CSEF

Centre for Studies in Economics and Finance

WORKING PAPER NO. 258

Contracts with Wishful Thinkers

Giovanni Immordino, Anna Maria C. Menichini, Maria Grazia Romano

September 2010

This version October 2012

University of Naples Federico II

University of Salerno

Bocconi University, Milan

CSEF - Centre for Studies in Economics and Finance
DEPARTMENT OF ECONOMICS – UNIVERSITY OF NAPLES
80126 NAPLES - ITALY
Tel. and fax +39 081 675372 – e-mail: csef@unisa.it

WORKING PAPER NO. 258

Contracts with Wishful Thinkers

Giovanni Immordino*, Anna Maria C. Menichini*, Maria Grazia Romano*

Abstract

In a setting with a wishful thinking agent and a realistic principal, the paper studies how incentive contracts should be designed to control for both moral hazard and self-deception. The properties of the contract that reconciles the agent with reality depend on the weight the agent attaches to anticipatory utility. When this is small, principal and agent agree on full recollection. For intermediate values the principal bears an extra cost to make the agent recall bad news. For large weights the principal renounces inducing signal recollection. We extend the analysis to the case in which the parameter of anticipatory utility is private information. The distinction between the two settings assumes practical relevance if preferences can be related to personality characteristics as in this case the parameter of anticipatory utility could be learned through psychological testing.

JEL Classification: D03, D82, D86.

Keywords: Self-deception, anticipatory utility, contracts.

Acknowledgement: We are indebted to Utpal Bhattacharya, Yeon-Koo Che, Andy Newman and Marco Pagano for many useful insights. We also thank participants to the the First Workshop ME@Velia (Ascea Marina), to the 6th CSEF-IGIER Symposium in Economics and Institutions (Capri), as well as seminar participants at the Université de Paris Dauphine (Paris) and Cass Business School (London) for their comments. A previous version has been circulated under the title "Optimal Compensation Contracts for Optimistic Managers.

* Università di Salerno and CSEF. Contact address: Dipartimento di Scienze Economiche e Statistiche, University of Salerno, Via Ponte Don Melillo, 84084 Fisciano (SA), Italy. E-mail: giimmo@tin.it.

Table of contents

1 Introduction

Screening emotional aspects of would-be employees before hiring them is a common practice among firms, rooted in the widespread psychological evidence showing that most individuals hold overly positive evaluations about the self, exaggerated perceptions of control or unrealistic optimism (Taylor and Brown, 1988).[1] Despite this widespread practice, it is not clear the role of psychological traits in determining employee's compensation and in driving the firms' hiring decisions.

In this paper we study, within a moral hazard setting, an employment contract between an endogenously optimistic agent and a realistic principal. Optimism is modeled assuming that the agent enjoys anticipatory utility, i.e., derives utility from the anticipation of his future payoff: the greater his future payoff, the greater his current utility. A greater anticipatory utility can be achieved by suppressing current bad information affecting future payoffs, thus expecting good outcomes more often than is warranted. But because distorted beliefs distort actions, optimism has an influence on decisions and exacerbates incentive problems. We study how the need to control for both optimism and moral hazard affects the design of incentive contracts.

To analyze this problem, we develop, within a simple contract-theoretic framework, a model that unifies various themes from psychology and economics. According to an influential literature in psychology (Taylor and Brown, 1998) normal mental functioning is skewed in a positive direction and processes of self deception - the active misrepresentation of reality to the conscious mind - are characteristic of mental health (Trivers, 2000). The resulting biases guide the processing of information, such that mildly negative or ambiguous information is distorted to be more positive than may actually be the case. In particular, individuals adjust to threatening events by constructing benign interpretations of the same.[2] One of the many forces that may

[1]Such attitudes are especially documented for businessmen. For example, Cooper et al. (1988) argue that entrepreneurs see their own chances for success higher than that of their peers, while Malmendier and Tate (2005, 2008) find evidence that CEOs over-estimate their firms' future performance.

[2]Freud (1940, 1957) argues that when events from the internal and external world are highly threatening, people may deny or repress their implications in order to avoid intolerable anxiety. Denial involves a distortion of negative experiences so complete that it can block out memory of the experience altogether (cited in Taylor at al., 1989).

favour mechanisms of self-deception is that positive illusions may give intrinsic benefits (Taylor and Brown, 1989; Trivers, 2000). As Taylor describes it: "It is just as easy to construe future events in a manner that promises success and happiness rather than one that portends failure. Self-deception can be healthful and bolstering if it doesn't involve gambling one's resources beyond salvage". The beneficial effect of self-deception is modeled in the economic literature by assuming that, prior to the resolution of uncertainty, individuals experience feelings of anticipation. Because of them, through imperfect memory, they select their beliefs so as to enjoy the greatest comfort or happiness, thus leading to cognitive dissonance.

However, there are limits to the extent of self-deception. "At one level, [the normal human mind] constructs beneficent interpretations of threatening events that raise self-esteem and promote motivation; yet at another level, it recognizes the threat or challenge that is posed by these events" (Taylor, 1988). To capture the "consciousness/awareness" that rejoins individuals with reality, most of the economic literature has assumed individuals to be Bayesian information processors (Bénabou and Tirole, 2002): they are aware of the flaws of memory and in making choices take into account the possibility of having suppressed unfavourable information.

Our paper builds on this literature and, by applying a game of belief management à la Bénabou and Tirole (2002) into a contract-theoretic framework, inquires into how a principal should reward a forgetful agent in the awareness that well designed rewards can limit his tendency to self-deception. In this respect, the contracting framework represents a further mechanism through which the individual can be reconciled with reality and one of the main contributions of this paper.

In our model a risk-neutral principal hires a risk-neutral agent for a project. When the principal offers the contract, the parties are symmetrically informed. If the agent accepts, he chooses a level of effort that affects the project's probability of success. After signing, but before choosing his unobservable effort, the agent receives a private signal about the profitability of the task. A good signal implies a high return in case of success, a bad signal only intermediate return. Finally, in case of failure, the return is zero regardless of type of signal. If the signal is informative about the return from effort, the agent would benefit from having accurate news. However, since he derives utility from the anticipation of his final payoff, the suppression of a bad signal may induce a positive interim emotional effect. We assume that the principal cannot observe the agent's choice. Thus, to induce him to choose the right action, she makes compensation contingent on project revenues. More specifically, parties can write a complete contract specifying the rewards contingent on the various outcomes, the effort levels to be exerted contingent on the recalled signals, and the probability that bad news will be remembered accurately. Does the optimal contract always ensure complete recall?

We show that if the agent's anticipatory utility is sufficiently low, there is no conflict of interest between principal and agent's desired recall. There is a conflict for intermediate values of this parameter, and the principal chooses to bear the extra cost necessary to have the agent recall the bad signal and exert the right level of effort. Finally, for a sufficiently high weight on anticipatory utility, the principal becomes indifferent between inducing signal recollection and not, and the optimal contract is characterized by a pooling equilibrium reminiscent of adverse selection models.

Why does the optimal contract look like this? Informed agents face a trade-off between ensuring that the level of effort they choose reflects accurate news and savoring emotionally gratifying good news. However, the agent's preferred level of memory may differ from the principal's. As a result, in writing the contract, the principal may want to affect this dimension of the agent's choice. If in the agent's utility the weight of emotions is sufficiently small, accurate news becomes a priority for the agent too, there is no conflict over information recollection, and a contract can attain the optimal recall at no extra cost.[3] For larger weights of anticipatory utility, the agent's trade-off tilts away from accurate news towards good news, so that enticing proper information recollection becomes costly. For intermediate values of the parameter, the principal chooses to move the agent's trade-off towards accuracy by making it costly, when the true signal is bad, to recall a good signal. This is done by increasing the cost to the agent when he exerts the effort expected for the good rather than the bad signal. But if the weight on emotions instead is sufficiently large, the optimal contract calls for a pooling equilibrium in which the agent exerts

[3] No extra cost with respect to the resources needed to solve the moral hazard problem.

the same level of effort and receives the same payments regardless of signal type. Intuitively, when the weight on anticipatory utility is high, the agent will recall the signal accurately only if he does not anticipate a lower payment when the signal is bad.[4]

The analysis so far implicitly relies on the assumption that the principal knows what weight the agent attaches to anticipatory utility. We extend the analysis to the opposite extreme, assuming that this parameter is the agent's private information. In this case, the optimal contract will be designed so as to induce all agents with a parameter of anticipatory utility below a threshold level to recall the bad signal, and all those above it to forget. Relative to the perfect information benchmark, the set of agents who will be induced (by the contract) to recall the bad signal is now larger and also includes some agents who, due to the conflict of interest, impose an extra recollection cost on the principal.

The distinction between the two information settings we have just discussed is of particular relevance if the parameter capturing anticipatory utility, where not observable, can be learned, for example through psychological testing. In this case, the perfect information assumption becomes a plausible starting point to study the role of emotional aspects in agent's compensation and recruitment, with the incomplete information case capturing situations in which psychological tests are not carried out. However, the extent to which the parameter of anticipatory utility can be captured through personality testing practices, and thus through personality traits, is a question that has not been dealt with so far in the literature. To our knowledge, this is the first paper to conjecture a link between these two aspects.

The issue is then to relate psychological characteristics to our measure of anticipatory utility. In this respect, conscientiousness and emotional stability, defined in the psychological literature respectively as dependable and cool, seem to capture our parameter of anticipatory utility. Thus, people who are unreliable, daydreamer or anxious and emotional may for self-assurance be more prone to self-deception, i.e., to distort their assessments of the likelihood of future events (discard bad news). This leads them to expect good outcomes too often, or more often than they ought

[4] There also exists an outcome-equivalent equilibrium where investors prefer not to elicit information recollection, the manager never recalls a bad signal, and the level of effort is the same as in the pooling equilibrium.

to, and thus to overoptimism.

By relating the parameter of anticipatory utility to psychological factors, employee's optimism is founded directly upon the psychological traits that may give rise to it. This should help to base future empirical work directly on agents' personality traits rather than on their likely consequences. In this view, hiring and compensation policies could be explained by personality traits. For instance, one of the implications of our analysis is that the agent's "realism" –captured by high conscientiousness and emotional stability– is always valuable for the principal, but more so in less risky sectors.

This view that conjectures a link between anticipatory utility and personality traits finds some preliminary support in a recent literature that relates time preference to psychological factors. For instance Borghans, Duckworth, Heckman and ter Weel (2008) and Almlund, Duckworth, Heckman and Kautz (2011) identify some of these factors with conscientiousness, self-control, and consideration of future consequences. Recent experimental evidence by Daly, Delaney and Harmon (2009) confirms this conjecture, finding that conscientiousness associates negatively with the discount rate.

The paper is organized as follows. After presenting a brief review of the relevant economic and experimental literature in the next subsection, Section 2 presents the model. Section 3 sets out the results concerning the conflict between principal's and agent's optimal recall. Section 4 characterizes the optimal contract and identifies the type of the agent that the principal would choose. Section 5 makes further progress in the characterization of the optimal contract by assuming a quadratic cost function. Section 6 extends the analysis to the case in which the parameter of anticipatory utility is the agent's private information. Section 7 concludes. Proofs are in the Appendices.

1.1 Related literature

Economic literature. The paper is related to a vast economic literature that studies cognitive dissonance, anticipatory feelings and motivation effects in incentive design.

Cognitive dissonance (see, among others, Akerlof and Dickens, 1982; Rabin, 1994; Carrillo and Mariotti, 2000; Bénabou and Tirole, 2002) occurs when a person does something that is inconsistent with his beliefs. Most people feel uncomfortable in managing this discrepancy between behavior and beliefs. Thus, they prefer to reduce cognitive dissonance, and, rather than doing that through changing their behavior, they do it by changing their beliefs. In Akerlof and Dickens (1982), for example, agents choose their beliefs to minimize the unpleasant feelings arising from learning the risks involved in a hazardous job. In Rabin (1994), "increasing people's distaste for being immoral can increase the level of immoral activities. This can happen because people will feel pressure to convince themselves that immoral activities are in fact moral" and thus acceptable.

In the papers based on anticipatory utility (see, among others, Loewenstein, 1987; Caplin and Leahy, 2001; Bernheim and Thomadsen, 2005; Brunnermeier and Parker, 2005; Köszegi, 2006), it is assumed that individuals derive utility not only from current consumption, but also from anticipating future consumption. As a result, current instantaneous utility depends positively on future consumption. Such a setting allows Loewenstein (1987) to explain why one might want to get an unpleasant experience over with quickly – so as to shorten the period of dread – and at the same time delay a pleasant one, so as to savor it. Caplin and Leahy (2001) incorporate anticipatory utility into an expected utility model, and show that time inconsistency arises as a result.[5]

In underlying potential psychological effects of the contracting framework, the paper is linked to Bénabou and Tirole (2003) and Fang and Moscarini (2005).[6] The first authors, in a setting in which the agent has imperfect knowledge about his ability and undertakes a certain task only if he has sufficient confidence in his ability to succeed, study how an informed principal should reward the agent knowing that rewards can undermine intrinsic motivation. Similarly, Fang and Moscarini (2005) study the design of wage contracts that provide incentives and affect work

[5] Köszegi (2006) adapts Caplin and Leahy (2001) framework to allow for cheap-talk communication, and for an action to be taken by one of the parties, while Bernheim and Thomadsen (2005) study the role of anticipatory emotions in decision making under memory imperfections, showing that, to increase anticipatory utility, individuals can influence the memory process.

[6] Other papers on motivation effects in incentive design are Ishida (2006), Swank and Visser (2007) and Crutzen, Swank and Visser (2010).

morale. Thus, whereas in our paper rewards serve to limit wishful thinking, in these papers they manipulate motivation. This is to be ascribed to the different information structures of the two settings: while in the above contributions the principal has ex ante information about the agent's characteristics, in our paper parties are ex-ante symmetrically informed.

Finally, the paper is related to a companion one (Immordino, Menichini and Romano, 2011) in which the authors show that the emotional aspects may render it impossible to implement the first-best output, thus providing a negative result. Specifically, although parties are symmetrically informed and contracts are complete, it may be impossible to achieve the first-best if the weight on emotions is too high. In the present paper, instead, starting from a second best world, we study the properties of the contract that reconciles the agent with reality. Therefore, the two analyses complement each other.

Experimental literature. The paper is also related to some recent experimental literature documenting people's tendency to process information in a biased manner. For example, individuals tend to view themselves as intelligent and will process information in a biased way to support that belief. Mobius, Niederle, Niehaus and Rosenblat (2011), in an experiment with college students who perform an IQ test for which they receive noisy signals of their performance, find that subjects systematically discount bad news about their own intelligence. Similarly, Eil and Rao (2011) find that subjects are asymmetric updaters: close to proper Bayesian updating for positive news about their intelligence and beauty, but systematically underresponding to negative news. Both studies find that agents are averse to information and willing to pay to avoid learning their type.

A differential response to good and bad news is also documented in Mayraz (2011) who designs an experiment in which subjects observing a financial asset's historical price chart have to predict the price of the asset. They receive both an accuracy bonus for predicting the price at some future point, and an unconditional award that is either increasing or decreasing in this price. It turns out that subjects gaining from high prices make significantly higher predictions than those gaining from low prices, with the magnitude of the bias independent of the amount paid for accurate predictions. Last, recent experimental evidence shows that belief distortion responds to incentives. In particular, Mijovic-Prelec and Prelec (2010) propose an experimental study showing that self-deceptive judgements can be elicited with financial incentives, with the latter affecting even the degree of self deception. With the many studies showing positive incentive effects of monetary rewards, i.e. higher incentives lead to higher effort (Bailey et al., 1998; Sprinkle, 2000), a further step forward in the literature should go in the direction of providing evidence of the core mechanism of the paper, that is, that incentive contracts can ensure at the same time effort provision and memory recollection.

2 The model

Players and environment

In our model a risk-neutral principal hires a risk-neutral agent for a project. When the principal offers the contract, the parties are symmetrically informed. If the agent accepts, he chooses a level of effort $a \in [0, 1]$ that affects the project's probability of success. The effort has disutility $c(a)$, with $c(0) = 0$, $c'(a) \geq 0$, $c''(a) > 0$ and $c'''(a) \geq 0$. In order to ensure interior solutions, we also assume that $c'(0) = 0$ and $c'(1) \geq v_H$. After signing, but before choosing his unobservable effort, the agent receives a private signal $\sigma \in \{L, H\}$ about the profitability of the task. A good signal implies a high return in case of success, a bad signal only an intermediate return (in case of success). The probability of a good signal H is q, that of a bad signal L is $(1-q)$, with $q \in [0, 1]$. Finally, in case of failure, the return is v_0 regardless of type of signal. In our setting, good (bad) news means that the outcome is v_H (v_L) or v_0 with probability a and $1-a$, respectively.[7] Thus, the project has three possible outcomes, $\widetilde{v} \in \{v_0, v_L, v_H\}$, with $v_0 < v_L < v_H$ and each outcome occurs with probabilities $1-a, (1-q)\,a$ and qa, respectively.[8] From now on, for the sake of simplicity and without loss of generality, we normalize $v_0 = 0$.

Since the signal gives information on the return from effort, in choosing its level the agent

[7] In other words, we assume for simplicity that the signal is perfectly correlated with the return $\tilde{v}$, implying that $\Pr(\tilde{v} = v_0|\sigma = L) = \Pr(\tilde{v} = v_0|\sigma = H) = 1 - a$, and $\Pr(\tilde{v} = v_L|\sigma = H) = \Pr(\tilde{v} = v_H|\sigma = L) = 0$.

[8] This modelling choice that distinguishes between project's characteristics and agent's effort allows us to deal in the simplest possible way with two imperfections: moral hazard and imperfect recall.

would benefit from accurate news. But if the agent derives utility from the anticipation of his final payoff, the suppression of a bad signal may induce a positive emotional effect. This is modeled assuming, as in Caplin and Leahy (2001, 2004) and Köszegi (2006), that total utility is a convex combination of the actual physical outcome and the anticipation of it, with weights $1-s$ and s, respectively, where s is the realization of a random variable distributed over the compact support $S \equiv [0,1]$ according to the twice continuously differentiable and atomless cumulative distribution function $F(s)$, with density $f(s)$. We assume that the parameter s is observed by the principal (this assumption is relaxed in Section 6).

At the time of the effort decision a bad signal can be forgotten (voluntarily repressed). We then denote by $\widehat{\sigma} \in \{L, H\}$ the recollection at that time of the news σ, and by $\lambda \in [0,1]$ the probability that bad news will be remembered accurately, that is, $\lambda \equiv \Pr(\widehat{\sigma} = L|\sigma = L)$. We assume that the agent can costlessly increase or decrease the probability of recollection.[9] Finally, we denote by "Agent 1" the agent's self at time 1 and by "Agent 2" the agent's self at time 2.[10]

The principal cannot observe the agent's action directly. Hence, to induce the "right" action, she can only offer the agent rewards contingent on observable, verifiable project revenues. We denote by $C \equiv \{w_0, w_L, w_H\}$ the contract that the principal offers the agent, where w_i is the reward corresponding to $v = v_i$, for any $i = 0, L, H$. We assume that the agent has limited liability, so that $w_i \geq 0$ for any i. Finally, we maintain the standard assumption of individuals as rational Bayesian information processors.[11]

Timing

The precise sequence of events unfolds as follows (Figure 1):

t=0: The principal offers a contract C to the agent to run a project.

t=1: If Agent 1 refuses the contract, the game ends. If he accepts, he observes a private signal $\sigma \in \{L, H\}$ and, when the signal is bad ($\sigma = L$), chooses whether to recall or not.

[9] Assuming costly recollection would not change our results qualitatively.

[10] In the present paper we will omit the details of the derivation of the memory game between the two selves of the agent since this has already been derived in our companion paper (Immordino, Menichini and Romano, 2011).

[11] By modeling agents as Bayesian, we are treating them as fairly sophisticated. A departure from this assumption though, while seemingly realistic, would be arbitrary, as there may be several ways of being less sophisticated.

t=2: Agent 2 observes $\widehat{\sigma} \in \{L, H\}$, updates his beliefs on the outcome v accordingly, selects the effort level a and enjoys utility from the anticipation of his future prospects, i.e., $E_2[U_3]$, where E_2 denotes current available information.

t=3: The project payoff is realized and the payment is executed.

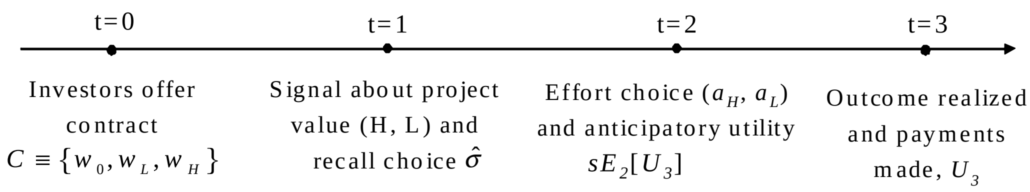

Fig. 1: Time-line.

Equilibrium

The equilibrium concept is Perfect Bayesian Equilibrium. The game can be solved first identifying the agent's optimal effort choice a, given his beliefs about σ and a contract C. Then, finding the equilibrium of the memory game for any given contract.[12] Finally, by computing the principal's contract offer using the agent's optimal effort choice rule and his inference.

3 The conflict over optimal recollection between principal and agent

In this section we show that the principal always prefers perfect recollection (Proposition 1). Then, we show that if the weight on anticipatory utility is sufficiently high, the agent prefers to forget bad news when he is offered the second-best contract (Proposition 2). This points to a potential conflict of interest between principal and agent over the memory strategy.

To establish the former result, we proceed in two steps. First we solve the principal's maximization problem under the assumption that λ is exogenous and common knowledge. Then we find the optimal level of λ from the principal's point of view.

[12] To derive the equilibrium of the memory game, i) for any realized σ, Agent 1 chooses his message $\widehat{\sigma}$ to maximize his expected utility, correctly anticipating the inferences that Agent 2 will draw from $\widehat{\sigma}$, and the action that he will choose; ii) Agent 2 forms his beliefs using Bayes' rule to infer the meaning of agent 1's message, knowing his strategy.

Given that the agent's effort is not observable, the principal must offer an incentive-compatible contract that induces the agent to choose the desired level of effort.

Faced with the contract $C \equiv \{w_0, w_L, w_H\}$ and the recalled signal $\widehat{\sigma}$, at $t = 2$ the agent chooses a level of effort $a_{\widehat{\sigma}}$ such that

$$a_{\widehat{\sigma}} = \arg\max_{a \in [0,1]} \{E_2 [U_3] \equiv -c(a) + E_2 [u(C, a)|\widehat{\sigma}]\}, \tag{1}$$

where $E_2 [u(C, a)|\widehat{\sigma}]$ is the sum of the agent's material payoff, $(1 - s)E_2 [u(C, a)|\widehat{\sigma}]$, and the anticipatory utility experienced by savoring the future material payoff, $sE_2 [u(C, a)|\widehat{\sigma}]$. Notice that when $\widehat{\sigma} = L$, the agent is sure that $\sigma = L$, and the expected payoff simplifies to

$$E_2[u(C, a)|L] = aw_L + (1 - a) w_0.$$

But when $\widehat{\sigma} = H$, the agent is unsure whether he actually received a good signal or instead received a bad signal and censored it. The expected payoff is

$$E_2[u(C, a)|H] = a (r(\lambda)w_H + (1 - r(\lambda))w_L) + (1 - a) w_0,$$

where $r(\lambda)$ is the likelihood of an accurate signal recollection and, by the Bayes' rule, it is given by

$$r(\lambda) \equiv \Pr(\sigma = H|\widehat{\sigma} = H, \lambda) = \frac{q}{q + (1 - q)(1 - \lambda)}.$$

Denote by $a(\lambda, C) \equiv \{a_L, a_H(\lambda)\}$ the vector of effort levels that solves problem (1), where for $\widehat{\sigma} = L$ is

$$w_L = c'(a_L), \tag{2}$$

and for $\widehat{\sigma} = H$ is

$$(r(\lambda)w_H + (1 - r(\lambda))w_L) = c'(a_H).^{13} \tag{3}$$

When the principal makes her offer, the agent does not know σ. So to induce him to accept, the contract has to satisfy the following ex-ante participation constraint (see equation 12 in the Appendix A)

$$E_0 [U (C, a(\lambda, C), \lambda)] \geq 0, \tag{4}$$

[13] Where $(r(\lambda)w_H + (1 - r(\lambda))w_L) \in [c'(0), c'(1)]$ for any r.

where E_0 denotes the expectation at time $t = 0$. Finally, by limited liability, the agent's payments must always be non-negative, i.e.

$$w_i \geq 0 \; \forall i \in \{0, L, H\}. \tag{5}$$

If a $\lambda-$type agent accepts contract C, the principal's expected profit in period 0 is

$$E_0 \left[\Pi \left(C, a(\lambda, C), \lambda\right)\right] = \Pr{}_0 \left(v_H|\lambda\right) \left(v_H - \left(w_H - w_0\right)\right) + \Pr{}_0 \left(v_L|\lambda\right) \left(v_L - \left(w_L - w_0\right)\right) - w_0, \tag{6}$$

with $\Pr_0 \left(v_H|\lambda\right) = q a_H$, and $\Pr_0 \left(v_L|\lambda\right) = (1-q) \left((1-\lambda) a_H + \lambda a_L\right)$.

The principal's problem reduces to the choice of effort levels a_H, a_L and payments w_H, w_L, w_0 that maximize her expected profits (6) subject to the incentive constraint (1), the participation constraint (4), and the limited liability constraints (5). Notice that the limited liability constraint on w_0 is binding. Thus, from now on, we set $w_0 = 0$. Moreover, the constraint on transfers (5) limits the principal's ability to punish the agent and implies, as shown in the Appendix, that the participation constraint (4) is slack.[14] Thus, we neglect it. Let us denote by P^λ the principal's programme for given λ.

Let us define the λ-first-best world as a setting where effort is observable and λ exogenous.

Solving program $\mathcal{P}^\lambda$, with the incentive constraints (2) and (3) binding, the principal attains her λ-second-best expected utility.

Proposition 1 shows that the accuracy of the agent's information is always valuable to the principal.

Proposition 1 *An increase in the probability that bad news will be remembered accurately has a positive effect on the principal's λ-second-best expected utility.*

The intuition behind this result is the following. As information becomes more precise, i.e. as λ increases, w_H decreases because the agent's expected benefit from exerting any level of effort when he recalls a good signal increases (see the expression for the incentive constraint (3) in the Appendix). This is a positive indirect effect. However, an increase in λ implies a negative direct

[14]Notice that moral hazard and limited liability make delegation costly to the principal.

effect on the probability of success following a bad signal $\Pr_0(v_L|\lambda)$, as the weight of a_L increases and that of a_H decreases. Since the positive indirect effect outweighs the negative direct effect, having the possibility of choosing the agent's type, the principal would prefer one with $\lambda = 1$.

We next proceed to show the existence of a conflict of interest between principal and agent over the recollection strategy.

In our companion paper, we have shown that whenever the weight attached to anticipatory utility s is large, the emotional gain from forgetting (see condition 11 in Immordino, Menichini and Romano, 2011) may induce the agent to forget a bad signal. In the present paper, we skip the details of the derivation and write the condition for $\lambda = 1$ ensuring that the agent has an incentive to remember (non-forgetfulness constraint):

$$c(a_H) - c(a_L) \geq s a_H (w_H - w_L) + (a_H - a_L)(w_L - w_0), \tag{7}$$

where $a_H \equiv a_H(1)$. This condition is obtained by comparing the expected utility of the agent when he observes a bad signal and decides to forget ($\lambda = 0$) or to recall it ($\lambda = 1$). Notice that we restrict the analysis to the pure strategy equilibria of the memory game ($\lambda \in \{0, 1\}$). This is without loss of generality because mixed strategy equilibria always coexist with the pure strategy ones (see Benabou and Tirole, 2003). Then, since we assume that when the agent is indifferent between recalling or forgetting, he chooses the action preferred by the principal, in case of multiplicity the pure strategy equilibrium with perfect recall ($\lambda = 1$) will prevail.

From the above constraint we conclude that the agent has an incentive to remember when the extra cost he incurs to exert effort a_H rather than a_L ($c(a_H) - c(a_L)$) exceeds the sum of the emotional gain from forgetting (the component of (7) $s a_H (w_H - w_L)$ due to the uncertainty about the true project return in case of success) plus the gain due to obtaining w_L rather than w_0 with an increased probability ($a_H - a_L$). It is clear that the incentive to recollect the bad signal is greater the lower the parameter capturing anticipatory utility s.

Let us denote by C^{SB} the second-best contract that solves programme $\mathcal{P}^{\lambda}$ when $\lambda = 1$.

The next proposition shows that agents with sufficiently low anticipatory emotions will prefer to recall bad news when offered contract C^{SB}.

Proposition 2 *There exists a threshold s^{SB} such that for all $s \leq s^{SB}$ the second-best contract C^{SB} satisfies the non-forgetfulness constraint (7) and implements the second-best level of effort.*

Taken together, Propositions (1) and (2) highlight a potential conflict in a setting where λ is endogenous: the principal always prefers perfect signal recollection, but the agent, if the weight placed on anticipatory utility is great, $s > s^{SB}$, could prefer to forget bad news (unless the contract is appropriately modified).

4 The optimal contract

In this section we explore the implications of violating condition (7) and derive the optimal contract. We thus assume that $s > s^{SB}$. In such cases, the second-best outcome cannot be achieved because contract C^{SB} fails to induce the agent to recollect his private information correctly (condition 7 violated). This gives rise to a third-best scenario, in which effort is unverifiable and the agent elects to forget bad news. In such circumstances, the principal's problem, which we denote by $\mathcal{P}^{TB}$, is to choose a vector of effort levels and a contract that solve programme $\mathcal{P}^{\lambda}$ above, with $\lambda = 1$, under the further non-forgetfulness constraint (7) that the agent is indifferent between forgetting and remembering bad news.

To derive our results we introduce the following technical assumption on the cost function, for instance satisfied for any power function, that guarantees a regularity condition on the non-forgetfulness constraint.

Assumption 1 *The ratio between average and marginal costs is always increasing or decreasing, that is:*

$$\frac{\partial}{\partial a}\left(\frac{c(a)/a}{c'(a)}\right) \underset{(or \geq 0)}{\leq 0} \quad \text{for any } a \text{ in } [0,1].$$

This assumption states that average costs grow always at a higher or lower rate than marginal costs.

Satisfaction of condition (7) produces two possible equilibria: a *separating* equilibrium, denoted by the superscript S, where $a_H^S > a_L^S$, and a *pooling* equilibrium, denoted by the superscript P, where $a_H^P = a_L^P$.

Proposition 3 *If Assumption 1 holds, then there exists a threshold $s^S > s^{SB}$ such that the separating equilibrium arises for all $s \in (s^{SB}, s^S]$ and the pooling equilibrium for all $s > s^S$.*

Thus, depending on the weight placed on anticipatory utility, we get the three possible scenarios depicted in Figure 2**.** When s is sufficiently small ($s \leq s^{SB}$), the agent recollects the signals correctly ($\lambda = 1$) and the principal designs a contract that rewards effort but not memory. Due to moral hazard, the principal achieves the second-best. For $s^{SB} < s \leq s^S$, the emotional impact of bad news may induce the agent to suppress it and instead "recall" good news. To induce accurate memory recollection, the principal has to design a separating contract that punishes forgetfulness and rewards memory. Last, for $s > s^S$, the principal removes any incentive to forget bad news by asking for the same level of effort regardless of the signal.

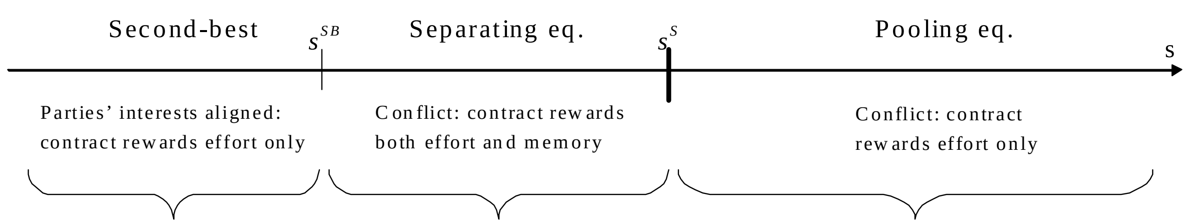

Fig. 2: Optimal contract as a function of s. The figure shows that a second-best, separating or pooling contract may arise depending on the weight placed on anticipatory utility.

The next proposition analyzes the effect on the principal's utility of a change in s.

Proposition 4 *The principal's expected utility is constant for $s \in \left[0, s^{SB}\right]$, decreasing for $s \in \left(s^{SB}, s^S\right]$ and constant for $s > s^S$.*

The result in Proposition 4 has implications for the type of agents the principal wants to hire. Having the possibility of selecting agents on the basis of the characteristics captured by the parameter s, she will hire any agent with $s \leq s^{SB}$. The issue is then to translate agent's preferences into observable personality characteristics. The taxonomy of traits with the broadest shared consensus endorses the so-called five-factor model (FFM henceforth). According to this model, the main personality traits can be summarized under five broad categories: extraversion, emotional stability, agreeableness, conscientiousness, and openness to experience. Among these, conscientiousness and emotional stability seem to be mostly related to our measure of anticipatory utility. Conscientiousness measures how dependable (trustworthy and reliable), organized, self-disciplined and efficient a person is, rather than unreliable, disorganized, self-indulgent, engaging in fantasy, daydreamer. Emotional stability is concerned with how calm, self-confident, self-conscious and cool an individual is rather than anxious, emotional and vulnerable.[15] Thus, people who are more emotional, anxious, daydreamer for self-assurance may be more prone to self-deception, i.e., to distort their assessments of the likelihood of future events (discard bad news). This leads them to expect good outcomes too often, or more often than they ought to, and thus to overoptimism.[16]

If we can establish a link between anticipatory utility and personality traits, the implication of this result is that the principal prefers agents relatively more conscientious and emotionally stable.

In the next section, we will use an explicit functional form to derive the properties of the optimal contract and perform comparative statics analysis.

5 The quadratic cost case

Assume that $c(a) = ca^2/2$, with $c \geq v_H$.[17] This functional form, which satisfies Assumption 1, allows us to interpret the results we have obtained so far, as well as to make further progress in

[15] Several studies have tested the validity of the FFM for predicting job performance, earnings and occupational choices, with conscientiousness and emotional stability having the most predictive power (Barrick and Mount, 1991; Salgado, 1997; Salgado and Rumbo, 1997; Barrick, Mount and Judge, 2001; Nyhus and Pons, 2005; Kaplan, Klebanov, and Sorensen, 2008; Ham, Junankar and Wells, 2009).

[16] In studying whether overoptimism may explain an excess of merger activity, Malmendier and Tate (2008) use, in addition to measures of optimism based on CEOs' personal portfolio decisions, a press-based indicator constructed by retrieving all leading business publications articles that characterize sample CEOs as "Confident" (confident, optimistic) versus "Cautious" (cautious, reliable, practical, conservative, frugal, steady, or negating one of the "Confident" terms). Interestingly, the terms negating optimism (the traits described by "Cautious") resemble closely the trait described by conscientiousness.

[17] This condition ensures interior solutions.

the characterization of the optimal contract. All the results presented in this section (and the following) are derived in Appendix B.

Solving problem P^{λ}, we derive the second-best contract $C^{SB} = \{0, v_L/2, v_H/2\}$ that implements the second-best level of effort, that for a quadratic cost function turns out to be $a_L^{SB} = v_L/2c$ and $a_H^{SB} = v_H/2c$ (see Appendix B for details). Using these payments in the non-forgetfulness constraint (7), we deduce that the agent chooses perfect signal recollection only if

$$s \leq s^{SB} \equiv \frac{v_H - v_L}{2v_H} \leq \frac{1}{2}. \tag{8}$$

As in the general case, the threshold s^{SB} is the greatest weight placed on anticipatory utility that makes the agent indifferent between recalling and forgetting bad news. Notice that since the ratio in (8) is less than 1/2, condition (7) is violated if $s > 1 - s$, that is, whenever the weight attached to anticipatory utility is greater than that of physical outcome. Finally, it is interesting that the importance of the signal is inversely related to the distance between the intermediate return v_L and the good return v_H. Indeed, the signal is worthless if $v_L = v_H$, while it is crucial when $v_L = 0$. As a consequence, the significance of the conflict of interest between principal and agent over the probability of recall λ also depends on the distance between the return for an extremely good project, v_H and for a business-as-usual result, v_L. In the following, we will interpret this distance as a measure of the riskiness of the project. For given v_H, as the intermediate outcome v_L increases, the agent's tendency to forget bad news increases (s^{SB} is decreasing in v_L).[18] Generally, as the distance between v_H and v_L decreases, the agent's tendency to forget bad news increases (s^{SB} increases in $\Delta v = v_H - v_L$).

We have seen above that when $s > s^{SB}$, the second-best contract C^{SB} cannot be achieved. This gives rise to a third-best scenario $\mathcal{P}^{TB}$ in which, with quadratic costs, the non-forgetfulness constraint (7) can now be written as:

$$(a_H - a_L)\left[(1-2s)\,a_H - a_L\right] = 0 \tag{9}$$

[18] If $v_H = v_L$, $s^{SB} = 0$, but this is a degenerate case since there is no memory problem. Instead, when $v_L = 0$, $s^{SB} = 1/2$, and the agent wants to forget bad news only if the weight of anticipatory utility is greater than that of the physical outcome, that is if $s > 1/2$.

where the repayments have been replaced by the incentive feasible payments (see the Appendix B for details).

Satisfaction of the previous equality produces two possible equilibria: a *separating* equilibrium, denoted by the superscript S, where $a_L = \phi(s) a_H$, with $\phi(s) \equiv (1-2s)$, which is possible only if $s \leq 1/2$, and a *pooling* equilibrium, denoted by the superscript P, where $a_H = a_L$.[19] Proposition 5 solves the principal's problem in these two equilibria.

Proposition 5 *In the separating equilibrium, the optimal levels of effort are*

$$\begin{aligned} a_H^S &= a_H^{SB} + \gamma(1-q)(1-2s) \\ a_L^S &= a_L^{SB} - \gamma q \end{aligned}$$

and are implemented by the following state-contingent rewards

$$w_0^S = 0,\ w_H^S = w_H^{SB} + c\gamma(1-2s)(1-q),\ w_L^S = w_L^{SB} - c\gamma q,$$

where $\gamma \equiv \frac{v_H\left(s-s^{SB}\right)}{c\left(q+(1-q)(1-2s)^2\right)}$ *is positive for any* $s \geq s^{SB}$.

In the pooling equilibrium, the optimal level of effort is given by

$$a_H = a_L = a^P = q a_H^{SB} + (1-q) a_L^{SB},$$

and is implemented by the following state-contingent rewards

$$w_0^P = 0, w_H = w_L = w^P = q w_H^{SB} + (1-q) w_L^{SB}.$$

The above proposition makes it clear that there are two opposite ways to elicit accurate recollection. In the separating equilibrium, which arises for all $s \in (s^{SB}, s^S]$, the principal increases the cost of forgetting, i.e., the left-hand side of condition (7), by asking for a level of effort higher than second-best when the agent recalls good news, lower when he recalls bad news. In order to implement those effort levels, which are farther apart,[20] an agent recollecting a

[19] If $s > 1/2$, the binding non-forgetfulness constraint (9) would imply $a_L < 0$.

[20] It is immediate to verify that $\Delta w^S = w_H^S - w_L^S > w_H^{SB} - w_L^{SB} = \Delta w^{SB}$ and $\Delta a^S = a_H^S - a_L^S > a_H^{SB} - a_L^{SB} = \Delta a^{SB}$.

good (bad) signal must be offered a reward higher (lower) than the second-best. In the pooling equilibrium, which arises for $s > s^S$, the principal eliminates any incentive to suppress bad news by offering a flat contract paying a constant amount w^P regardless of result $\{v_H, v_L\}$ and asking for the same level of effort a^P following both signals. As neither effort levels nor rewards depend on signal recollection, the agent is indifferent between recalling and forgetting bad news.[21]

Finally, there exists an outcome-equivalent equilibrium in which the principal prefers not to elicit recollection because it is too costly. In these cases the principal opts for an *accommodating* strategy that accepts the agent's forgetfulness ($\lambda = 0$) by neglecting constraint (7), so that the agent never recollects the bad signal. As a result, he always exerts high effort a_H^A, and a_L is off the equilibrium path. It turns out that $a_H^A = a^P$ and $w_H^A = w^P$, so that this is welfare-equivalent to the pooling equilibrium (see Appendix B for details).

The above results have implications not only on the design of the compensation contract for emotional agents, but also on the selection process, i.e., on which type of agent the principal would want to choose if she could. As already argued in commenting Proposition 4, the principal prefers agents relatively more conscientious and emotionally stable. The use of a specific functional form allows us to further argue that the preference over these features should be more pronounced for less risky firms/industries/occupations (see condition 8). Although seemingly counterintuitive, this can be rationalized considering that in our setup incentive contracts play the dual role of inducing effort and eliciting memory. To see this, recall the comparative statics results on s^{SB} (equation 8). When the distance between v_H and v_L increases, s^{SB} shifts rightward, leftward when it decreases. Thus when v_H and v_L are distant (high-risk firms), to induce effort the principal must offer high-powered incentive contracts that also alleviate the agent's memory problem. Thus, by means of a standard second-best contract, the principal manages to resolve even the memory problem of more emotional agents. When v_H and v_L are close to each other (low-risk firms), a low-powered incentive contract suffices to induce effort. But because s^{SB} is smaller, this may conflict with the memory problems of more emotional agents, calling for a high-

[21] It is interesting to observe that in the pooling equilibrium a^P is the average between a_H^{SB} and a_L^{SB}, and w^P is the average between w_H^{SB} and w_L^{SB}.

powered incentive contract, which results in a separating contract. Thus, despite the attenuated moral hazard, lower-risk firms may be faced with the problem of overoptimism. Since this is costly, they prefer more stable agents. Thus, principals with riskier projects can "afford" to employ more emotional agents since, by offering high-powered incentives to induce effort, they can also control their tendency to self-deception at no extra cost. Those with less risky projects, instead, have less difficulty in inducing effort, but are confronted with the problem of controlling the agent's optimism. Since this is costly, they prefer to resort to more stable agents.

6 The agent has private information

The analysis conducted so far has assumed that the parameter of anticipatory utility is public information. As argued in the introduction, this seems a plausible assumption if this parameter can be related to personality traits and learned through psychological testing. This conjecture is corroborated by the common practice in many institutions, industries and businesses of screening non-cognitive traits of would-be employees through psychological testing. However, even if these tests are common practice, perfect knowledge of s on the part of the principal could be difficult to achieve. We thus extend the analysis to the case in which the weight on s is the agent's private information and we look at the effects of this setup modification on the optimal contract. For tractability, we focus on quadratic cost functions.

If only the agent knows s, the principal can offer no contract (second-best, separating, or pooling) contingent on s. For any contract $C = \{w_0, w_L, w_H\}$, both the preferred recollection strategy and the level of effort chosen will depend on the agent's type s.

Let us denote by $\widehat{S}$ a subset of the agent's type such that all those with $s \in \widehat{S}$ prefer to recall their private information when offered the contract C. Then the principal's decision problem is to choose the set $\widehat{S}$, the level of effort and the contract C that maximize her expected profits (6), subject to the incentive constraints both for accurate recollection and forgetfulness (1), the participation constraints (4), the non-forgetfulness constraint (7) for all agents with $s \in \widehat{S}$, and the limited liability constraints (5).[22]

[22] As in the complete-information case, the limited liability constraint on w_0 is binding, so that $w_0 = 0$. This

Given the recalled signal $\widehat{\sigma}$ and the memory strategy $\lambda(s)$ equal to 1 if $s \in \widehat{S}$ and to 0 if $s \notin \widehat{S}$, a type s agent faced with the contract C chooses the level of effort $a_{\widehat{\sigma}}(\lambda(s))$ that solves problem (1). Under the quadratic cost assumption this gives a level of effort $a_L = w_L/c$ for a non-forgetful agent observing bad news, and $a_H(1) = w_H/c$ for one observing good news. On the other hand, the level of effort of an agent who chooses to forget bad news turns out to be $a_H(0) = qw_H/c + (1-q)\, w_L/c$ regardless of the observed signal. Notice that $a_H(1)$ and $a_H(0)$ follow from (3) with $\lambda = 1$ and $\lambda = 0$, respectively, and at equilibrium $a_H(0)$ is the average between $a_H(1)$ and a_L. Thus, the effort of a forgetful agent is higher than the effort of an agent with accurate memory if the private signal is bad, and it is lower if the private signal is good. Since effort affects the probability of success, when the project value in case of success is low, the most successful projects are those run by forgetful agents.

To simplify the analysis of the principal's decision problem, we assume that the weight s is distributed uniformly over $[0,1]$, and denote by $\widehat{s}$ the supremum of $\widehat{S}$. Then, Proposition 6 states our first result.

Proposition 6 *When s is the agent's private information, the optimal contract $C^{AI} \equiv \{w_0, w_L^{AI}, w_H^{AI}\}$ is such that all agents with $s \leq \widehat{s}$ will recall a bad signal, while those with $s > \widehat{s}$ will forget it, with $\widehat{s} \in (s^{SB}, 1/2]$.*

To grasp the intuition for $\widehat{s} \in (s^{SB}, 1/2]$, consider that the principal's profits are the average between profits produced by agents who choose to recall the bad signal and those produced by agents who choose to forget it. For the latter group, the principal prefers not to elicit information recollection, opting instead for an *accommodating* strategy that accepts the agent's forgetfulness by neglecting constraint (7) ($\lambda = 0$). In this setting, suppose there is a contract such that $\widehat{s} < s^{SB}$.The principal could do better by offering the second-best contract C^{SB}: this induces recollection from all agents up to s^{SB}, and maximizes the profits generated by those agents who choose to forget.[23] Thus $\widehat{s}$ cannot be lower than s^{SB}. In fact, there is another contract, C^{AI},

implies, as shown in Appendix B, that the participation constraints for both types of agents are satisfied and can be ignored.

[23]Neglecting the memory problem, the second best contract solves the moral hazard problem under limited liability.

that does better than C^{SB}, as it induces recollection from a set of agents strictly larger than s^{SB}. Indeed, from the previous section we know that for all $s^{SB} < s \leq 1/2$, the emotional impact of bad news may induce the agent to suppress it and recall good news instead. To induce accurate recollection, the principal has to design a costly separating contract that punishes forgetfulness and rewards memory. Thus, it may seem surprising that the optimal contract with imperfect information is such that the principal decides to induce recollection also from agents with $s > s^{SB}$. However, for those whose weight is slightly greater than s^{SB}, the extra cost of inducing recollection is small and the increase in profits obtained switching from accommodating to separating is large. In other words, the direct effect of an increase in the threshold $\widehat{s}$ is first-order while the indirect effect via the non-forgetful constraint is second-order.

The next proposition compares equilibrium rewards in the asymmetric information setting to the second-best payments.

Proposition 7 *If s is agent's private information, the distance between equilibrium rewards when the good and the bad signals are observed grows relative to the second-best.*

If we interpret private information as the situation in which firms do not rely on psychological testing, these results, along with those in the previous section, have interesting implications. In particular, for given riskiness, industries and businesses that commonly use psychological testing for recruitment, by comparison with those that do not, should be less prone to reality denial and take on more dependable and emotionally stable employees, offering them less high-powered incentive schemes. This is because under public information, having the possibility to choose the type of the agent, the principal will hire only less emotional ones (those with $s \leq s^{SB}$, as from Proposition 4). Conversely, when s is private information, the principal, being unable to screen agents, will offer the same contract to all types, thus hiring also forgetful agents (those with $s > \widehat{s}$, as from Proposition 6). Last, to reduce the set of forgetful agents, the principal designs more high-powered incentive schemes relative to the perfect information case (Proposition 7). This in turn suggests that high-powered incentive schemes in less risky sectors not relying on psychological tests may be driven by behavioral causes rather than by the need to control incentive problems.

7 Conclusion

We have modeled an employment contract between an optimistic agent and a realistic principal. After showing the existence of a potential conflict over memory strategy, we have shown that the agent's optimism may be affected by monetary incentives. More specifically, we have found that for sufficiently low anticipatory emotions, principal and agent's preferences over optimal recollection are perfectly aligned, so that the second-best contract C^{SB} that solves the moral hazard problem also satisfies a non-forgetfulness constraint. However, if the agent places a large weight on anticipatory utility, the second-best outcome cannot be achieved because contract C^{SB} fails to induce the agent to recall his private information correctly. This gives rise to a third-best world in which the principal must distort effort levels and payments to make the agent indifferent between forgetting and remembering bad news.

What happens in our setting if effort is verifiable but the signal is still private information? If payments are contingent on the outcome, so that a better outcome is associated with a higher payment, the agent will always have an incentive to forget bad news. To prevent this, the principal can offer a flat contract and obtain the first-best utility. In other words, not only the presence of an emotional agent, but also a second imperfection is required to make our analysis interesting.

Last, the paper has conjectured a link between behavioral characteristics and agent's compensations. This is a significant issue that warrants further investigation and suggests a potentially fruitful avenue for future research. As argued by James Heckman, "there are major challenges in integrating personality psychology and economics. Economists need to link the traits of psychology with the preferences, constraints and expectation mechanisms of economics" (Heckman, 2011, p.3).

A Appendix A

In the analysis to follow, the limited liability constraint on w_0 is always binding. Thus, throughout all the proofs, we set $w_0 = 0$.

Proof of Proposition 1. In order to solve problem $\mathcal{P}^{\lambda}$, we differentiate (1) wrt a_L and a_H, solve the incentive constraints for w_L and w_H, substitute $w_L(a_L)$ and $w_H(a_H)$ in the objective function and maximize with respect to a_L and a_H. By the strict concavity of the agent's objective function, a necessary and sufficient condition for the incentive constraint to be satisfied when $\widehat{\sigma} = L$ is

$$w_L = c'(a_L).$$

Instead, when $\widehat{\sigma} = H$, the necessary and sufficient condition is

$$(r(\lambda)w_H + (1 - r(\lambda))w_L) = c'(a_H).$$ [24]

Define $w_L(a_L) \equiv c'(a_L)$. Substituting it in the above equation, we obtain

$$w_H(a_H, a_L) = \frac{(1 - \lambda(1-q))c'(a_H) - (1-\lambda)(1-q)c'(a_L)}{q}. \tag{10}$$

Substituting $w_L(a_L)$ and $w_H(a_H, a_L)$ in (6) and rearranging, the objective function becomes:

$$\begin{aligned} E_0[\Pi(a_L, a_H)] \;=\; & q a_H v_H + (1-q)((1-\lambda)a_H + \lambda a_L) v_L + \\ & -(1 - \lambda(1-q)) a_H c'(a_H) - \lambda(1-q) a_L c'(a_L). \end{aligned} \tag{11}$$

Substituting $w_L(a_L)$ and $w_H(a_H, a_L)$ in the participation constraint

$$\begin{aligned} E_0[U(a_L, a_H)] &= q a_H w_H + (1-q)((1-\lambda)a_H + \lambda a_L) w_L + \\ & -((1-\lambda(1-q))c(a_H) + \lambda(1-q)c(a_L)) = \\ &= [1 - \lambda(1-q)] a_H \left[c'(a_H) - \frac{c(a_H)}{a_H}\right] + \lambda(1-q) a_L \left[c'(a_L) - \frac{c(a_L)}{a_L}\right], \end{aligned} \tag{12}$$

which, by the convexity of the cost function, is strictly positive and can be neglected.[25]

Differentiating (11) with respect to a_L and a_H gives the following necessary and sufficient conditions

$$\frac{\partial E_0[\Pi(a_L, a_H)]}{\partial a_L} = v_L - \left(c'\left(a_L^{SB}\right) + a_L^{SB} c''\left(a_L^{SB}\right)\right) = 0 \tag{13}$$

$$\frac{\partial E_0[\Pi(a_L, a_H)]}{\partial a_H} = \frac{q}{(1 - \lambda(1-q))}(v_H - v_L) + v_L - \left[c'\left(a_H^{SB}(\lambda)\right) + a_H^{SB}(\lambda) c''\left(a_H^{SB}(\lambda)\right)\right] = 0 \tag{14}$$

By the envelope theorem, the derivative of the principal's expected profits with respect to λ is

$$\frac{\partial E_0\left[\Pi\left(a_L^{SB}, a_H^{SB}(\lambda)\right)\right]}{\partial \lambda} = (1-q)\left[a_L^{SB}(v_L - c'(a_L^{SB})) - a_H^{SB}(\lambda)(v_L - c'(a_H^{SB}(\lambda)))\right],$$

that is positive only if

$$a_L^{SB}(v_L - c'(a_L^{SB})) > a_H^{SB}(\lambda)(v_L - c'(a_H^{SB}(\lambda))). \tag{15}$$

Define the function $f(a) \equiv a(v_L - c'(a))$, with first derivative given by $f'(a) = v_L - (c'(a) - ac''(a))$, and observe that

[24] Where $(r(\lambda)w_H + (1 - r(\lambda))w_L) \in [c'(0), c'(1)]$ for any r.

[25] This result will hold throughout all the proofs in which s is observable.

- $f'(a_L^{SB}) = 0$ by (13) and $f'(a_H^{SB}(\lambda)) < 0$ by (14),
- $f''(a) = -2c''(a) - ac'''(a) < 0$ for any a, since $c''(a) \geq 0$ and $c'''(a) \geq 0$ by assumption.

Hence, $f(a)$ is decreasing for any $a \in [a_L^{SB}, a_H^{SB}(\lambda)]$ and condition (15) is satisfied. □

Proof of Proposition 2. For any given contract C^{SB}, condition (7) becomes

$$c\left(a_H^{SB}\right) - c\left(a_L^{SB}\right) \geq s a_H^{SB}(w_H^{SB} - w_L^{SB}) + \left(a_H^{SB} - a_L^{SB}\right) w_L^{SB}, \tag{16}$$

where a_H^{SB} is the solution to the following equation

$$w_H^{SB} = c'\left(a_H^{SB}\right). \tag{17}$$

By using (2) and (17) in (16) and rearranging terms, gives

$$(1-s)a_H^{SB}(c'\left(a_H^{SB}\right) - c'\left(a_L^{SB}\right)) - \left[\left(a_H^{SB}c'\left(a_H^{SB}\right) - c\left(a_H^{SB}\right)\right) - \left(a_L^{SB}c'\left(a_L^{SB}\right) - c\left(a_L^{SB}\right)\right)\right] \geq 0 \tag{18}$$

If $s = 1$, (18) becomes

$$a_H^{SB}c'\left(a_H^{SB}\right) - c\left(a_H^{SB}\right) \leq a_L^{SB}c'\left(a_L^{SB}\right) - c\left(a_L^{SB}\right),$$

that is never true since $a_H^{SB} > a_L^{SB}$ and the function $h(a) = a\,c'(a) - c(a)$ is increasing in a ($h'(a) = a\,c''(a) > 0$ by assumption).

If $s = 0$, (18) becomes

$$a_H^{SB}c'\left(a_L^{SB}\right) - c\left(a_H^{SB}\right) \leq a_L^{SB}c'\left(a_L^{SB}\right) - c\left(a_L^{SB}\right),$$

that is always true since a_L^{SB} is the $\arg\max$ of the function $k(a) = a\,c'\left(a_L^{SB}\right) - c(a)$. Moreover, for any incentive feasible contract, the function

$$\Psi(s, a_H, a_L) \equiv (1-s)a_H(c'(a_H) - c'(a_L)) - [(a_H c'(a_H) - c(a_H)) - (a_L c'(a_L) - c(a_L))], \tag{19}$$

defined on $[0,1]^3$, is strictly decreasing in s for all $a_H > a_L$. As a consequence, there exists $s^{SB} \in (0,1)$, such that $\Psi(s^{SB}, a_H^{SB}, a_L^{SB}) = 0$ and $\Psi(s, a_H^{SB}, a_L^{SB}) > 0$ if and only if $s > s^{SB}$. □

Proof of Proposition 3. The proof will be developed in 4 steps. To show that if a separating equilibrium exists then it is unique we first prove that for each s larger than a threshold s^{SB} and for each a_H there exists at most one a_L smaller than a_H which strictly satisfies the non-forgetfulness constraint (19) (**Step 1**); second, we show that if s is large (larger than a threshold $\bar{s}$), then there is no positive a_L smaller than a_H which strictly satisfies the non-forgetfulness constraint (19) (**Step 2**); then, we state that if s is larger than $\bar{s}$ a pooling equilibrium prevails (**Step 3**); finally, we use the previous steps to conclude that there exists a threshold $s^S > s^{SB}$

such that the separating equilibrium arises for all $s \in (s^{SB}, s^{S}]$ and the pooling equilibrium for all $s > s^{S}$ (**Step 4**).

Step 1. *For any* (s,a_H) *in* $[0,1]^2$ *there exists at most one* $a_L(s,a_H) < a_H$ *such that* $\Psi(s,a_H,a_L(s,a_H)) = 0$, *and it is lower than* $(1-s)a_H$.

- The derivative of (19) with respect to a_L is $(a_L - (1-s)a_H)\, c''(a_L)$ which is non-negative if and only if $a_L \geq (1-s)a_H$.
- Moreover, (19) evaluated at $a_L = (1-s)\, a_H$ is:

$$\Psi(s,a_H,(1-s)a_H) = (1-s)a_H c'(a_H) - c((1-s)a_H) - (a_H c'(a_H) - c(a_H)),$$

  which is negative for all s and a_H. Indeed, by the convexity of $c(a)$

$$\begin{aligned}\Psi(s,a_H,(1-s)a_H) &< (1-s)(a_H c'(a_H) - c(a_H)) - (a_H c'(a_H) - c(a_H)) = \\ &= -s(a_H c'(a_H) - c(a_H)) < 0.\end{aligned}$$

The above amounts to say that the function $\Psi(s,a_H,\cdot)$ is first strictly decreasing and then increasing; and it is negative at its minimum. This implies that, for any s and a_H, the function $\Psi(s,a_H,\cdot)$ crosses the x-axes at most twice: once in the upward sloping part when $a_L = a_H$, since $\Psi(s,a_H,a_H) = 0$ for all s and a_H, and once in the downward sloping part if $\Psi(s,a_H,0) \leq 0$. Thus, if $a_L(s,a_H)$ exists, it is lower than $(1-s)a_H$. Finally, notice that since the previous result is true for $s \in [0,1]$, then it is a fortiori true for each s larger than a threshold s^{SB}.

Step 2. *There exists an* $\bar{s} \in \left[s^{SB}, 1\right]$ *such that for any* (s,a_H) *in* $[\bar{s},1] \times [0,1]$ *no* $a_L(s,a_H) > 0$ *exists.*

Let $a_L = 0$. Then $\Psi(s,a_H,0) = c(a_H) - s a_H c'(a_H)$ and for any a_H there exists a unique $s(a_H) = c(a_H)/a_H c'(a_H)$ such that $\Psi(s(a_H),a_H,0) = 0$.

By Assumption 1, $s(a_H)$ is a strictly monotone function of a_H. In particular, if $\frac{\partial}{\partial a}\left(\frac{c(a)/a}{c'(a)}\right) \geq 0$ then $s([0,1]) = [s(0), s(1)]$, whilst if $\frac{\partial}{\partial a}\left(\frac{c(a)/a}{c'(a)}\right) \leq 0$ then $s([0,1]) = [s(1), s(0)]$. Notice that $s(0)$ and $s(1)$ are both smaller than 1. Indeed, $s(0) = \lim_{a_H \to 0} c(a_H)/a_H c'(a_H) = 1/\underline{n}$ (by de l'Hôpital Theorem) with $\underline{n} = \inf\left\{n \in N : c^{(n)}(0) \neq 0\right\}$ higher than 2 since $c'(0) = 0$ by assumption, while $s(1) = c(1)/c'(1) < 1$ by the convexity of $c(\cdot)$.

Next we will argue that if $\frac{\partial}{\partial a}\left(\frac{c(a)/a}{c'(a)}\right) \geq 0$ then $\bar{s} = s(1)$. The following four steps will help to prove the result:

1. For all $a_L < (1-s)a_H$ and positive, the function $\Psi(s,a_H,\cdot)$ is decreasing and, then, $\Psi(s(a_H),a_H,a_L) < 0$;

2. The function $s(\cdot)$ is strictly increasing and, then, for all $a_H > a'_H$, we have that $s(a_H) > s(a'_H)$;

3. For any a_H the function $\Psi(s, a_H, 0)$ is decreasing in s. Combining this with the result in point 2, gives $0 = \Psi(s(a'_H), a'_H, 0) > \Psi(s(a_H), a'_H, 0)$ for all $a_H > a'_H$;

4. Recalling that $\partial\Psi(s, a_H, \cdot)/\partial a_L < 0$ for all $a_L < (1 - s)a_H$, $\Psi(s(a_H), a'_H, 0) > \Psi(s(a_H), a'_H, a_L)$.

Combining points 3 and 4, gives $\Psi(s, a_H, a_L) < 0$ for all a_H and for all $s > s(a_H)$. Since $\max_{a_H \in [0,1]} s(a_H) = s(1) < 1$, this implies that for any (s,a_H) in $[s(1), 1] \times [0, 1]$ no $a_L(s, a_H) > 0$ exists.

Finally, to show that $s^{SB} < s(1)$ suppose, by contradiction, that $s^{SB} \geq s(1)$. By definition, s^{SB} is such that $\Psi(s^{SB}, a_H^{SB}, a_L^{SB}) = 0$ and, since $\Psi(s, a_H, \cdot)$ is decreasing, $\Psi(s^{SB}, a_H^{SB}, 0) > 0$. But this contradicts the assumption that $s^{SB} \geq s(1)$. This concludes the proof of the claim for the case $\frac{\partial}{\partial a}\left(\frac{c(a)/a}{c'(a)}\right) \geq 0$. The proof for the case $\frac{\partial}{\partial a}\left(\frac{c(a)/a}{c'(a)}\right) \leq 0$ is similar (except for the fact that $\bar{s} = s(0)$) and will be omitted.

Step 3. *For any $s \in [\bar{s}, 1]$ the pooling equilibrium arises.*

From Step 2 it follows that when $s \in (\bar{s}, 1]$ there is no contract that induces positive effort levels $a_L < a_H$ and satisfies the non forgetful constraint. Thus, the only incentive feasible contract is the pooling one. Consider then $s = \bar{s}$. If $\frac{\partial}{\partial a}\left(\frac{c(a)/a}{c'(a)}\right) \geq 0$, $\bar{s} = s(1)$ and the unique separating equilibrium candidate is $a_L = 0$ and $a_H = 1$. However, this allocation is dominated by the pooling equilibrium candidate, which always satisfies the non forgetful constraint. Indeed, by assumption $c'(1) \geq v_H$, then the incentive feasible contract when the $a_L = 0$ and $a_H = 1$ gives negative utility to the principal. On the other hand, if $\frac{\partial}{\partial a}\left(\frac{c(a)/a}{c'(a)}\right) \leq 0$, $\bar{s} = s(0)$ and the unique separating equilibrium candidate is $a_L = 0$ and $a_L = 0$, that is again dominated by the pooling equilibrium candidate since it gives zero utility to the principal. We then conclude that for any $s \in [\bar{s}, 1]$ the pooling equilibrium arises.

Step 4. *There exists a threshold $s^S \in \left(s^{SB}, 1/2\right]$ such that the separating equilibrium arises for all $s \in (s^{SB}, s^S]$ and the pooling equilibrium for all $s > s^S$.*

Notice that by definition of s^{SB} and $a_L(s, a_H)$, $a_L^{SB} = a_L(s^{SB}, a_H^{SB})$ and, then, $E_0\left[\Pi^{SB}\right] = E_0\left[\Pi^S\left(s^{SB}\right)\right] > E_0\left[\Pi^P\right]$. By continuity, for all $s = s^{SB} + \varepsilon$, with $\varepsilon > 0$ and small enough, the separating contract exists and dominates the pooling one. Similarly, into a neighborhood of $\bar{s}$ the pooling contract dominates any incentive feasible separating contract (by Step 3 and by continuity arguments). Thus, a threshold s^S does exist.

Suppose, by way of obtaining a contradiction, that it is not unique and that there exists a pair of effort levels a'_L and a'_H and a weight $s' > s^S$ such that $\Psi(s', a'_H, a'_L) \geq 0$, $a'_L < a'_H$ and

$E_0\left[\Pi^S\left(a'_L, a'_H\right)\right] \geq E_0\left[\Pi^P\right]$. Since $\Psi(s, a_H, a_L)$ is decreasing in s, then $\Psi(s, a'_H, a'_L) > 0$ for all $s < s'$. But this contradicts the assumption that $s' > s^S$ and concludes the proof. $\square$

Proof of Proposition 4. The principal's expected profit is equal to the second best expected profit $E_0\left[\Pi^{SB}\right] \equiv E_0\left[\Pi(a_L^{SB}, a_H^{SB})\right]$ for all $s \leq s^{SB}$, it is equal to the separating expected profit $E_0\left[\Pi^S(s)\right] \equiv E_0\left[\Pi(a_L(s, a_H^S), a_H^S)\right]$ for all $s \in (s^{SB}, s^S]$ and it is equal to the pooling expected profit $E_0\left[\Pi^P\right] \equiv E_0\left[\Pi(a_L^P, a_H^P)\right]$ for all $s > s^S$. Observe that:

- both the second best contract and the pooling contract do not depend on s, then $\partial E_0\left[\Pi^{SB}\right]/\partial s = 0$ for each $s \leq s^{SB}$ and $\partial E_0\left[\Pi^P\right]/\partial s = 0$ for each $s > s^S$;
- for each $s \in (s^{SB}, s^S]$: $\partial E_0\left[\Pi^S(s)\right]/\partial s = \partial E_0\left[\Pi^S(a_L(s, a_H^S), a_H^S)\right]/\partial a_L \cdot \partial a_L(s, a_H^S)/\partial s \leq 0$, since $\partial E_0\left[\Pi^S(a_L(s, a_H^S), a_H^S)\right]/\partial a_L > 0$ ($a_L^S < a_L^{SB}$) and $\partial a_L(s, a_H^S)/\partial s = a_H(c'(a_H) - c'\left(a_L(s, a_H^S)\right))/c''\left(a_L(s, a_H^S)\right)\left(a_L(s, a_H^S) - (1-s)a_H^S\right) \leq 0$ ($a_L(s, a_H^S) < (1-s)a_H^S$).

The proof is completed noting that $E_0\left[\Pi^{SB}\right] = E_0\left[\Pi^S(s^{SB})\right]$ and $E_0\left[\Pi^S(s^S)\right] \geq E_0\left[\Pi^P\right]$. $\square$

B Appendix B: The case with quadratic costs

In order to compute the λ-second-best levels of effort and rewards, we solve problem $\mathcal{P}^\lambda$ under the assumption of quadratic costs. We solve (2) and (3) for w_L and w_H. Substituting into the objective function (6), differentiating with respect to a_H and a_L, and solving gives

$$a_H^{SB}(\lambda) = \frac{r(\lambda)\, v_H + (1 - r(\lambda))\, v_L}{2c} \tag{20}$$

$$a_L^{SB} = \frac{v_L}{2c}, \tag{21}$$

where v_L and $r(\lambda)\, v_H + (1 - r(\lambda))\, v_L$ are the gains from effort with a good signal and a bad signal, respectively. Since the marginal benefit of effort is greater when the agent recollects a good signal than a bad one (i.e., $r(\lambda)\, v_H + (1 - r(\lambda))\, v_L \geq v_L$), then in the former case the level of effort implemented by the principal is higher. Moreover, the marginal benefit of effort when good news is recollected increases with the accuracy of information. As a consequence, $a_H^{SB}(\lambda)$ increases with λ. Notice also that the second-best level of effort a_H^{SB} in Section 5 follows from (20) for $\lambda = 1$. Using (20) and (21) in the incentive constraints we get

$$w_H^{SB} = \frac{v_H}{2} \tag{22}$$

$$w_L^{SB} = \frac{v_L}{2}. \tag{23}$$

Interestingly, the λ-second-best rewards are independent of λ. Moreover, it is easy to verify that the participation constraint (4) is satisfied by these second-best payments and efforts. Incentives to induce a λ-type agent to exert effort when good news is recollected depend on λ through the expected rewards in the case of success, not through each state-contingent payment.

Using the λ-second-best payments and effort levels in the principal's objective function (6) and differentiating with respect to λ, in keeping with Proposition 1, we find that the principal always prefers the agent to remember a bad signal

$$\frac{\partial E_0 \Pi^{\lambda SB}(\lambda)}{\partial \lambda} = \frac{q^2 (1-q)(v_H - v_L)^2}{4c(1-\lambda(1-q))^2} \geq 0.$$

Moreover, using the second-best payments (22), (23) in constraint (7) and solving for s, we find that the agent would choose $\lambda = 1$ only if condition (8) is satisfied.

Finally, the principal's second-best expected utility (with $\lambda = 1$) is given by

$$E_0 \Pi^{SB} = q\frac{v_H^2}{4c} + (1-q)\frac{v_L^2}{4c}. \tag{24}$$

Proof of Proposition 5. In order to induce the agent to recall the signal, the principal offers a contract that satisfies the non-forgetfulness constraint (7) with equality. Substituting (2) ($w_L = ca_L$), and (3) with $\lambda = 1$ ($w_H = ca_H$) in (7), the non-forgetfulness constraint simplifies to (9) ($(a_H - a_L)[(1-2s)a_H - a_L] = 0$). Satisfaction of the previous equality for $a_H \neq a_L$ gives rise to $a_L(a_H) = \phi a_H$, where $\phi \equiv 1 - 2s$ to simplify notation. In order to solve the principal's problem in the separating scenario, we substitute $a_L(a_H)$ in the incentive constraint (2) to obtain $w_L(a_H)$. We substitute $w_L(a_H)$, $w_H(a_H)$ and $a_L(a_H)$ in the objective function, which gives:

$$E_0[\Pi(a_H)] = qa_H[v_H - ca_H] + (1-q)\phi a_H[v_L - c\phi a_H]. \tag{25}$$

Differentiating with respect to a_H gives the following necessary and sufficient condition

$$\frac{\partial E_0[\Pi(a_H)]}{\partial a_H} = q(v_H - 2ca_H) + (1-q)\phi(v_L - 2c\phi a_H) = 0 \tag{26}$$

Solving (26) with respect to a_H gives

$$a_H^S = \frac{qv_H + (1-q)\phi v_L}{2c(q + (1-q)\phi^2)}, \tag{27}$$

that is lower than 1 since $c > v_H > v_L$ by assumption. Substituting (27) in $a_L(a_H)$ in (3) when $\lambda = 1$, and in (2) and rearranging terms we obtain the effort levels and payments in the Proposition.

In order to solve the principal's problem in the pooling scenario, we impose $a_L = a_H$. Substituting $a_L = a_H = a$ in (3) and in (2), we have

$$w_H(a) = w_L(a) = ca. \tag{28}$$

Substituting $\lambda = 1$, $a_L = a_H = a$, and (28) in (6) and rearranging terms, the objective function becomes

$$E_0[\Pi(a)] = qav_H + (1-q)av_L - ca^2. \tag{29}$$

Differentiating with respect to a gives the following necessary and sufficient condition:

$$\frac{\partial E_0[\Pi(a)]}{\partial a} = qv_H + (1-q)v_L - 2ca = 0 \tag{30}$$

Solving (30) for a and considering that $a_L^P = a_H^P = a$, we obtain

$$a_L^P = a_H^P = \frac{qv_H + (1-q)v_L}{2c}. \tag{31}$$

Substituting (31) in (28) and rearranging terms we obtain the effort level and payments in the proposition. □

We now verify that there exists $s^S \in (s^{SB}, 1/2]$ such that the separating equilibrium arises for all $s \in (s^{SB}, s^S]$, the pooling equilibrium for all $s > s^S$. To this aim, we work out the expected utilities and show that, for all $s > s^S$, $E_0\Pi^P > E_0\Pi^S$. Substituting out a_H^P in (29) and a_H^S in (25) we obtain the principal's expected profit in the pooling and separating equilibrium respectively, i.e.,

$$E_0\Pi^P = E_0\Pi^{SB} - \frac{q(1-q)(v_H - v_L)^2}{4c}. \tag{32}$$

$$\begin{aligned} E_0\Pi^S = {} & \frac{qv_H + (1-q)\phi v_L}{2c(q + (1-q)\phi^2)} * \\ & * \left\{ q\left[v_H - \frac{qv_H + (1-q)\phi v_L}{2(q + (1-q)\phi^2)}\right] + (1-q)\phi\left[v_L - \phi\frac{qv_H + (1-q)\phi v_L}{2(q + (1-q)\phi^2)}\right]\right\}. \end{aligned} \tag{33}$$

By equating $E_0\Pi^S$ and $E_0\Pi^P$ and solving for s, after some tedious algebra we obtain

$$\tilde{s} = 1 - \frac{v_H v_L}{q(v_H - v_L)^2 + v_L(2v_H - v_L)}.$$

Recalling that a separating equilibrium arises only for $s \leq 1/2$ (otherwise $a_L^S < 0$) and noticing that $s_2 > 1/2$, we conclude that the separating equilibrium arises for all $s \in (s^{SB}, 1/2]$, whilst the pooling equilibrium arises for all $s > 1/2$. □

Proof of the accommodating equilibrium.

In the accommodating scenario, the principal's problem $\mathcal{P}^A$ is to choose the levels of effort, a_L^A and a_H^A, and the contract, $C^A = \{0, w_L, w_H\}$, that maximize their expected profits (6), subject to the limited liability constraints and the incentive constraints, given $\lambda = 0$. From (2), $w_L = ca_L$. From (3) and $\lambda = 0$, we obtain

$$qw_H + (1-q)w_L = ca_H. \tag{34}$$

Substituting out $\lambda = 0$ and (34) in (6), we have

$$\begin{aligned} E_0\left[\Pi\left(a_H\right)\right] &= a_H\left[q\left(v_H - w_H\right) + (1-q)\left(v_L - w_L\right)\right] \\ &= a_H\left[\left(qv_H + (1-q)\,v_L\right) - c\,a_H\right] \end{aligned} \tag{35}$$

which does not depend on a_L. Indeed, in the accommodating scenario, the agent never recollects the bad signal, hence a_L is out of the equilibrium path. Differentiation of (35) with respect to a_H gives the following necessary and sufficient condition

$$\frac{\partial E_0\left[\Pi\left(a_H\right)\right]}{\partial a_H} = \left(qv_H + (1-q)\,v_L\right) - 2ca_H = 0 \tag{36}$$

whence, solving for a_H :

$$a_H^A = \frac{\left(qv_H + (1-q)\,v_L\right)}{2c}. \tag{37}$$

Substituting out (37) in (34), we have

$$qw_H^A + (1-q)w_L^A = \frac{\left(qv_H + (1-q)\,v_L\right)}{2}. \tag{38}$$

Hence, the accommodating equilibrium is characterized by any contract $C^A = \{0, w_L^A, w_H^A\}$, such that $qw_H^A + (1-q)w_L^A$ satisfies (38), and by the levels of effort a_H^A given by (37) and a_L^A given by (2).

Finally, since $a_H^A = a^P$ and $qw_H^A + (1-q)w_L^A = w^P$ (see (34) and (28)), the accommodating equilibrium is welfare equivalent to the pooling equilibrium. $\square$

The principal's expected utility is weakly decreasing in s.

Observe that:

1. from (24), $\partial E_0\left[\Pi^{SB}\right]/\partial s = 0$ for each $s \in [0, s^{SB})$;
2. from (33), $\partial E_0\left[\Pi^{S}\right]/\partial s = -2\,(1-q)\,a_H^S\left[v_L - 2ca_H^S\phi\right] < 0$ for each $s \in [s^{SB}, 1/2)$, since

$$v_L - 2ca_H^S\phi \geq 0 \iff v_L \geq -2w_L^S \iff v_L \geq v_L - 2\gamma cq,$$

   that is true for all $s \geq s^{SB}$;
3. from (32), $\partial E_0\left[\Pi^{P}\right]/\partial s = 0$ for each $s \in [s^S, 1]$.

The proof is completed noting that $E_0\left[\Pi^{SB}\right] = E_0\left[\Pi^S(s^{SB})\right]$ and $E_0\left[\Pi^S(s^S)\right] \geq E_0\left[\Pi^P\right]$. $\square$

Proof of Proposition 6. The contract has to be such that all agents with $s \in \widehat{S}$ prefer to recall the signal, which is ensured if we impose the non-forgetfulness constraint (7) for all $s \in \widehat{S}$. But this is equivalent to $a_L \leq \phi(s)\, a_H(1)$ for all $s \in \widehat{S}$, where $\phi(s) \equiv (1-2s)$. Then, noticing that $\phi(s)$ is decreasing in s, this condition is clearly satisfied for all $s \in \widehat{S}$ if and only if

$$a_L \leq \phi(\widehat{s})a_H(1). \tag{39}$$

Now, we demonstrate that $\widehat{S}$ is an interval. Suppose, by contradiction, that there exists an agent with $s = s' < \widehat{s}$ which prefers to forget bad news. Since the contract offered by the principal is the same for all agents, from the incentive constraints we know that the effort chosen by each agent does not depend on his type. Then, an agent with $s = s'$ prefers to forget bad news only if $a_L \geq \phi(s')a_H(1)$, which is possible only if $s' \geq \widehat{s}$ since $a_L \leq \phi(\widehat{s})a_H(1)$ and ϕ is decreasing. This contradicts our assumption and implies that $\widehat{S} = [0, \widehat{s}]$.

Next, we show that all contracts that satisfy incentive constraints satisfy also the participation constraints. The participation constraint of an agent with $s \in \widehat{S}$ is

$$q\left(a_H(1)w_H - \frac{ca_H^2(1)}{2}\right) + (1-q)\left(a_L w_L - \frac{ca_L^2}{2}\right).$$

Substituting $w_L = ca_L$ and $w_H = ca_H(1)$ it becomes

$$q\frac{ca_H^2(1)}{2} + (1-q)\frac{ca_L^2}{2},$$

which is always positive. Similarly, by substituting $a_H(0) = qa_H(1) + (1-q)\,a_L$, $w_L = ca_L$ and $w_H = ca_H(1)$ in the participation constraint of an agent with $s \notin \widehat{S}$ gives

$$a_H(0)\left(qw_H + (1-q)\,w_L\right) - \frac{ca_H^2(0)}{2} = \frac{c}{2}\left(qa_H(1) + (1-q)\,a_L\right),$$

that is always positive.

Since $\widehat{S}$ is an interval, the principal's expected profits can be written as

$$\begin{aligned} &\textstyle\int_0^{\widehat{s}} [q\, a_H(1)\,(v_H - w_H) + (1-q)\, a_L(v_L - w_L)]ds + \\ &+ \textstyle\int_{\widehat{s}}^1 a_H(0)\,[q\ (v_H - w_H) + (1-q)\,(v_L - w_L)]ds. \end{aligned} \tag{40}$$

Substituting incentive constraints into (40) and rearranging terms we get

$$\begin{aligned} E_0\left[\Pi\left(a_H(1),\, a_L,\, \widehat{s}\right)\right] &= [a_H(1)q(v_H - ca_H(1)) + a_L(1-q)(v_L - ca_L)]\widehat{s} + \\ &+ [a_H(1)q + a_L(1-q))][q(v_H - ca_H(1)) + (1-q)(v_L - ca_L)](1-\widehat{s}). \end{aligned} \tag{41}$$

Thus, the principal's problem simplifies to choosing $a_H(1)$, a_L and $\widehat{s}$ that maximize (41), subject to (39).

Next, we show that constraint (39) is binding at equilibrium. Suppose, by contradiction, that this is not true. If (39) is not binding, the principal's expected profit (41) is linear in s and

$$\frac{\partial E_0\left[\Pi\left(a_H(1),\, a_L,\, \widehat{s}\right)\right]}{\partial \widehat{s}} = q(1-q)(a_H(1) - a_L[(v_H - ca_H(1)) - (v_L - ca_L)]. \tag{42}$$

The optimal $\widehat{s}$ is 1 if (42) is positive and 0 otherwise. However, $\widehat{s} = 1$ is not possible since a_L cannot be negative and constraint (39) would require $a_L \leq \phi(1)a_H(1) < 0$. On the other hand, if $\widehat{s} = 0$, the first order conditions on $a_H(1)$ and a_L would imply $a_H(0) = \left(qv_H + (1-q)\,v_L\right)/2c$. It is easy to verify that the second best level of efforts which satisfy the first order conditions are such that (42) is positive. Thus, $\widehat{s} \in (0, 1/2]$ and constraint (39) is binding at equilibrium.

In the following we show that $\widehat{s} > s^{SB}$. Substituting (39) in (41) and rearranging terms gives

$$\begin{aligned} E_0\left[\Pi\left(a_H(1), \widehat{s}\right)\right] &= a_H(1)[q(v_H - ca_H) + (1-q)\phi(\widehat{s})(v_L - c\phi(\widehat{s})a_H(1))]\widehat{s} + \\ &+ a_H(1)[q + (1-q)\phi(\widehat{s}))[q(v_H - ca_H(1)) + (1-q)(v_L - c\phi(\widehat{s})a_H(1))](1-\widehat{s}). \end{aligned} \tag{43}$$

The principal's problem simplifies to the choice of $a_H(1) \in [0,\, 1]$ and $\widehat{s} \in (0,\, 1/2]$ that maximize (43). Differentiating (43) with respect to $a_H(1)$ gives the following necessary and sufficient condition for an interior solution

$$\begin{aligned} \frac{\partial E_0\left[\Pi\left(a_H(1), \widehat{s}\right)\right]}{\partial a_H(1)} &= -2c\left[4\,(1-q)\,\widehat{s}\left(q\widehat{s}^2 + (1-q)\widehat{s} - 1\right) + 1\right]a_H\,(1) \\ +2\widehat{s}\,(1-q)\left[\widehat{s}q(v_H - v_L) - (qv_H + (1-q)\,v_L)\right] &+ qv_H + (1-q)v_L = 0 \end{aligned}$$

Solving for $a_H(1)$, we obtain

$$a_H(1)(\widehat{s}) = \frac{2\widehat{s}^2 q\,(1-q)\,(v_H - v_L) + (qv_H + (1-q)v_L)\,(1 - 2\widehat{s}\,(1-q))}{2c\left[4\,(1-q)\,q\widehat{s}^3 + \left(1 - 2\,(1-q)\,\widehat{s}\right)^2\right]} \tag{44}$$

Substituting (44) in (43), differentiating $E_0\left[\Pi\left(a_H(1)(\widehat{s}),\, \widehat{s}\right)\right]$ with respect to $\widehat{s}$ and rearranging terms gives the following necessary and sufficient condition for an interior solution

$$2(1-q)\left[2c(1 - 3s^2q - 2s(1-q))a_H(1)(\widehat{s}) - q(v_H - v_L)(1-2s) - v_L\right]a_H(1)(\widehat{s}) = 0. \tag{45}$$

Since at equilibrium $a_H(1)(\widehat{s}) > 0$, the first order condition (45) for an interior solution reduces to $\eta(\widehat{s}) \equiv [2c(1 - 3s^2q - 2s(1-q))a_H(1)(\widehat{s}) + 2qs(v_H - v_L) - (qv_H + (1-q)v_L)] = 0$. By substituting (44) in $\eta(\widehat{s})$ and rearranging terms gives:

$$\eta(\widehat{s}) = \frac{qs\varphi(\widehat{s})}{\left[4\,(1-q)\,q\widehat{s}^3 + \left(1 - 2\,(1-q)\,\widehat{s}\right)^2\right]}$$

where $\varphi(\widehat{s}) \equiv 2((q(1-q)\widehat{s}^3+1))(v_H-v_L)-((1-q)(v_H-v_L)+v_H)(3-2(1-q)\widehat{s})\widehat{s}$. Since at equilibrium $\widehat{s}>0$, the first order condition (45) for an interior solution reduces to $\varphi(\widehat{s})=0$.

Observe that

$$\varphi'(\widehat{s}) = 6q(1-q)(v_H-v_L)s^2+(4(1-q)s-3)\left((2-q)v_H-(1-q)\,v_L\right) \geq 0$$

for all $\widehat{s} \in [0,1/2]$. Indeed, $\varphi''(\widehat{s}) = 4(1-q)(3q(v_H-v_L)s+((2-q)v_H-(1-q)\,v_L)) > 0$ for all $s>0$, so that $\varphi'(\widehat{s})$ is increasing for positive s, and $\varphi'(\widehat{s}=1/2) = 1/2((v_H-v_L)q-3v_H-v_L)q-2v_H+v_L<0$.

Hence, the function $\varphi(\widehat{s})$ is decreasing for all $\widehat{s} \in [s^{SB},\, 1/2]$. Moreover,

$$\varphi(s^{SB}) = \frac{(v_H-v_L)}{4v_H^3}(q\left(v_H^2-v_L^2\right)\left((q\left(v_H-v_L\right)+\left(v_H+v_L\right)\right)+2v_Hv_L^2) > 0$$

and

$$\varphi(\widehat{s}=1/2) = (\frac{1}{4}q^2(v_H-v_L)+\frac{1}{4}q(v_H-6v_L)-v_L),$$

which has ambiguous sign depending on q, v_H and v_L (negative if q is low and positive if q is large and v_L small with respect to v_H). Since the derivative of $E_0\left[\Pi\left(a_H(1)(\widehat{s}),\, \widehat{s}\right)\right]$ with respect to $\widehat{s}$ is positive when $\widehat{s}=s^{SB}$, then the optimal $\widehat{s}$ is larger than s^{SB}, and it is equal to 1/2 if $\varphi(\widehat{s}=1/2)>0$ and lower than 1/2 otherwise.□

Proof of Proposition 7. Let us define the distance between equilibrium rewards when the good and the bad signals are privately observed by the agents' as $\Delta w(\widehat{s}) \equiv w_H(1)(\widehat{s})(1-\phi(\widehat{s}))$ and the distance between second best rewards as $\Delta w^{SB} \equiv \frac{v_H}{2}(1-\frac{v_H}{v_L})$. In order to show that $\Delta w(\widehat{s})$ is always larger than Δw^{SB} we will prove the following claims.

Claim 1: $\inf_{\widehat{s}\in(s^{SB},1/2]} w_H(1)(\widehat{s}) \geq \frac{v_H}{2}$ *for all* $\frac{v_L}{v_H} \in [0,1]$.

The claim is immediately proved by noticing that $\partial w_H(1)(\widehat{s})/\partial\widehat{s} \geq 0$ for all $\widehat{s} \in (s^{SB},\, 1/2]$ and, then, $\inf_{\widehat{s}\in(s^{SB},1/2]} w_H(1)(\widehat{s}) = w_H(1)(s^{SB}) = \frac{v_H}{2}$.

Claim 2: $\inf_{\widehat{s}\in(s^{SB},1/2]}(1-\phi(\widehat{s})) \geq (1-\frac{v_L}{v_H})$ *for all* $\frac{v_L}{v_H} \in [0,1]$.

Observe that: $\inf_{\widehat{s}\in(s^{SB},1/2]}(1-\phi(\widehat{s})) \geq (1-\frac{v_L}{v_H})$ iff $\sup_{\widehat{s}\in(s^{SB},1/2]} \phi(\widehat{s}) \leq \frac{v_L}{v_H}$. Since $\partial\phi(\widehat{s})/\partial\widehat{s} \leq 0$ for all $\widehat{s} \in (s^{SB},\, 1/2]$, then $\sup_{\widehat{s}\in(s^{SB},1/2]} \phi(\widehat{s}) = \phi(s^{SB}) = \frac{v_L}{v_H}$. □

References

[1] Akerlof, George A., and William T. Dickens, 1982, The Economic Consequences of Cognitive Dissonance, *American Economic Review*, 72, 307–319.

[2] Almlund, Mathilde, Angela Lee Duckworth, James J. Heckman, Tim D. Kautz, 2011, Personality Psychology and Economics, NBER Working Paper 16822.

[3] Bénabou, Roland, 2009, Groupthink: Collective Delusions in Organizations and Markets, CEPR Discussion Paper n. 7193.

[4] Bénabou, Roland, and Jean Tirole, 2002, Self-Confidence and Personal Motivation, *Quarterly Journal of Economics*, 117, 871–915.

[5] Bénabou, Roland, and Jean Tirole, 2003, Intrinsic and Extrinsic Motivation, *Review of Economic Studies*, 70, 489–520.

[6] Barrick, Murray R., and Michael K. Mount, 1991, The Big Five Personality Dimensions and Job Performance: A Meta-Analysis, *Personnel Psychology*, 44 (1), 1-26.

[7] Barrick, Murray R., Michael K. Mount, and Timothy A. Judge, 2001, Personality and Performance at the Beginning of the New Millennium: What do We Know and Where Do We Go Next?, *International Journal of Selection and Assessment*, 9 (1), 9-30.

[8] Bernheim, B. Douglas, and Raphael Thomadsen, 2005, Memory and Anticipation, *The Economic Journal*, 115, Issue 503, 271–304.

[9] Bowles, Samuel, Herbert Gintis, Melissa Osborne, 2001, The Determinants of Earnings: A Behavioral Approach, *Journal of Economic Literature*, 39 (4), 1137-1176.

[10] Borghans, Lex, Angela Duckworth, James Heckman, and Bas ter Weel, 2008, The Economics and Psychology of Personality Traits, *Journal of Human Resources*, 43 (4), 974–1061.

[11] Brunnermeier, Markus, and Jonathan Parker, 2005, Optimal Expectations, *American Economic Review*, 90, 1092-1118.

[12] Caplin, Andrew, and John Leahy, 2001, Psychological Expected Utility Theory and Anticipatory Feelings, *Quarterly Journal of Economics*, 116, 55–79.

[13] Carrillo, Juan, and Thomas Mariotti, 2000, Strategic Ignorance as a Self-Disciplining Device, *Review of Economic Studies*, 66 , 529–544.

[14] Crutzen, Benoît S. Y. , Otto H. Swank and Bauke Visser, 2010, Confidence management: on interpersonal comparisons in teams, mimeo.

[15] Daly, Michael, Liam Delaney, and Colm P. Harmon, 2009, Psychological and Biological Foundations of Time Preferences, *Journal of the European Economic Association*, 7 (2-3), 659–669.

[16] Eil, David and Justin Rao (2011), "The Good News-Bad News Effect: Asymmetric Processing of Objective Information about Yourself," *American Economic Journal: Microeconomics*, 3(2), 114-138.

[17] Fang, Hanming, and Giuseppe Moscarini, 2005, Morale Hazard, *Journal of Monetary Economics*, 52, 749-778.

[18] Ham, Roger, P.N. (Raja) Junankar, and Robert Wells, 2009, Antagonistic Managers, Careless Workers and Extraverted Salespeople: An Examination of Personality in Occupational Choice, IZA DP 4193.

[19] Heckman, James J., 2011, Integrating Personality Psychology into Economics, NBER Working Paper No. 17378.

[20] Immordino, Giovanni, Anna Maria C. Menichini and Maria Grazia Romano, 2011, A Simple Impossibility Result in Behavioral Contract Theory, *Economics Letters*, 113, 307-309.

[21] Ishida, Junichiro, 2006, Optimal Promotion Policies with the Looking-Glass Effect, *Journal of Labor Economics,* 24, 857-878.

[22] Kaplan, Steven N., Mark M. Klebanov, and Morten Sorensen, 2008, Which CEO Characteristics and Abilities Matter?, NBER Working Papers 14195, National Bureau of Economic Research.

[23] Köszegi, Botond, 2006, Emotional Agency, *Quarterly Journal of Economics*, 21, 121–156.

[24] Köszegi, Botond, 2010, Utility from Anticipation and Personal Equilibrium, *Economic Theory*, 44 (3), 415-444.

[25] Loewenstein, George, 1987, Anticipation and the Valuation of Delayed Consumption, *Economic Journal*, 97, 666-684.

Malmendier, Ulrike, and Geoffrey Tate, 2008, Who Makes Acquisitions? CEO Overconfidence and the Market's Reaction, *Journal of Financial Economics*, 89 (1), 20-43.

[26] Mayraz, Guy, 2011, "Wishful Thinking", CEP Discussion Papers dp1092, Centre for Economic Performance, LSE.

[27] McCrae, R. R., Costa, P.T. Jr., and Busch, C.M., 1986, Evaluating Comprehensiveness in Personality Systems: The California Q-Set and the Five Factor Model, *Journal of Personality*, 57 (1), 17-40.

[28] Mobius, Markus M., Muriel Niederle, Paul Niehaus and Tanya S. Rosenblat, 2011, "Managing Self-Confidence: Theory and Experimental Evidence," Working Paper 17014, National Bureau of Economic Research.

[29] Nyhus, Ellen K, and Empar Pons, 2005, The Effects of Personality on Earnings, *Joumal of Economic Psychology*, 26 (3),363–84.

[30] Rabin, Matthew, 1994, "Cognitive dissonance and social change," *Journal of Economic Behavior and Organization*, 23, 177-194.

[31] Salgado, Jesus F., 1997, The Five Factor Model of Personality and Job Performance in the European Community, *Journal of Applied Psychology*, 82 (1), 30-43.

[32] Salgado, Jesus F., and Andrés Rumbo, 1997, Personality and Job Performance in Financial Services Managers, *International Journal of Selection and Assessment*, 5 (2), 91–100.

[33] Swank, Otto H. and Bauke Visser, 2007, Motivating through Delegating Tasks or Giving Attention, *Journal of Law, Economics and Organization*, 23, 731-742.

[34] Taylor, Shelley, 1989, *Positive Illusions: Creative Self-Deception and the Healthy Mind*, New York, Basic Books.

[35] Taylor, Shelley E., and Jonathon D. Brown, 1988, Illusion and Well-Being: A Social Psychological Perspective on Mental Health, *Psychological Bulletin*, 103 (2), 193–210.

[36] Trivers, Robert, 2000, The Elements of a Scientific Theory of Self-Deception, *Annals NY Acad Sciences*, 907, 114-131.