_h(e^{**})}{\pi(e^{**})}\right\},$$

where e^* is the optimal level of effort in the first-best case if the optimal constant policy happens to be buy if the signal realization is high and do not buy if it is low.

To see that this is an upper bound for S^*, suppose the optimal first-best policy is to never buy. Then the payoff is 0. If instead it is to always buy, then the payoff is $\gamma y_h + (1-\gamma)y_l$. If it is to buy if high and not buy if low, then $S^* = (u_h - \psi)/(1-(1-\pi)\delta) < (u_h - \psi)/\pi < u_h/\pi$.

Notice that if we replace S^* in (A.16) by either 0 or $\gamma y_h + (1-\gamma)y_l$, it is clear that the inequality in (A.16) holds. It remains to show that if we replace S^* by $u_h(e^{**})/\pi(e^{**})$ then (A.16) holds as well. We will show something stronger by setting $\delta = 1$ in (A.16), which (after rearranging terms) becomes

$$\pi(e^{**})(\gamma\beta_h y_h - (1-\gamma)\beta_l y_l) - ((\gamma\beta_h - (1-\gamma)\beta_l)\, u_h(e^{**}) \geq 0. \tag{A.17}$$

The first term is $(\gamma(\alpha+\beta_h\eta(e)) + (1-\gamma)(\alpha-\beta_l\eta(e))) \times (\gamma\beta_h y_h - (1-\gamma)\beta_l y_l)$, while the second term equals $(\gamma\beta_h - (1-\gamma)\beta_l) \times (\gamma(\alpha+\beta_h\eta(e))y_h + (1-\gamma)(\alpha-\beta_l\eta(e))y_l)$. Substituting in (A.17) yields, after some algebra,

$$\gamma(1-\gamma)(y_h - y_l)(\beta_h(\alpha - \beta_l\eta(e^*)) + \beta_l(\alpha + \beta_h\eta(e^{**}))) > 0.$$

Thus, optimal effort is zero also in the $\pi' > 0$ case, proving that the optimal effort in problem (A.13) is zero.

Step 2: U_{01} is dominated by an alternative policy. Since the optimal effort is zero, the expected profit under U_{01} is $U_{01}(V) = (1-\pi(0))(E_0 - V) + \pi(0)\delta U(V/\delta)$. If $E_0 - V < \delta U(V/\delta)$, then $U_{00}(V)(\geq \delta U(V/\delta)) > U_{01}(V)$. If $E_0 - V > \delta U(V/\delta)$, then $U_{11}(V) > U_{01}(V)$ (and if $E_0 - V = \delta U(V/\delta)$ then $U_{11}(V) = U_{01}(V)$). Thus U_{01} is dominated, strictly so except for the knife-edge case $E_0 - V = \delta U(V/\delta)$, and thus can be deleted from the problem. □

A.6. The $\Sigma = \{0, 1\}$ case: proof of Claim 2

$U_{00}(V)$ is the value of the following problem:

$$\begin{aligned} U_{00}(V) = &\max_{V_l, V_h, e} \pi(e)\delta U(V_h) + (1-\pi(e))\delta U(V_l) \\ \text{s.t.} \; & -\psi(e) + \pi(e)\delta V_h + (1-\pi(e))\delta V_l = V, \\ & \psi'(e) = \pi'(e)\delta(V_h - V_l), \\ & V_h \geq 0, \quad V_l \geq 0, \quad e \geq 0. \end{aligned}$$

Let $V > 0$, and notice that the objective function in the U_{00} problem can be written as follows after solving the promise-keeping and incentive constraints for V_h and V_l:

$$\pi(e)\delta U\left(\frac{V + \psi(e) + (1-\pi(e))\frac{\psi'(e)}{\pi'(e)}}{\delta}\right) + (1-\pi(e))\delta U\left(\frac{V + \psi(e) - \pi(e)\frac{\psi'(e)}{\pi'(e)}}{\delta}\right). \tag{A.18}$$

If U is decreasing in V, then the following argument shows that $e = 0$ is optimal:

$$\pi\delta U(V_h) + (1-\pi)\delta U(V_l) \leq \delta U(\pi V_h + (1-\pi)V_l) = \delta U\left(\frac{V+\psi}{\delta}\right) < \delta U\left(\frac{V}{\delta}\right), \tag{A.19}$$

where the first inequality follows from concavity, the equality from the functional forms of V_h and V_l, and the last inequality from the premise that U is decreasing in V. Hence, choosing $e = 0$ dominates choosing a positive effort level. This argument also applies to the value of V for which

U has zero slope. As we have shown in the proof of part (i)-(a) of Proposition 2, U cannot have zero slope on an interval if $U = U_{00}$.

Assume now that $U'(V) > 0$. The derivative of (A.18) with respect to e is

$$\delta\pi'\Delta U + \pi(1-\pi)\left(\frac{\psi'}{\pi'}\right)'\Delta U', \tag{A.20}$$

where $\Delta U = U(V_h) - U(V_l)$ and $\Delta U' = U'(V_h) - U'(V_l)$. Notice that the derivative with respect to e of the objective function is zero at $e = 0$. We are going to show that it is positive near $e = 0$, and hence the optimal effort must be positive for such V.

When $e > 0$, we can rewrite the above expression as

$$\Delta U\left[\delta\pi' + \pi(1-\pi)\frac{\left(\frac{\psi'}{\pi'}\right)'\Delta U'}{\Delta U}\right]. \tag{A.21}$$

Consider first $\pi' > 0$. Then $\Delta U > 0$ since $V_h > V_l$ and $U'(V) > 0$. Thus, it suffices to show that the term in square brackets converges to a positive value as e goes to zero. For by continuity it will be positive for small values of e, and the whole expression will be positive.

The first term converges to $\delta\pi'(0) > 0$, while the second term converges to

$$\pi(0)(1-\pi(0))\lim_{e\to 0^+}\frac{\left(\frac{\psi'}{\pi'}\right)'\Delta U'}{\Delta U}.$$

Since the limit in this expression is $0/0$, we use L'Hopital's rule to obtain

$$\begin{aligned}\lim_{e\to 0^+}\frac{\left(\frac{\psi'}{\pi'}\right)'\Delta U'}{\Delta U} &= \lim_{e\to 0^+}\left(\frac{\psi'}{\pi'}\right)'\lim_{e\to 0^+}\frac{\Delta U'}{\Delta U}\\ &= \lim_{e\to 0^+}\left(\frac{\psi'}{\pi'}\right)'\lim_{e\to 0^+}\frac{(1-\pi)U''(V_h)+\pi U''(V_l)}{(1-\pi)U'(V_h)+\pi U'(V_l)}\\ &= 0,\end{aligned}$$

since the first term is zero while the second is $U''(V)/U'(V) \le 0$. Hence, the term in the square brackets in (A.21) is positive at $e = 0$ and by continuity for $e > 0$ small enough. Since the derivative of the objective function is zero at $e = 0$ but positive for any e in a right neighborhood of $e = 0$, the optimal level of effort must be positive.

Notice that the same goes through if instead $\pi' < 0$, for in this case $\Delta U < 0$ for e small as $V_l > V_h$. This completes the proof of the claim. □

A.7. The $\Sigma = \{0, 1\}$ case: additional results

Claim 4. *Suppose the function U is strictly increasing for some V. Let $\hat{V}$ be the largest value V for which U has zero slope. Then $U(V) = U_{10}(V)$ for $V \ge \hat{V}$.*

Proof. Let $\hat{V} = \sup\{V|U'(V) = 0\}$. Since $U_{11} = \gamma y_h + (1-\gamma)y_\ell - V$, U cannot be equal to U_{11} at $\hat{V}$. Suppose that $U(\hat{V}) = U_{00}(\hat{V})$. Then following the same argument as in the proof of part (i)-(a) of Proposition 2, we conclude U must have zero slope at $\hat{V}/\delta > \hat{V}$, a contradiction with the definition of $\hat{V}$. Hence $U(\hat{V}) = U_{10}(\hat{V})$.

Next we will show that $U(V) = U_{10}(V)$ for all $V \geq \hat{V}$. We need to show that neither U_{11} nor U_{00} can intersect the U_{10} to the right of $\hat{V}$. It is clear that such an intersection cannot occur with U_{11} because $U_{11} = \gamma y_h + (1-\gamma)y_\ell - V$ and the function U has an upward sloping part. (It is easy to see graphically that for the U to have an upward sloping part, an intersection with U_{11} can only occur to the left of $\hat{V}$.) We thus only need to rule out an intersection of U_{00} and U_{10}. From part (*iii*) of Proposition 2 we know that $U(V) = U_{10}(V)$ for $V \geq V^*$ whenever $V^* > 0$. Therefore if an intersection occurs, U_{00} has to intersect U_{10} on $(\hat{V}, V^*)$ at least twice. To rule this out, it is enough to show that U_{00} cannot intersect U_{10} from below to the right of $\hat{V}$. Towards a contradiction, suppose that at V', U_{00} crosses U_{10} from below. Such a crossing would require that at V' the slope of U_{00} be bigger than the slope of the upper envelope, U.[32] (And if the functions are not differentiable at V', the same holds for their right derivatives.) That is, $U'_{00+}(V') > U'_+(V)$. However, this cannot occur for the following reason. As shown in the proof of Claim 2, when U is decreasing in V, the optimal effort in the U_{00} problem is zero, and $V_h = V_\ell = V/\delta$. Thus $U'_{00+}(V) = U'_+(V/\delta) \leq U'_+(V)$ for all $V > \hat{V}$, where the inequality follows from concavity of U. A contradiction. □

A.8. *Use of lotteries: proof of Lemma 3*

We will ignore lotteries over contracts since the argument goes through the same way with them. Suppose that $e \geq \tilde{e}$, which implies that $E_h(e) \geq \delta S^* > \delta S(V_h)$ for all $V_h < V^*$. We will show that $\sigma_h = 1$ at the optimum. Towards a contradiction, consider a solution with $\sigma_h \in [0, 1)$, which implies that $w_h = 0$. Then an increase in σ_h by $d\sigma_h$ and w_h by dw_h such $dw_h = (\delta V_h/\sigma_h)d\sigma_h$, leaves the constraint set unaltered and increases the principal's payoff by $(E_h(e) - \delta S(V_h))d\sigma_h > 0$, contradiction. Hence, $\sigma_h = 1$ at the optimum.

Suppose that $e \in (0, \tilde{e})$, so that $E_h(e) < \delta S^*$. We will show that $\sigma_h = 1$ cannot be optimal. Towards a contradiction, consider a solution with $\sigma_h = 1$ and thus with $w_h > 0$. Alter the contract by setting $V_h = V^*$ and $\sigma_h < 1$ such that $w_h = (1-\sigma_h)V^*$ to keep the constraint set unaffected. The principal's expected profit increases by $(1-\sigma_h)(\delta S^* - E_h(e))$, contradiction. Hence, $\sigma_h = 1$ is not optimal if $e \in (0, \tilde{e})$, completing the proof of the first part of the lemma.

To prove the second assertion, assume that $e \in (0, \tilde{e})$ and, towards a contradiction, that σ_h arbitrarily close to one is optimal. Since $\gamma > \hat{\gamma}$, by part (*i*)-(*a*) of Proposition 2 $\sigma_h w_h + (1 - \sigma_h)V_h > 0$, and since $w_h = 0$ is optimal, it follows that $V_h > 0$. If σ_h is arbitrarily close to one, then V_h is arbitrarily large and thus $V_h \geq V^*$. As a result, $S(V_h) = S^*$ and $S'(V_h) = 0$. Then the first-order condition (A.12) for $i = h$ (which must hold since σ_h is interior by the premise) becomes $E_h(e) - \delta S^* = 0$, contradicting $e < \tilde{e}$. □

A.9. *Use of lotteries: proof of Proposition 3*

We will prove the proposition through a series of lemmata.[33] We will proceed by setting $\sigma_\ell = 0$, and at the end of the proof we will consider what happens if $\sigma_\ell > 0$. From the premises,

[32] Indeed, the intersection between U_{00} and U_{10} creates a non-convexity; around the point of the intersection, lotteries over contracts would be used, and U will be a convex hull of U_{00} and U_{10} (to the right of $\hat{V}$). It is easy to see graphically that the curve that intersects from below (U_{00}) cannot have a larger slope than the convex hull.

[33] To simplify an already long proof, throughout we are going to proceed as if U was continuously differentiable and use first-order conditions and derivatives. At the cost of more notation, one could do it with superdifferentials as in Proposition 2. Also, one can replace in each instance the continuity of the derivative by the continuity of the right derivative, which holds for a concave function such as U.

notice that $\pi'(e) > 0$ for all e (since $\gamma > \hat{\gamma}$) and $E_h(0) > 0$ (since $\gamma > \tilde{\gamma}$). Moreover, $E_h(0) > \delta S(0)$. To see this, notice that if $V = 0$, then only $e = 0$ is feasible. Hence,

$$S(0) = \pi(0)[\sigma_h E_h(0) + (1-\sigma_h)\delta S(0)] + (1-\pi(0))\delta S(0),$$

which rearranges to

$$S(0) = \frac{\pi(0)\sigma_h E_h(0)}{1-\delta(\pi(0)(1-\sigma_h) + (1-\pi(0)))}.$$

It is easy now to verify that $E_h(0) > \delta S(0)$ if $E_h(0) > 0$.

Step 1: Candidate for an interior σ_h. Consider the following problem, which is obtained by setting $\sigma_\ell = 0$ and solving the promise-keeping and incentive constraints for V_h and V_ℓ (given that $w_h = 0$ when $\sigma_h < 1$):

$$\begin{aligned}\max_{e,\sigma_h} -\psi + \pi &\left(\sigma_h E_h + (1-\sigma_h)\delta S\left(\frac{V + \psi + (1-\pi)\frac{\psi'}{\pi'}}{\delta(1-\sigma_h)}\right)\right) \\ &+ (1-\pi)\delta S\left(\frac{V + \psi - \pi\frac{\psi'}{\pi'}}{\delta}\right) \\ \text{s.t.}\quad & V + \psi - \pi\frac{\psi'}{\pi'} \geq 0, \quad e \geq 0, \quad \sigma_h \in [0,1).\end{aligned}$$

The first constraint is the non-negativity constraint $V_\ell \geq 0$, which is the relevant one given that $\pi' > 0$. At the optimum, either $V + \psi - \pi\psi'/\pi' = 0$ (the constraint binds) or $V + \psi - \pi\psi'/\pi' > 0$. We will consider each case separately.

Lemma 8. *If V is sufficiently small, then there is a unique point $(e_0(V), \sigma_h(V))$, with $e_0(V) > 0$ and $\sigma_h(V) \in (0,1)$, that solves the first-order conditions with $V + \psi - \pi\frac{\psi'}{\pi'} = 0$.*

Proof. Since the constraint binds, $e_0(V)$ is pinned down uniquely by it. Notice that (i) it increases in V, and (ii) it is differentiable with derivative $e_0'(V) = 1/[\pi(\psi'/\pi')']$. Using $V + \psi - \pi\frac{\psi'}{\pi'} = 0$, the first-order condition for an interior σ_h given by (11) becomes

$$\begin{aligned}E_h(e_0(V)) = \delta S&\left(\frac{\psi'(e_0(V))/\pi'(e_0(V))}{\delta(1-\sigma_h)}\right) \\ &- \delta S'\left(\frac{\psi'(e_0(V))/\pi'(e_0(V))}{\delta(1-\sigma_h)}\right)\frac{\psi'(e_0(V))/\pi'(e_0(V))}{\delta(1-\sigma_h)}.\end{aligned}$$

This is one equation in one unknown, σ_h. Notice that the left-hand side is independent of σ_h. Consider the right-hand side. The derivative with respect to σ_h is strictly positive as long as the argument of S and S' is less than V^*, and becomes zero afterwards.

As σ_h goes to 1, the right-hand side converges to δS^* (the second term vanishes for σ_h large enough, since S' is zero once $S = S^*$), which is bigger than $E_h(e_0(V))$ if V is small enough (i.e., $e_0(V) < \tilde{e}$, where $\tilde{e}$ is defined in Lemma 3).

If $\sigma_h = 0$, then since $E_h(0) > \delta S(0)$ and ψ'/π' goes to zero as V goes to zero (and all this is continuous in V), it follows that the left-hand side is strictly greater than the right-hand side for V sufficiently small. Thus $\sigma_h = 0$ cannot be a solution for V small enough.

Since the left-hand side starts above and ends below the right-hand side, the Intermediate Value Theorem yields a unique interior value for $\sigma_h(V)$ that satisfies the first-order condition above, and thus $(e_0(V), \sigma_h(V))$ is the unique critical point in this case. □.

Lemma 9. *If V is sufficiently small, then there is no solution to the first-order conditions with $V + \psi - \pi(\psi'/\pi') > 0$.*

Proof. Notice that any critical point (e, σ_h) in this case satisfies $e \in [0, e_0(V))$ and the following equations, which are the first-order conditions with respect to e and σ_h:

$$-\psi' + \sigma_h u_h' + \pi'\delta((1-\sigma_h)S(V_h) - S(V_l)) + \pi(1-\pi)\left(\frac{\psi'}{\pi'}\right)'(S'(V_h) - S'(V_l)) \le 0,$$
$$E_h - \delta S(V_h) + \delta S'(V_h)V_h = 0,$$

where $V_h = (V + \psi + (1-\pi)(\psi'/\pi'))/(\delta(1-\sigma_h))$ and $V_l = (V + \psi - \pi(\psi'/\pi'))/\delta$, and the argument e has been omitted to simplify the notation. We will show that when V is sufficiently small, these equations do not have a solution.

We first prove that there is a threshold $\hat{\sigma}_h(V)$ such that, for any pair (e, σ_h) that solves the first-order conditions, $\sigma_h > \hat{\sigma}_h(V)$. To see this, let $\hat{\sigma}_h(V) = \sigma_h$ that solves

$$E_h(0) = \delta S\left(\frac{V + a(e_0(V))}{\delta(1-\sigma_h)}\right) - \delta S'\left(\frac{V + a(e_0(V))}{\delta(1-\sigma_h)}\right)\left(\frac{V + a(e_0(V))}{\delta(1-\sigma_h)}\right), \tag{A.22}$$

where $a = \psi + (1-\pi)(\psi'/\pi')$ to simplify the notation. Following the same steps as in Lemma 8, one can show that there is a unique $\hat{\sigma}(V)$, which is interior, that solves (A.22). Moreover, since both $E_h(e)$ and $\delta S(V_h(e)) - \delta S'(V_h(e))V_h(e)$ are strictly increasing in e, it follows that *any* solution (e, σ_h) to the first-order conditions above must be such that $\sigma_h > \hat{\sigma}_h(V)$.

We next show that $\lim_{V\to 0} \hat{\sigma}_h(V) = 1$. Towards a contradiction, suppose not, i.e., suppose $\hat{\sigma}_h(V)$ converges to a number strictly less than one. Since $e_0(V)$ goes to zero as V goes to zero, it follows that (A.22) becomes $E_h(0) = \delta S(0)$, but $E_h(0) > \delta S(0)$, contradiction.

Finally, we show that for V sufficiently small, there is no solution to the first-order condition with respect to e. Let V be small enough so that any critical point has $\sigma_h \ge 1 - \varepsilon(V)$, where $\varepsilon(V)$ goes to zero as V goes to zero, and insert this value into the first-order condition for e. Consider the left-hand side of the first-order condition with respect to e. It is easy to verify that it is larger than

$$-\psi'(e_0(V)) + (1-\varepsilon(V))u_h'(e_0(V)) - \pi'(0)\delta S\left(\frac{V}{\delta}\right) - \frac{1}{4}\left(\frac{\psi'(e_0(V))}{\pi'(e_0(V))}\right)S'(0), \tag{A.23}$$

which converges to $u_h'(0) - \pi'(0)\delta S(0)$ as V vanishes.

This expression is positive when $E_h(0) > 0$. To see this, recall that $u_h = \pi E_h$ and thus $u_h' = \pi' E_h + \pi E_h'$. Hence $u_h' - \pi'\delta S(0) = \pi(0)E_h'(0) + \pi'(0)(E_h(0) - \delta S(0)) > 0$. By continuity, (A.23) is positive for V positive near zero. Hence, the first-order condition with respect to e cannot be satisfied when V is sufficiently small. □

Lemma 10. *If V is sufficiently small, then the critical point in Lemma 8, which consists of a positive e and an interior σ_h, solves the optimal long-term contracting problem.*

Proof. If the problem has a solution, then it satisfies the Kuhn–Tucker conditions. From the lemmata above it follows that when V is sufficiently small, there is a unique critical point. So if the problem has a solution it must be that point, which has $e > 0$ and $\sigma_h \in (0, 1)$. Moreover, by choosing V small enough one can ensure that $e < \tilde{e}$, and thus the omitted corner $\sigma_h = 1$ cannot be optimal by Lemma 3. □

Step 2: Verification of interior equilibrium σ_h. We must show that $\sigma_h \in (0, 1)$ can indeed be optimal for small values of V, since lotteries over contracts with $\sigma_h \in \{0, 1\}$ could make purchase lotteries redundant.

To this end, we will first analyze the (right) derivatives of U_{00} and U_{10} at $V = 0$. Recall that U_{00} is the value of not buying today irrespectively of the signal realization, and U_{10} is the value of buying today if and only if the signal realization is high, which solves

$$
\begin{aligned}
U_{10}(V) = & \max_{w_h, V_\ell, e} \pi(e)\left(E_h(e) - w_h\right) + (1 - \pi(e))\delta U(V_\ell) \\
\text{s.t.} \quad & -\psi(e) + \pi(e) w_h + (1 - \pi(e)) V_\ell = V, \\
& \psi'(e) = \pi'(e)\left(w_h - \delta V_\ell\right), \\
& V_\ell \geq 0, \quad w_h \geq 0, \quad e \geq 0.
\end{aligned}
$$

Solving for w_h and V_ℓ from the constraints and using the premise that $\gamma > \hat{\gamma}$, we obtain

$$
\begin{aligned}
U_{10}(V) = \max_{e \geq 0} \pi(e) & \left(E_h(e) - \left(V + \psi(e) + (1 - \pi(e)) \frac{\psi'(e)}{\pi'(e)} \right) \right) \\
& + (1 - \pi(e))\delta U \left(\frac{V + \psi(e) - \pi(e)\frac{\psi'(e)}{\pi'(e)}}{\delta} \right)
\end{aligned}
$$

subject to $V + \psi(e) - \pi(e)\psi'(e)/\pi'(e) \geq 0$ (i.e., $V_\ell \geq 0$.)

Lemma 11. *The (right) derivative of U_{10} diverges to infinity as V goes to zero, while the derivative of U_{00} is positive and finite at $V = 0$.*

Proof. In the U_{10} problem, it is easy to show that when V is sufficiently small, $V_\ell = 0$, and thus the optimal effort level is $e_0(V)$. Hence, $U_{10}(V)$ is

$$
U_{10}(V) = \pi(e_0(V)) \left(E_h(e_0(V)) - \frac{\psi'(e_0(V))}{\pi'(e_0(V))} \right) + (1 - \pi(e_0(V))\delta U(0).
$$

Differentiating and using $e_0'(V) = (\pi(\psi'/\pi'))^{-1}$ and $u_h' = \pi' E_h + \pi E_h'$, we obtain

$$
\begin{aligned}
U_{10}'(V) &= \left[\pi' \left(E_h - \frac{\psi'}{\pi'} - \delta U(0) \right) + \pi \left(E_h' - \left(\frac{\psi'}{\pi'} \right)' \right) \right] e_0'(V) \\
&= \frac{u_h' - \psi' - \pi' \delta U(0)}{\pi' \left(\frac{\psi'}{\pi'} \right)'} - 1.
\end{aligned}
$$

The limit of the numerator of the right-hand side as V goes to zero is $u_h'(0) - \pi'(0)\delta U(0)$, which is positive since $E_h(0) > 0$. The limit of the denominator is zero. Hence

$$
U_{10}'(0) = \lim_{V \to 0^+} U_{10}'(V) = +\infty.
$$

Regarding U_{00}, if at the optimum the constraint ($V_\ell \geq 0$) binds (which implies that effort is positive and thus this can only happen if there is an upward sloping portion of U) then

$$U_{00}(V) = \pi(e_0(V))\delta U\left(\frac{V + \psi(e_0(V)) + (1-\pi(e_0(V)))\frac{\psi'(e_0(V)}{\pi'(e_0(V))}}{\delta}\right) + (1-\pi(e_0(V)))\delta U(0).$$

Differentiating this expression and using the derivative of $e_0(V)$, we obtain

$$U'_{00}(V) = \pi'\delta(U(V_h) - U(0))e'_0(V) + \pi U'(V_h)\left(1 + (1-\pi)\left(\frac{\psi'}{\pi'}\right)' e'_0(V)\right)$$
$$= \delta\frac{\pi'}{\pi}\frac{U(V_h) - U(0)}{\left(\frac{\psi'}{\pi'}\right)'} + U'(V_h).$$

As V goes to zero, the last term on the right-hand side converges to $U'(0)$. Regarding the first term, notice that

$$0 \leq \delta\frac{\pi'}{\pi}\frac{U(V_h) - U(0)}{\left(\frac{\psi'}{\pi'}\right)'} \leq \delta\frac{\pi'}{\pi}\frac{U'(0)\left(\frac{\psi'}{\pi'}\right)}{\left(\frac{\psi'}{\pi'}\right)'} = \delta\frac{\gamma\beta_h - (1-\gamma)\beta_\ell}{\pi}U'(0)\frac{\psi'}{\left(\frac{\psi'}{\eta'}\right)'},$$

where the first inequality follows from U being upward sloping in the region considered, the second one from concavity of U, and $e = e_0(V)$ in each argument. Now, $U'(0)$ is finite since $E_h(0) > 0$.[34] Hence, the last term converges to a 0/0 expression as V vanishes, since $\psi'(0) = (\psi'/\eta')'(0) = 0$. Applying L'Hopital's rule, we obtain that this term vanishes given our premise $(\psi'/\eta')''(0) > 0$ and the fact that $(\psi'/\eta')'(0) = 0$ implies $\psi''(0) = 0$. Hence, $U'_{00}(0) = U'(0) < \infty$.

If the constraint ($V_\ell \geq 0$) does not bind, then it follows from the Envelope Theorem that

$$U'_{00}(V) = \pi U'(V_h) + (1-\pi)U'(V_l).$$

As V goes to zero, so does e (since the feasible set of values of e shrinks to $\{0\}$). Hence V_h and V_l go to zero and thus $U'_{00}(0) = U'(0) < \infty$. □

The following lemma now completes the proof of Proposition 3.

Lemma 12. *When $V > 0$ is sufficiently small, $\sigma_h \in (0, 1)$ at the optimal contract.*

Proof. Let $\tilde{e}$ solve $E_h(\tilde{e}) = \delta S^*$, and let $\tilde{V} = \inf\{V | e(V) \geq \tilde{e}\} > 0$. Since we are assuming that $\gamma > \tilde{\gamma}$, it follows that $U_{11}(0) > U_{10}(0) > U_{00}(0)$. So consider the lottery over contracts that randomizes between the contract that involves always buying, zero effort, and $V = 0$ with probability p; and buying only if high, positive effort, and a positive V with probability $1 - p$. (Graphically, this is the line segment that starts at the intercept of U_{11} and is tangent to U_{10} at V.)

[34] This follows since $U(V) = U_{11}(V) = \gamma y_h + (1-\gamma)y_\ell - V$ at $V = 0$ under our premises, and thus $U'(0)$ is equal to the finite slope of a convex combination of $U_{11}(0)$ with $U_{00}(V)$ or $U_{10}(V)$ for some $V > 0$ if, contrary to what we are trying to prove, the purchase lotteries are dominated by the contract ones.

Notice that this value of $V > 0$ need not be smaller than $\tilde{V}$; moreover, it is not clear at this point that U_{00} is below this line segment for *all* values in the interval between 0 and V.

Let $\gamma \to \tilde{\gamma}^+$, so that $U_{11}(0)$ converges to $U_{10}(0)$. Since $U'_{10}(0) = \infty$ and $U'_{00}(0) < \infty$, it follows that for values of γ sufficiently close to $\hat{\gamma}$ both $V < \tilde{V}$ and U_{00} is below the line segment joining $U_{11}(0)$ and $U_{10}(V)$.

Since $V < \tilde{V}$, it follows that $\sigma_h = 1$ cannot be optimal. So it is either zero, or interior. We want to rule out the former possibility. Recall that we have set $\sigma_\ell = 0$. Since U_{00} is below U_{10} at V, $\sigma_\ell = \sigma_h = 0$ cannot optimal. Furthermore, $\sigma_\ell = 1$ and $\sigma_h = 0$ cannot be optimal either by Claim 1. Hence, there are three possibilities left: (i) $\sigma_\ell = 0$ and $\sigma_h \in (0, 1)$; (ii) $\sigma_\ell \in (0, 1)$ and $\sigma_h \in (0, 1)$; or (iii) $\sigma_\ell \in (0, 1)$ and $\sigma_h = 0$.

If (i) dominates (ii) and (iii) we are obviously done since we have shown that this alternative is better than a lottery over contracts with deterministic purchase decisions. If (ii) dominates (i) and (iii) we are also done, since by transitivity this alternative will dominate a lottery over payments, as it dominates (i). And we claim that we can rule out the case in which (iii) dominates (i) and (ii) by making V small enough. To see this, notice that $U_{11}(0) = \gamma y_h + (1-\gamma)y_\ell > U_{00}(0) = \delta(\gamma y_h + (1-\gamma)y_\ell)$ and $U_{11}(0) = \gamma y_h + (1-\gamma)y_\ell > U_{01}(0) = [1 - \pi(0) + \pi(0)\delta](\gamma y_h + (1-\gamma)y_\ell)$. As a result, and using obvious notation, $U_{11}(0) > U_{0\sigma_\ell}(0) = [(1-\pi(0))(\sigma_\ell + (1-\sigma_\ell)\delta) + \pi(0)\delta](\gamma y_h + (1-\gamma)y_\ell)$ for all $\sigma_\ell \in [0, 1]$.[35] Since the inequalities are strict, the same is true for $V > 0$ sufficiently small. Hence, case (iii) will be dominated by U_{11} for V small enough, and thus by either cases (i) or (ii). This completes the proof. □

A.10. Use of lotteries: additional results

Claim 5. *If $\gamma \le \tilde{\gamma}$ (so that $E_h(0) = E_\ell(0) \le 0$), then $\sigma_\ell = 0$ for all V.*

Proof. Suppose that $\gamma \le \tilde{\gamma}$ so that $E_h(0) = E_\ell(0) \le 0$. Using part (ii) of Proposition 2, at $V = 0$ $\sigma_h = \sigma_\ell = 0$ are optimal, and the payoff to the principal is $U(0) = \delta U(0)$ implying that $U(0) = 0$. Then $-S(V) + S'(V)V = 0$ at $V = 0$.[36] Moreover, $-S(V) + S'(V)V$ is decreasing in V since S is concave, so it is negative for all V. Hence the left-hand side of (11) for $i = \ell$ is always strictly negative, implying that $\sigma_\ell = 0$. □

Claim 6. *If $\gamma < \hat{\gamma}$ and $e > 0$, then $\sigma_h \ge \sigma_\ell$. Moreover, at the optimum either $\sigma_\ell = \sigma_h = 0$, or $0 = \sigma_\ell < \sigma_h$, or $\sigma_\ell < \sigma_h = 1$.*

Proof. Recall that when σ_i is interior, the corresponding first-order condition is $E_i(e) - \delta[S(V_i) - S'_i(V_i)V_i] = 0$ (assuming for simplicity that S is differentiable at V_i). Notice that $S(V_i) - S'_i(V_i)V_i$ is increasing in V_i since S is concave, and strictly increasing if $V_i < V^*$. Moreover, $E_h(e) > E_\ell(e)$ for any $e > 0$, and under $\gamma < \hat{\gamma}$ we have that $V_\ell > V_h$ (see Proposition 2). Hence $E_\ell(e) - \delta[S(V_\ell) - S'_\ell(V_\ell)V_\ell] < E_h(e) - \delta[S(V_h) - S'_h(V_h)V_h]$, so both σ_h and σ_ℓ cannot be interior, and $\sigma_\ell \le \sigma_h$. Since $\sigma_h = \sigma_\ell = 1$ is inconsistent with $e > 0$, we are left with three possible cases: (1) $\sigma_\ell = \sigma_h = 0$, (2) $0 = \sigma_\ell < \sigma_h$, and (3) $\sigma_\ell < \sigma_h = 1$. □

[35] This is the value of purchasing with probability σ_ℓ after the low signal realization when $V = 0$.

[36] Notice that $\lim_{V \to 0} S'(V)V$ must be finite and equal to zero, for otherwise the first-order condition (11) for $i = \ell$ that should hold with a weak inequality at $V = 0$ would be violated.

A.11. Optimal effort: proof of Proposition 4

Towards a contradiction, suppose that $e(V) > e^{**}$. Recall that $e^{**} > \tilde{e}$. Thus by Lemma 3, $\sigma_h = 1$. Recall that the first-order condition with respect to effort in the first-best case is $f^* \equiv -\psi'(e) + u_h'(e) - \pi'(e)\delta S^* = 0$. Using $\sigma_h = 1$, the corresponding first-order condition in the second-best case can be written as $f \equiv -\psi'(e) + u_h'(e) + \sigma_\ell u_\ell'(e) - \pi'(e)(1 - \sigma_\ell)\delta S(V_\ell) + \mu(\eta''(e)\psi'(e)/\eta'(e) - \psi''(e)) = 0$. Then $f = f^* + \sigma_\ell u_\ell'(e) + \pi'(e)\delta[S^* - (1 - \sigma_\ell)S(V_\ell)] + \mu(\eta''(e)\psi'(e)/\eta'(e) - \psi''(e))$. Since $\eta''\psi'/\eta' - \psi'' < 0$, $\mu \geq 0$, $u_\ell' < 0$, $S^* \geq S(V_\ell) \geq (1 - \sigma_\ell)S(V_\ell)$, and $\pi' < 0$ when $\gamma < \hat{\gamma}$, we have $f \leq f^*$, and hence $e \leq e^{**}$, with strict inequalities unless $\mu = 0$. □

Appendix. Supplementary material

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.jet.2015.12.003.

References

Abreu, D., Pearce, D., Stacchetti, E., 1986. Optimal cartel equilibria with imperfect monitoring. J. Econ. Theory 39, 251–269.

Abreu, D., Pearce, D., Stacchetti, E., 1990. Toward a theory of discounted repeated games with imperfect monitoring. Econometrica 58, 1041–1063.

Athey, S., Levin, J., 2001. The value of information in monotone decision problems. Stanford University Working Paper.

Aubin, J., 1979. Mathematical Methods of Game and Economic Theory. North-Holland, New York.

Blackwell, D., Girshick, M., 1954. Theory of Games and Statistical Decisions. Dover Publications Inc., New York.

Borwein, J., Lewis, A., 2006. Convex Analysis and Nonlinear Optimization. Springer, New York.

Chade, H., Schlee, E., 2002. Another look at the Radner–Stiglitz nonconcavity in the value of information. J. Econ. Theory 107, 421–452.

Eso, P., Szentes, B., 2007. The price of advice. Rand J. Econ. 38, 863–880.

Gromb, D., Martimort, D., 2007. Collusion and the organization of delegated expertise. J. Econ. Theory 137, 271–299.

Grossman, S., Hart, O., 1983. An analysis of the principal–agent problem. Econometrica 51, 7–45.

Inderst, R., Ottaviani, M., 2012. Competition through commissions and kickbacks. Am. Econ. Rev. 102, 780–809.

Jewitt, I., 2007. Information order in decision and agency problems. Nuffield College Working Paper.

Jewitt, I., Kadan, O., Swinkels, J., 2008. Moral hazard with bounded payments. J. Econ. Theory 143, 59–82.

Kashyap, A., Kovrijnykh, N., forthcoming. Who should pay for credit ratings and how? Rev. Financ. Stud. http://dx.doi.org/10.1093/rfs/hhv127.

Krahmer, D., Strausz, R., 2011. Optimal procurement contracts with pre-project planning. Rev. Econ. Stud. 78, 1015–1041.

Levitt, S., Snyder, C., 1997. Is no news bad news? Information transmission and the role of "early warning" in the principal–agent model. Rand J. Econ. 28, 641–661.

Lewis, T., Ottaviani, M., 2008. Search agency. Northwestern University Working Paper.

MasColell, A., Whinston, M., Green, J., 1995. Microeconomic Theory. Oxford University Press, New York.

Mason, R., Valimaki, J., 2015. Getting it done: dynamic incentives to complete a project. J. Eur. Econ. Assoc. 13, 62–97.

Quah, J., Strulovici, B., 2009. Comparative statics, informativeness, and the interval-dominance ordering. Econometrica 77, 1949–1992.

Rogerson, W., 1985a. The first-order approach to principal–agents problems. Econometrica 53, 1357–1367.

Rogerson, W., 1985b. Repeated moral hazard. Econometrica 53, 69–76.

Sannikov, Y., 2008. A continuous-time version of the principal–agent problem. Rev. Econ. Stud. 75, 957–984.

Spear, S., Srivastava, S., 1987. On repeated moral hazard with discounting. Rev. Econ. Stud. 54, 599–617.

Szalay, D., 2005. The economics of clear advice and extreme options. Rev. Econ. Stud. 72, 1173–1198.

Szalay, D., 2009. Contracts with endogenous information. Games Econ. Behav. 65, 586–625.

Toxvaerd, F., 2006. Time of the essence. J. Econ. Theory 129, 252–272.