Optimal Incentives without Expected Utility*

Geoffrey Castillo† Víctor González-Jiménez‡

January 22, 2021

Abstract

This paper investigates the optimal design of incentives when agents distort probabilities, disregarding expected utility theory as an adequate representation of preference. We show that the type of probability distortion displayed by the agent and its degree determine whether an incentive-compatible contract can be implemented, the strength of the incentives included in the optimal contract, and the location of incentives on the output space. Our framework demonstrates that incorporating descriptively valid theories of risk in a principal-agent setting leads to incentive contracts that are typically observed in practice such as salaries, lump-sum bonuses, and high-performance commissions.

JEL Classification : D82, D86, J41, M52, M12.

Keywords: Contracts, Risk Attitude, Incentives, Probability Weighting, Mechanism Design.

---

*We are indebted to Daniel Garcia, Cole Williams, and Gyula Seres for valuable comments.

†University of Vienna, Vienna Center for Experimental Economics. E-mail: `geoffrey.castillo@univie.ac.at`

‡University of Vienna, Department of Economics and Vienna Center for Experimental Economics. E-mail: `victor.gonzalez@univie.ac.at`

1 Introduction

The theory of incentives is one of the basic building blocks of economics.[1] It shows how a principal can set up a contract to incentivize an agent whose actions are unobservable. Over decades this theory has been refined and applied to nearly all fields of economics. The contracts it predicts, however, often do not match those observed in practice (Lazear and Oyer, 2007; Prendergast, 1999; Salanié, 2003). Notably, the bulk of the literature captures risk attitudes with expected utility. Expected utility, while theoretically appealing, is not an accurate description of choice under risk (Starmer, 2000).[2]

In our paper we investigate whether relaxing the assumption of expected utility maximization changes the type of contract predicted by the theory. In particular, we consider agents who *distort* probabilities as documented by abundant evidence from decision theory (Abdellaoui, 2000; Gonzalez and Wu, 1999; Tversky and Fox, 1995; Wu and Gonzalez, 1996). [3] [4] This assumption underlies the most prominent alternative models of decision under risk, such as rank-dependent utility (Quiggin, 1982) and cumulative prospect theory (Tversky and Kahneman, 1992). We take these models and incorporate them to the theory of incentives, thus bridging the gap between the two literatures.

The adopted models of risky decision-making are not only descriptively valid, but they also satisfy a number of desirable normative properties such as first-order stochastic dominance and transitivity. Our approach thus differs from earlier research (De La Rosa, 2011; Santos-Pinto, 2008; Spinnewijn, 2013) that relied on simple cognitive biases, such as general overconfidence. There, agents would for example violate first-order stochastic dominance.

Our main contribution is to show how the principal can take advantage of

[1] See Mirrlees (1976) and Holmstrom (1979) for seminal contributions, and Laffont and Martimort (2002) and Bolton and Dewatripont (2005) for reviews.

[2] See also the references throughout this paper.

[3] See Wakker (2010, p. 204) for an extensive list of papers documenting this pattern.

[4] This pattern of choice is not only restricted to behavior in laboratory experiments, but is a regularity observed in settings with sizeable stakes Bombardini and Trebbi (2012), and everyday situations such as insurance purchase (Barseghyan et al., 2013) and gambling (Jullien and Salanié, 2000; Snowberg and Wolfers, 2010).

an agent who distorts probabilities. We consider different types of probability distortions as well as loss aversion and diminishing sensitivity. With these we find that the contracts offered mimic those observed in real-life, such as fixed wages, lump-sum bonuses, and long-shots.

We first look at agents who display optimism or pessimism. These probability distortions reflect an irrational belief that either best performance levels, in the case of optimism, or worst performance levels, in the case of pessimism, are more likely to realize. The principal reacts to these probability distortions by offering a contract that concentrates incentives at performance levels which the agent perceives to be more likely. For example, when facing an optimistic agent, the principal offers a contract that provides a large payments if highest performance levels realize—in other words, a long-shot.

We further show that, when optimism or pessimism is severe, incentive-compatible contracts are either not needed or cannot be implemented. Under severe optimism the first-best contract, on its own, induces high effort; the excessive confidence that high performance levels will realize is enough to generate strong incentives. By contrast, the incentive-compatible contract under severe pessimism concentrates incentives at the lowest performances. Moreover, to avoid perverse incentives, such as agents wanting to destroy output, the principal needs to provide a high and fixed payment for all other realizations which ultimately make the incentive-compatible contract too costly. The principal would thus offer a contract with a constant payment—a fixed wage.

Second, we go beyond optimism and pessimism and consider also probability distortions stemming from the agents' cognitive limitations to perceive probabilities. These probability distortions are referred as likelihood insensitivity (Einhorn and Hogarth, 1985; Tversky and Wakker, 1995; Wakker, 2001). Agents who are likelihood-insensitive assign too much weight to highest and lowest performance levels, but perceive performance levels in the middle of the range to be similar. When facing these agents, the principal concentrates incentives at high or low performance levels while offering flat incentives in-between. The optimal contract can thus exhibit long-shots and strong punishments.

Using our framework, we consider a number of extensions. First, we look at agents who evaluate outcomes relative to a reference point. These agents not only suffer from probability distortion but also from loss aversion and diminishing sensitivity. Again, there is ample evidence for these patterns (see Baillon et al., 2020; Bruhin et al., 2010; Kahneman and Tversky, 1979; Kahneman et al., 1991; Tversky and Kahneman, 1992).[5] Depending on the circumstances, reference-dependence gives rise to richer contracts, featuring multiple jumps or non-convexities; or simpler, more common contracts featuring a fixed wage with a lump-sum bonus.

We also relax the assumption that the principal is perfectly informed about the agent's risk preferences. To that end we introduce an adverse selection stage before the model discussed so far. On top of motivating high effort, the principal must now design a contract that allows to distinguish among two types: expected utility maximizers and non-expected utility maximizers. We find that the principal can use the agents' probability distortions to screen and incentivize agents. For example, when non-expected utility maximizers are more efficient, they can be disincentivized to mimic expected utility maximizers with a contract that, in their eyes, offers informational rents. This could be achieved via a contract that heavily concentrates incentives at extreme performance levels. Further, we find that depending on the magnitude and location of the probabilities generated by the agents' choice of effort, one type can become more efficient than the other. This result highlights that the difficulty of the task—the mapping between effort level and probability of obtaining a performance level—has consequences for the design of incentives.

Broadly speaking our paper contributes to the behavioral contract theory literature. This literature incorporates into contract theory biases such as loss aversion, present bias, other-regarding preferences and incorrect beliefs (see Koszegi, 2014, for a review). We focus on incorporating probability distortions; to our knowledge we are the first to do so. This feature puts us closer to Spalt (2013) who shows that agents with cumulative prospect theory preferences can be optimally incentivized with stock options when effort is

[5]The reader interested in reference-dependent preferences outside of the laboratory is referred to footnote 1 in Baillon et al. (2020).

contractable. We find a similar result in which the first-best contract given to optimistic or likelihood-insensitive agents exhibits an option-like shape. But we go beyond and also look at what happens when effort is *not* contractable. Additionally, Gonzalez-Jimenez (2020) demonstrates that stochastic contracts are preferred to deterministic contracts when agents distort probabilities. That paper is silent about the shape of the optimal stochastic contracts, as well as the power and location of incentives over the performance interval.

We also contribute to the contract theory literature. One of our contributions is methodological; we show that when agents have non-expected utility preferences the solution derived from the first-order approach can decrease in performance, yielding perverse incentives. To avoid this problem, and guarantee that the solution is monotonic, we modify the first-order solution using Myerson (1981)'s ironing. We thus use a tool typically used in settings of adverse selection to provide an incentive-compatible solution in a hazard setting. Another contribution regards the paradox put forward by Salanié (2003): the complex theoretical solutions in contract theory do not match the simplicity of contracts observed in the field. We show that when individuals are overly pessimistic, the emerging contract is a salary; and, if they are also loss averse, the optimal contract consists of a salary and a lump-sum bonus given for high performance levels. These two contracts are among the most popular compensation practices.

2 Setup and probability weighting functions

Consider an agent (he) hired by the principal (she) to work on a task. The agent's action consists of exerting an effort $e \in \{\underline{e}, \bar{e}\}$. Exerting the high effort $\bar{e}$ generates more disutility than exerting the low effort $\underline{e}$. For simplicity, we assume that the agent faces the following cost function:

$$c(e) = \begin{cases} c \text{ if } e = \bar{e}, \\ 0 \text{ if } e = \underline{e}. \end{cases}$$

where $c > 0$.

To incentivize the agent, the principal offers a take-it-or-leave-it contract specifying a transfer $t(q)$.If the contract is accepted, the agent proceeds to work on the task and chooses the amount of effort. We assume that the transfer included in the contract $t(q)$ enters the agent's utility through the function u, which exhibits the standard property of diminishing returns, $u' > 0$ and $u'' < 0$. As is well known, this property generates risk-averse attitudes in an expected utility framework.

The agent's action cannot be observed by the principal. Additionally, the output q is a random variable that take values in the compact interval $[\underline{q}, \bar{q}]$. Hence, while the agent's action may influence his performance, there are also factors out of his control. By observing performance the principal thus cannot determine the agent's action with certainty. However, both parties know that q is distributed according to the conditional distribution function $F(q|e)$ that admits a probability density function $f(q|e)$.

The monotone likelihood ratio property describes the relationship between production and effort.

Assumption 1. *The monotone likelihood ratio property (MLRP) states that* $\frac{d}{dq}\left(\frac{f(q|\underline{e})}{f(q|\bar{e})}\right) \leq 0$.

The MLRP establishes how informative the realizations of q are about the agent's action. Specifically, high output realizations are more likely to be drawn from a distribution of output conditional on high effort. In contrast, low output realizations are more likely to be drawn from a distribution of output conditional on low effort.

The principal is risk-neutral and has the objective function

$$\Pi(t, e) = \int_{\underline{q}}^{\bar{q}} \Big(S(q) - t(q)\Big) f(q|e) \, \mathrm{d}q,$$

where S is a function that exhibits $S' > 0$, $S'' \leq 0$ for all q, and $S(\underline{q}) = 0$.

Moreover, under the aforementioned assumptions, the preferences of the agent can be written as

$$\mathbb{E}(U(t,e)) = \int\limits_{\underline{q}}^{\bar{q}} u\Big(t(q)\Big) f(q|e)\,\mathrm{d}q - c(e). \tag{1}$$

To simplify matters and relate to standard notation in the literature, we refer to a probability in our model as the decumulative probability or *rank* associated to an output level. In other words, a probability is in our model the likelihood that a realization better than some level $Q \in [\underline{q}, \bar{q}]$ takes place. This representation has no impact on the solution to the incentive design problem.To see why, note that the agent's preference in equation (1) is equivalent to the following representation in terms of ranks:[6]

$$\mathbb{E}(U(t,e)) = \int\limits_{\bar{q}}^{\underline{q}} u\Big(t(q)\Big)\,\mathrm{d}\Big(1 - F(q|e)\Big) - c(e). \tag{2}$$

When the agent perceives probabilities accurately, expected utility (EUT), in equations (1) and (2), captures his preferences. We relax this assumption by letting the agent exhibit probability distortions, which affect his risk attitudes. We model this feature by means of a *probability weighting function* w that transforms probabilities. We impose the following assumptions on w:

Assumption 2. *Let $p := 1 - F(q|e)$. A probability weighting function is a function $w : [0,1] \to [0,1]$ such that:*

- *$w(p)$ is $\mathcal{C}^2$;*
- *$w(0) = 0$ and $w(1) = 1$;*
- *$w'(p) > 0\ \forall p \in [0,1]$;*
- *For some $\tilde{p} \in [0,1]$, $-\infty < w''(p) < 0$ if $p < \tilde{p}$ and $\infty > w''(p) > 0$ if $p > \tilde{p}$;*
- *$\lim_{p\to 1} w'(p) = \infty$ and $\lim_{p\to 0} w'(p) = 0$ if $\tilde{p} = 0$;*

[6] Let $q_1, q_2 \in [\underline{q}, \bar{q}]$ with $q_2 > q_1$. Then,

$$\int_{q_1}^{q_2} f(q|e)\mathrm{d}q = F(q_2|e) - F(q_1|e) = 1 - F(q_1|e) - (1 - F(q_2|e)) = \int_{q_2}^{q_1} \mathrm{d}\left(1 - F(q|e)\right).$$

- $\lim_{p\to 0} w'(p) = \infty$ *and* $\lim_{p\to 1} w'(p) = 0$ *if* $\tilde{p} = 1$*;*
- $\lim_{p\to 0} w'(p) = \infty$ *and* $\lim_{p\to 1} w'(p) = \infty$ *if* $\tilde{p} \in (0, 1)$*;*
- *There exists* $\hat{p}$ *such that* $w(\hat{p}) = \hat{p}$ *if* $\tilde{p} \in (0, 1)$*.*

In words, the probability weighting function is an strictly increasing and continuous function that maps the unitary interval onto itself. The function exhibits at least two fixed points, one at impossibility $p = 0$ and one at certainty $p = 1$.

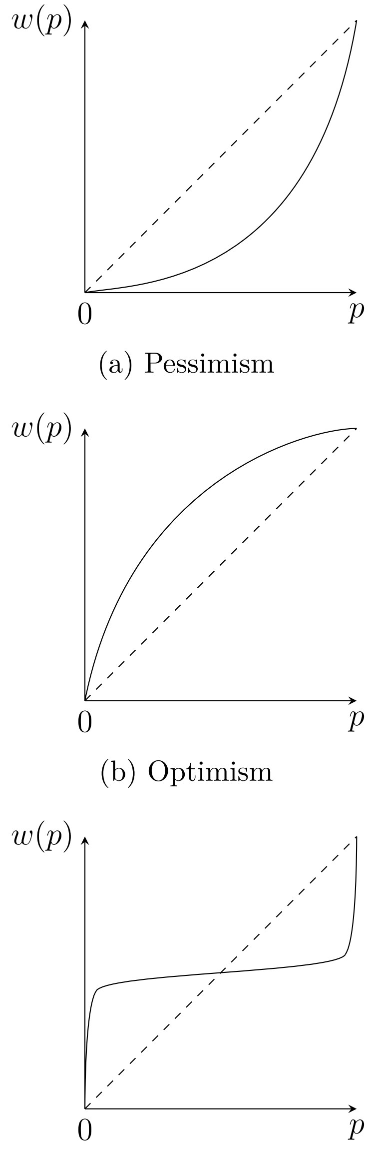
The function w can take three different shapes depending on the location of the inflection point $\tilde{p}$. When $\tilde{p} = 0$ the function is convex everywhere and probabilities associated to worst performance levels are given a larger weight than that given to probabilities associated to best performance levels. Figure 1a presents an example of a convex weighting function. In contrast, when $\tilde{p} = 1$ the function is concave everywhere and probabilities associated to best performance levels receive large weight while probabilities associated to worst performance levels receive small weight (Figure 1b). Finally, when $\tilde{p} \in (0, 1)$, the probability weighting function exhibits an inverse-S shape (Figure 1c). In this case the agent assigns very large weights to extreme performance levels.

The assumptions of extreme sensitivity to rare and almost-certain events, $\lim_{p\to 1} w'(p) = \infty$ and $\lim_{p\to 0} w'(p) = \infty$, are incorporated in the most prominent proposals of probability weighting functions. For instance, in the parametric forms proposed by Prelec (1998), Tversky and Kahneman (1992), and Goldstein and Einhorn (1987). Moreover, non-continuous proposals of probability weighting functions, e.g. Neo-additive (Chateauneuf et al., 2007) or Kahneman and Tversky (1979), include discontinuities at extreme probabilities to account for regularities in behavior that go in line with this extreme sensitivity to extreme probability events.

With the probability weighting function w, the preferences of the agent who exhibits sensitivity to probabilities are characterized by rank-dependent utility (RDU):

$$RDU(t, e) = \int_{\bar{q}}^{\underline{q}} u\Big(t(q)\Big)\, \mathrm{d}w\Big(1 - F(q|e)\Big) - c(e). \tag{3}$$

Figure 1: Examples of probability weighting functions



(a) Pessimism

$w(p)$
0
p

(b) Optimism

$w(p)$
0
p

(c) Likelihood insensitivity

Note: Dashed lines represent accurate perception of probability.

We also refer to agents with RDU preferences as non-EUT agents since their perception of probabilities prevents them from using mathematical expectations to evaluate possible outcomes.[7] We will assume that the principal can contract with either EUT or non-EUT agents and, as is standard in the literature, we assume that she is fully informed about the agent's risk preferences and designs and implements incentives accordingly.

3 Optimistic and Pessimistic agents

We start by studying the optimal design of incentives when the principal faces two specific types of non-EUT agents: pessimists and optimists. These agents deviate from expected utility due to motivational factors reflecting a proneness or a dislike for risk.

Pessimists dislike risk and assign large weights to the worst outcomes. In other words, they believe the worst outcomes realize more often. Pessimism is captured by a convex probability weighting function:

Definition 1. *Pessimism is characterized by a function w with the properties of Assumption 2 and the restriction $\tilde{p} = 0$.*

Optimists like risk and assign large weights to the best outcomes—they believe the best outcomes realize more often. Optimism is captured with a concave probability weighting function:

Definition 2. *Optimism is characterized by a function w with the properties of Assumption 2 and the additional restriction $\tilde{p} = 1$.*

We will also talk about agents who are more optimistic or more pessimistic than others using the following definition:

Definition 3. *An agent i is more pessimistic (optimistic) than some agent j if w_i, the weighting function of agent i, is such that $w_i = \theta \circ w_j$, for a continuous, strictly increasing, convex (concave) $\theta : [0, 1] \to [0, 1]$*

[7] It is noteworthy to emphasize that probability distortion is not the only departure from expected utility that we consider. In Section 5 we also consider reference-dependence.

More convex probability weighting functions generate more pessimism because more convexity makes the agent assign larger weights to high performance levels and smaller weights to low performance levels. The reasoning is mirrored for a more concave probability weighting function.[8]

3.1 First best

If effort is contractable, the principal can take advantage of the irrationalities exhibited by non-EUT agents by concentrating pay at performance levels that are, for the agents, more likely to realize. These first-best contracts are described in the next Lemma. All of our proofs are in Appendix A.

Lemma 1. *Under Assumptions 1 and 2, the first-best contract, t_{FB}, exhibits two possible shapes, both continuous:*

1. *Constant everywhere if the agent is EUT or pessimistic;*
2. *Strictly increasing in performance if the agent is optimistic.*

The first part of Lemma 1 establishes the standard risk-sharing argument of Borch (1960). When the agent is EUT and exhibits risk aversion the principal insures him with a contract that pays a fixed amount regardless of the performance level realized. The magnitude of the transfer ensures that the contract will be accepted by the agent and accounts for his degree of risk aversion. Since high effort is specified in the contract and the MLRP guarantees that high performance is likely, the principal is likely to obtain high production levels.

Instead, when facing an optimist, the first-best contract increases in performance. Such contract offers full insurance to this agent: it provides larger payments for realizations that are perceived to be more likely and lower payments for realizations that are perceived to be less likely. The principal is now taking advantage of the agent's sensitivity to probabilities: from her non-probability-distorted point-of-view it provides larger payments for unlikely events and smaller payments for more likely events.

[8] In our setting these comparisons can only be carried out with convex and concave functions that satisfy the properties specified in Assumption 2. For instance, the weighting function $w(p) = p^2$ cannot be included in these comparisons since its derivative is not steep enough at almost-certain events and goes against the assumption $\lim_{p\to 1} w'(p) = \infty$.

The principal would like to take advantage of a pessimistic agent in a similar way. This strategy would imply a contract that decreases in performance, offering large payments in case of low performance and low payments in case of high performance. This contract, however, would encourage the agent to destroy output to attain the highest possible payment. The principal instead proposes a contract that pays a constant transfer and satisfies the participation constraint. Since risk is eliminated, this solution eliminates the agent's pessimism, making it impossible to the principal to exploit the agent's probability distortions. But it prevents perverse incentives and does not make the principal incur excessive expenditures resulting from offering to the pessimistic agent a contract increasing in performance.

3.2 Second best

We now consider the more realistic case in which the agent's action is not contractable. The principal seeks to maximize her objective function by choosing a transfer that elicits high effort and that is accepted by the agent, who has reservation utility $\bar{U}$. Therefore, the maximization problem of the principal is now:

$$\max_{t(q)} \int_{\underline{q}}^{\bar{q}} \Big(S(q) - t(q)\Big) f(q|\bar{e}) \, \mathrm{d}q$$

$$\text{s.t.} \quad \int_{\underline{q}}^{\bar{q}} u(t) w'(1 - F(q|\bar{e})) f(q|\bar{e}) \, \mathrm{d}q - c \geq \int_{\underline{q}}^{\bar{q}} u(t) w'(1 - F(q|\underline{e})) f(q|\underline{e}) \, \mathrm{d}q,$$

$$\int_{\underline{q}}^{\bar{q}} u(t) w'(1 - F(q|\bar{e})) f(q|\bar{e}) \, \mathrm{d}q - c \geq \bar{U}.$$

In the absence of probability distortions, $w(p) = p$, the standard solution of Holmstrom (1979) applies: the second-best contract specifies transfers that strictly increase in performance. Moreover, under weak conditions those transfers are marginally increasing in performance.[9] We present this solution

[9] It is well-known that a linear contract emerges as second-best optimal under stringent

in the next Proposition. Figure 2a illustrates the resulting contract.

Proposition 1. *Under Assumption 1 and in the absence of probability distortions, $w(p) = p$, the optimal incentive scheme, t^B_{SB}, is continuous and strictly increasing in performance.*

Before presenting the second-best contract for non-EUT agents, we introduce an assumption crucial to our analysis. We strengthen the MLRP to ensure that, despite his tendency to distort probabilities, output realizations are sufficiently informative to the agent.

Assumption 3. *The modified monotone likelihood ratio property (W-MLRP) states that* $\frac{d}{dq}\left(\frac{w'(1-F(q|\underline{e}))f(q|\underline{e})}{w'(1-F(q|\bar{e}))f(q|\bar{e})}\right) < 0$

The W-MLRP implies that the principal, who is fully informed about the agent's risk attitudes, anticipates how probability distortions affect the agent's beliefs about the informativeness of his own action. The agent, on the other hand, is naive and does not evaluate the informativeness of performance realizations using mathematical expectations, which would be equivalent to anticipating the way in which the principal evaluates those realizations. Assumption 3 thus entails that the second-best contracts presented below take advantage of this naivete.

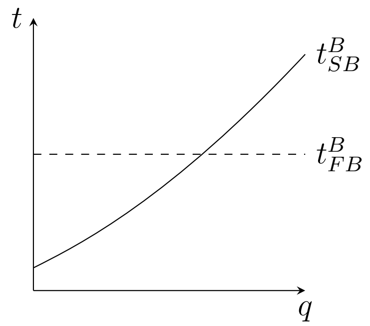
The W-MLRP is clearly more stringent than the standard MLRP. Under pessimism and optimism, however, the W-MLRP can be attributed to reasonable properties of the probability weighting function.

Remark 1. *Under Assumptions 1, 2 and pessimism (optimism), the W-MLRP is satisfied if the probability weighting function w exhibits more convexity (less concavity) at probabilities generated by low effort, $\underline{e}$, than at probabilities generated by high effort, $\bar{e}$.*

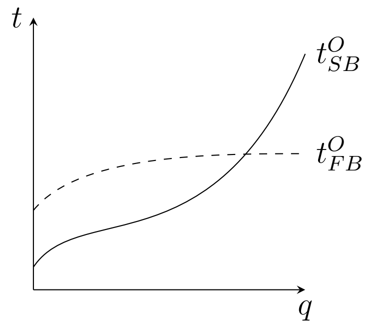
The standard MLRP implies that a high performance level occurs with higher probability when the agent chooses high effort $\bar{e}$. In other words,

conditions. Specifically, the production function must be additive on effort, the agent's utility function must belong to the CARA family, and the stochastic component of production must be normally distributed. Also, Holmstrom and Milgrom (1987) present a dynamic model where a linear contract can emerge as optimal.

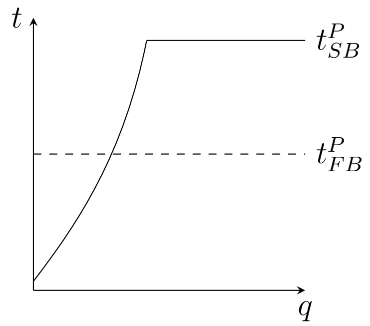
Figure 2: Illustration of Propositions 1, 2 and 3.



(a) Optimal contract under EUT



(b) Optimal contract under optimism



(c) Optimal contract under pessimism

Note: Dashed lines represent first-best contracts. Solid lines represent second-best contracts.

output realizations carry information about the effort. When probabilities are distorted by the agent, output realizations are as informative to him if the probability weighting function assigns different weights to the probabilities implied by high and low effort. This happens when the weighting function exhibits more convexity at the probabilities implied by high effort, making good news and bad news weight differently, and resulting in the W-MLRP.

We will assume the W-MLRP holds; note, however, that thanks to Remark 1 the results presented below can be obtained using the standard MLRP together with the necessary restrictions on the concavity or convexity of the weighting function.

The next two Propositions describe the properties of the second-best contracts that solve the principal's program when she faces an optimist or a pessimist and when effort is not contractable.

Proposition 2. *Under Assumptions 2 and 3 and* optimism*, two shapes of the optimal incentive scheme, t^O_{SB}, both continuous, are possible.*

1. *Identical to the first-best contract in Lemma 1 if optimism is severe.*
2. *Strictly increasing in performance with steep pay increments at the highest performance realizations if optimism is moderate.*

This contract is presented in Figure 2b. When optimism is severe, the first-best contract suffices to incentivize high effort. Remember that this first-best contract specifies a payment that increases in performance. Since the overly optimistic agent believes that high performance levels will realize, the contract convinces him that high effort is profitable. When the agent is only moderately optimistic, he needs to be incentivized with a contract that offers large rewards at the high-end of the output space. The combination of moderate optimism and strong incentive at high performance levels inflate the agent's perceived benefit of exerting high effort.

Proposition 3. *Under Assumptions 2 and 3 and* pessimism*, two shapes of the optimal incentive scheme, t^P_{SB}, both continuous, are possible.*

1. *Constant everywhere if pessimism is severe.*
2. *Smoothly increasing in performance up to some threshold after which pay is performance-insensitive if pessimism is moderate.*

Figure 2c presents the contract offered to pessimists. When the agent is moderately pessimistic, the principal must offer a contract that provides higher transfers in exchange of higher performance at the lower end of the output interval while also being performance-insensitive at the higher end of the output interval. These incentives, together with the irrational, yet moderate, confidence that low performance levels realize, motivate the agent to exert high effort. Note that implementing an increasing payments schedule at high performance levels is inefficient because the agent perceives these realizations to be unlikely. The principal would thus require excessively large transfers at those performance levels to motivate him. It is therefore more efficient to concentrate incentives at low performance levels.

The solution presented in Proposition 3 emerges after ironing is applied to the solution resulting from the first-order approach (Myerson, 1981). The first-order approach establishes that the optimal contract should have a segment where pay *decreases* in performance: the principal would like to avoid paying high incentives at performance levels that are considered unlikely. As already mentioned, however, this contract incentivizes the agent to destroy output; ironing prevents these perverse incentives.

On the other hand, severe pessimism limits the incentives that the principal can implement. As in the case of moderate pessimism, an incentive-compatible contract should concentrate incentives at the lower end of the output interval. As pessimism becomes more severe the concentration becomes more acute, thus extending the performance-insensitive segment of the contract. But paying a fixed and large transfer at performance levels perceived by the agent to be highly unlikely is an unnecessary and large expenditure. For sufficiently large pessimism, this expenditure outweighs the benefits, leading the principal to offer a constant payment.

We turn to compare the location of punishments and rewards in the second-best contracts presented in Propositions 1 to 3. We focus on the contracts that emerge when the agent has moderate optimism or moderate pessimism. As we have seen, in these cases incentive-compatibility requires a deviation from the first-best. The following corollaries formalize these comparisons.

Corollary 1. *Let* $q^* \in (\underline{q}, \bar{q})$ *satisfy* $w'\Big(1 - F(q^*|\bar{e})\Big) = 1$*. Under moderate optimism,* t^O_{SB} *(Proposition 2) specifies punishments and rewards at higher performance levels as compared to* t^B_{SB} *(Proposition 1) by offering higher transfers in* $q \in [\underline{q}, q^*]$ *and lower transfers in* $q \in (q^*, \bar{q}]$*.*

Corollary 1 shows that the second-best contract given to moderate optimists offers incentives at higher performance levels compared to the contract given to EUT agents. Since the moderate optimists perceive low output realizations to be unlikely, locating punishments at intermediate performance levels rather than at lower ones is more effective. These punishments yield contracts with higher-powered incentives at low performance levels.

These agents also perceive high output realizations to be more likely. The principal can take advantage of this perception by specifying steep but lower-powered rewards at these performance levels; the agent's moderate confidence that high output levels realize are complementary to the incentives provided at this end of the output space.

On the other hand, for the moderately pessimistic agent we have:

Corollary 2. *Let* $q^* \in (\underline{q}, \bar{q})$ *satisfy* $w'\Big(1 - F(q^*|\bar{e})\Big) = 1$*. Under moderate pessimism,* t^P_{SB} *(Proposition 3) specifies punishments and rewards at lower performance levels as compared to* t^B_{SB} *(Proposition 1) by offering lower transfers in* $q \in [\underline{q}, q^*]$ *and higher transfers in* $q \in (q^*, \bar{q}]$*.*

Compared to the EUT agent, the moderate pessimist receives a contract that specifies rewards and punishments at lower performance levels. This location of incentives implies that lower-powered incentives are given at the lower end of the output interval. Moreover, to incentivize the agent at realizations that are perceived more unlikely, he is given higher, yet flat, transfers.[10]

[10]The flatness at large output levels does not imply that the principal does not implement rewards on this segment of the output interval. In fact, to create rewards the constant pay offered by the second-best contract needs to be strictly larger than the constant pay offered by the first-best contract for high and intermediate output levels. Since the pessimistic agent thinks these levels of performance are unlikely, this way of incentivizing is more efficient than offering an increasing incentive scheme. To motivate the agent at this segment of the output interval the principal would need to incur a large and unnecessary expenditure.

To conclude this section, we investigate how more optimism or more pessimism, as defined in Definition 3, affects the second-best contract offered by the principal.

Corollary 3. *More optimism (Definition 3) implies a contract* t^O_{SB} *(Proposition 2) with transfers that:*

- *Increase more steeply in* q *at lower and higher performance levels if optimism is moderate.*
- *Increase more steeply in* q *at lower performance levels and that become flatter over a larger performance segment if optimism is severe.*

Corollary 4. *More pessimism (Definition 3) implies a contract* t^P_{SB} *(Proposition 3) with transfers that increase more steeply in* q *at lower performance levels and that become flatter over a larger performance segment.*

We first focus in the case in which an agent's higher degree of optimism does not make him an excessive optimist. In such case, more optimism yields contracts that concentrate punishments and rewards toward higher performance levels. Second, when the agent is a severe optimist, a higher degree of optimism leads to contracts specifying larger transfers in exchange for higher output at the lower end of the performance interval. Such scheme elicits high effort by offering strong incentives at realizations that are heavily underweighted by the agent while taking advantage of the agent's increasing optimism to keep him motivated without having to compensate him.

Conversely, more pessimism leads the principal to concentrate incentives at lower performance levels. Punishments for bad performance are thus located at lower performance levels as pessimism becomes more prominent. Also, more pessimism also implies that, to avoid unnecessary expenses, the incentive scheme is flatter for a larger segment.

4 Likelihood insensitivity and inverse S-shaped probability weighting functions

Thus far, we have studied the optimal design of incentives when the principal contracts with agents who deviate from EUT due to optimism or pessimism. Optimism and pessimism, however, cannot alone account for the common finding that individuals, when making risky decisions, exhibit an inverse S-shaped probability weighting function (see Wakker, 2010, p.204, and Fehr-Duda and Epper, 2011, for extensive lists of references documenting this pattern). This pattern is best understood as a consequence of *likelihood insensitivity* (Tversky and Wakker, 1995; Wakker, 2001), the cognitive limitations that prevent individuals from discriminating probabilities accurately. A likelihood-insensitive individual assigns excessively large weights to very small or very large probabilities—associated to near-certain and near-impossible events—and similar, or even indistinguishable, weights to intermediate probabilities, thus yielding an inverse S-shaped probability weighting function.

We look at the optimal incentive design when the principal faces likelihood-insensitive agents. We start by formalizing likelihood insensitivity.

Definition 4. *Likelihood insensitivity is characterized by a probability weighting function $w(p)$ with the properties of Assumption 2 and the restriction $\hat{p} = \tilde{p} = 0.5$.*

Likelihood insensitivity and pessimism or optimism can occur at the same time. When likelihood insensitivity coexists with pessimism the resulting probability weighting function has an interior fixed point at $\hat{p} \in (0, 0.5)$. In contrast, when likelihood insensitivity coexists with optimism, the probability weighting function has an interior fixed point in $\hat{p} \in (0.5, 1)$. We do not specify whether the likelihood-insensitive agent is also optimistic or pessimistic and instead assume generally that $\hat{p} \in (0, 1)$.

To compare agents according to their degree of likelihood insensitivity, we first need to introduce Tversky and Wakker (1995)'s definition of subadditivity:

Definition 5. *A function $\phi : [0, 1] \to [0, 1]$ is subadditive if $\phi(0) = 0$,*

$\phi(1) = 1$, ϕ is C^2 with $\phi' > 0$, and there exists constants ϵ, ϵ' such that

$$\phi(q) \geq \phi(r+q) - \phi(r)$$

whenever $0 < q < r < 1$ and $r + q \leq 1 - \epsilon$, and

$$1 - \phi(1-q) \geq \phi(r+q) - \phi(r)$$

whenever $0 < q < r < 1$ and $r \geq \epsilon'$.

We can then state the more-likelihood-insensitive-than relation:

Definition 6. *Agent i is more likelihood insensitive than agent j if the weighting function of agent i is such that $w_i = \phi \circ w_j$, for a subadditive function $\phi : [0,1] \to [0,1]$*

Intuitively, an agent is more likelihood insensitive when he assigns more decision weights to extreme performance levels and less decision weights to middle-ranged performance levels.

4.1 First best

The following lemma presents the first-best contract when effort is contractable and agents are likelihood insensitive. As we will see, the shape of the contract depends on whether optimism or pessimism accompanies likelihood insensitivity.

Lemma 2. *Under Assumptions 2 and 3 and likelihood insensitivity, the first-best contract, t^{LI}_{FB}, exhibits two possible shapes, both continuous:*

1. *constant up to threshold after which pay strictly increases in performance if the agent is optimistic;*
2. *constant everywhere if the agent is pessimistic.*

For a likelihood-insensitive agent who is also an optimist, $\tilde{q}$, the output level generating $w''\big(1 - F(\tilde{q}|e)\big) = 0$ is located at low output levels. In this case, the contract offers a fixed amount for low and medium performance realizations, and is increasing for high performance realizations. This way, the principal takes into account the agent's perception that low probabilities are more likely to realize and offers him perfect insurance. Instead, when the likelihood-insensitive agent is a pessimist—$\tilde{q}$ is located at high output levels—the principal implements a fixed pay contract with a transfer that satisfies the agent's participation constraint.

As in Lemma 1 and Propositions 2 and 4, to obtain the first-best contracts presented in Lemma 2 we iron the solution given by the first-order approach. We do so to eradicate segments in which payment decreases with performance. Full insurance concentrates pay at the extremes of the output interval, which the agent perceives to realize more often, and provides lower incentives at intermediate probabilities, where the agent exhibits insensitivity. For the pessimistic and likelihood-insensitive agent there is a large segment where payment decreases with performance. In that case the principal prefers to offer a fixed payment to remove any perverse incentives.

4.2 Second best

When effort is not contractable, likelihood insensitivity makes the principal's program more restrictive. The W-MLRP can no longer be related to properties of the probability weighting function, as it was the case for optimists and pessimists (Remark 1). The W-MLRP must instead be explicitly assumed. To see why, note that around the inflection point $\tilde{p}$, the probabilities implied by high or low effort are almost indistinguishable to the agent. Thus Remark 1, which states that the weighting function must be more convex at the probability implied by higher effort compared to the probability implied by low effort for the W-MLRP to hold, does not apply.

The next Proposition presents the second-best contract offered to the likelihood-insensitive agent. The solution depends on whether optimism or pessimism accompanies likelihood insensitivity, while the degree of likelihood insensitivity determines whether transfers increase in performance at intermediate levels.

Proposition 4. *Under Assumptions 2 and 3 and likelihood insensitivity, the second-best incentive scheme, t^{LI}_{SB} exhibits the following, continuous shapes.*

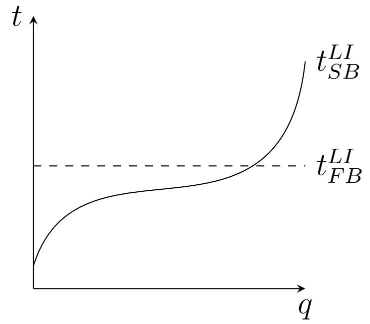
1. *Identical to the first-best contract, t_{FB}^{LI} in Lemma 2 under severe optimism.*
2. *A fixed amount is paid everywhere under severe pessimism.*
3. *Increasing everywhere in performance and with steep payment increments at extreme performance levels under moderate likelihood insensitivity and pessimism/optimism.*
4. *A fixed amount is paid for some finite, fixed, compact interval, but above and/or below this interval pay increases in performance under severe likelihood insensitivity and moderate pessimism/optimism.*

Under severe optimism or severe pessimism, the solution to the principal's program either does not require or cannot satisfy the incentive compatibility constraint. This is analogous to the solution presented in Propositions 2 and 3. Severe optimism, exacerbated by the agent's likelihood insensitivity, renders unnecessary the introduction of rewards and punishments compared to the full insurance contract. In this case the agent's confidence that high performance levels are more likely, coupled with the first-best contract having an increasing segment, is enough to motivate the agent. On the other hand, under severe pessimism incentive-compatible contracts are too costly and cannot be implemented.

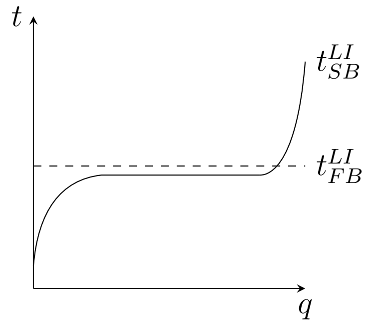
When optimism or pessimism are moderate, the optimal contract concentrates incentives at one or both extremes of the output interval. This contract thus combines the key features of the second-best contracts given to optimists and pessimists. When the agent exhibits moderate likelihood insensitivity the optimal contract increases in performance everywhere. Figure 3a presents an example of such a contract when such degree of likelihood insensitivity coexists with moderate pessimism. By contrast, when the agent exhibits severe likelihood insensitivity the contract includes a performance-insensitive segment to address the low weights that the agent assigns to intermediate performance levels. Figure 3b presents an example of such a contract when strong likelihood insensitivity coexists with moderate pessimism.

We next compare the power of the incentives as well as the location of rewards and punishments of the second-best contract presented in Proposition 4 to the contract given to the EUT agents.

Figure 3: Illustration of Proposition 4.



(a) Optimal contract under moderate likelihood insensitivity and moderate pessimism



(b) Optimal contract under strong likelihood insensitivity and moderate pessimism

Note: Figures present contracts when pessimism coexists with likelihood insensitivity. Dashed lines represent first-best contracts. Solid lines represent second-best contracts.

Corollary 5. *Let q_l^{**} satisfy $w'\Big(1-F(q_l^{**}|\bar{e})\Big)=1$ and $w''>0$, and let q_h^{**} satisfy $w'\Big(1-F(q_l^{**}|\bar{e})\Big)=1$ and $w''<0$. Under moderate optimism/pessimism, t_{SB}^{LI} (Proposition 4) specifies rewards and punishments at more extreme performance realizations than t_{SB}^{B} (Proposition 1), by paying lower transfers in $q\in[\underline{q},q_l^{**}]$ and higher transfers in $q\in[q_h^{**},\bar{q}]$.*

Corollary 5 shows that the second-best contract given to likelihood-insensitive agents specifies sizeable rewards at large performance levels and punishments at small performance levels. The principal thus concentrates sticks and carrots to where it matters for the agent.

We conclude this section by establishing how stronger likelihood insensitivity (Definition 6) affects the second best contract) (Proposition 4.

Corollary 6. *More likelihood insensitivity (Definition 6) implies a contract t_{SB}^{LI} (Proposition 4) with transfers that increase more steeply at one or both extreme performance levels, and that become flatter for a larger finite, fixed, compact interval.*

5 Extensions

5.1 Agents with Loss Aversion and Diminishing Sensitivity

We enrich the agent's risk preferences by considering Cumulative Prospect Theory (CPT, Tversky and Kahneman, 1992). Agents with these preferences evaluate potential transfers relative to a reference point $r>0$. Transfers below the reference point count as *losses* while transfers above count as *gains*. Typically the reference point r represents the status quo (Kahneman and Tversky, 1979; Tversky and Kahneman, 1981).

The main departure of CPT with respect to RDU and EUT is that the agent can exhibit different risk preferences for gains and for losses. This is captured with two ingredients. First, transfers enter the agent's utility differently whether they are classified as losses or as gains.

Assumption 4. *The value function, $V(t,r)$, is a piece-wise function,*

$$V(t,r) = \begin{cases} v\Big(t(q) - r\Big) & \text{if } t(q) \geq r, \\ -\lambda v\Big(t(q) - r\Big) & \text{if } t(q) < r, \end{cases}$$

with the following properties:

- $\lambda > 1$;
- $v(0) = 0$;
- $v' \geq 0$ *for all* $q \in [\underline{q}, \bar{q}]$;
- $v'' < 0$ *if* $t(q) \geq r$, *and* $v'' > 0$ *if* $t(q) < r$.

In the domain of losses the agent's utility is convex, generating risk seeking, while in the domain of gains it is concave, generating risk aversion. Moreover, the value function from Assumption 4 assumes loss aversion: transfers classified by the agent as losses loom larger than equally-sized transfers classified as gains. This property is captured by the parameter $\lambda > 1$: it enters the value function only for losses and so generates a kink at the reference point.

The second ingredient is that the shape of the probability weighting function can be different depending on the domain. Probabilities associated with gains are transformed with the probability weighting function w that we have already seen (Assumption 2). On the other hand, probabilities associated with losses are transformed with a probability weighting function z that applies transformations to cumulative probabilities, $F(q|e)$, rather than to decumulative probabilities, $1 - F(q|e)$. In other words, in the domain of losses the CPT agent orders possible transfers from the least-desirable, $t(\underline{q})$, to the closest to the reference point from below, and uses z to transform the probabilities that emerge from these—as the literature describes them—loss ranks.

We simplify the problem by assuming that z adopts the properties of the weighting function w.

Assumption 5. *A probability weighting function for losses is a function $z : [0,1] \to [0,1]$ satisfying the duality $z(F(q|e)) = 1 - w(1 - F(q|e))$ for any e.*

All in all, the utility of an agent with CPT preferences when incentivized with a contract $t(q)$ is

$$CPT(t,e,r) = \int_{q_r}^{\bar{q}} v\,(t(q)-r)\,w'\Big(1-F(q|e)\Big)f(q|e)\,\mathrm{d}q - \int_{\underline{q}}^{q_r} \lambda v(r-t(q))z'\Big(F(q|e)\Big)f(q|e)\,\mathrm{d}q - c(e), \tag{4}$$

where $q_r \in [\underline{q}, \bar{q}]$ is a performance level satisfying $t(q_r) = r$.

The maximization problem of the principal is now:[11]

$$\max_{t(q)} \quad \int_{\underline{q}}^{\bar{q}} \Big(S(q)-t(q)\Big)f(q|\bar{e})\,\mathrm{d}q$$

$$\text{s.t.} \quad CPT(t,\bar{e},r) \geq 0,$$

$$CPT(t,\bar{e},r) \geq CPT(t,\underline{e},r)$$

The optimal incentive scheme offered to agents with CPT preferences is presented in the next Proposition.

Proposition 5. *Under Assumptions 2 to 5, the optimal incentive scheme exhibits the following possible shapes:*

1. *$t(q) = r$ is paid.*
2. *$t(q) = r$ is paid for some finite, compact, fixed interval, below some threshold, but above this level the pay depends on performance as in Proposition 2, Proposition 3, or Proposition 4, depending on the shape of w.*
3. *pay depends on performance as in Proposition 2, Proposition 3, or Proposition 4, depending on the shape of w.*

When the agent has CPT preferences the second-best contract often includes a segment where transfers are constant. The reason behind this

[11] To further simplify the problem, we normalize the reservation utility of the agent to zero.

performance insensitive segment is loss aversion. A contract that exposes the agent to losses generates large disutility, which leads him to reject it. To prevent this rejection the principal can introduce large rewards that compensate the agent for facing such risk, or she can completely eradicate the possibility that the agent ends up in the domain of losses. The former solution is expensive since losses loom larger than equally sized gains by a factor of λ. Thus, the principal offers, wherever necessary, the minimum amount that locates him in the domain of gains: $t(q) = r$. This fixed payment is offered to the agent for all performance levels where the second-best contract that best fits the agent's probability perceptions falls short from his reference point.

Moreover, the second-best contract in Proposition 5 often specifies transfers that vary with performance in the same way as in Propositions 2, 3, or 4, depending on which of these contracts best fits the agent's probability perception. That is because in the domain of gains, the CPT agent exhibits risk attitudes that are equal to those of the RDU agent, and the second-best contract that motivates the RDU agent also suffices to incentivize the CPT agent in that domain.

Proposition 5 can lead to simple contracts often observed in practice. For instance, when the CPT agent has a convex weighting function, the optimal contract is binary: it exhibits a fixed salary and a lump-sum bonus. The former ensures that the agent does not face losses while the latter is optimal when pessimism is severe in the domain of gains. The emergence of these binary incentive schemes is also documented by Herweg et al. (2010). The difference between their setting and ours is that they do not consider probability transformations, so their agents are not characterized by CPT, and the reference point is endogenous, à la Koszegi and Rabin (2006).

Another simple contract emerges when the CPT agent displays optimism. Then, the optimal contract consists of a performance-insensitive salary plus a pay-for-performance segment that generates larger payments in the contingencies of higher performance. Such a contract is similar to a fixed-wage contract with commission, typically used to incentivize salesforce.

Complicated contracts can also be obtained when the agent exhibits likelihood insensitivity and moderate pessimism or optimism. There, the optimal contract exhibits multiple jumps or non-convexities. The first jump occurs when pay becomes large enough to bring the agent to the domain of gains; the second, when the region of insensitivity is reached; and the last one, when the region of insensitivity is escaped. This contract can be seen as an incentive contract with multiple commission bonuses.

5.2 Adverse selection followed by moral hazard

Thus far we have considered a situation in which the principal knows the agent's risk attitude. In this section we relax that assumption: the principal first needs to screen among the different types of agents before incentivizing high effort.

We consider two types of agents: EUT and non-EUT. We simplify matters by assuming that all non-EUT agents have RDU preferences with likelihood insensitivity: their weighting function is inverse-S shaped. Various studies support this assumption. For instance, Bruhin et al. (2010) find that the majority of individuals deviate from accurate perceptions of probability and exhibit inverse-S shaped probability weighting functions, whereas a minority conforms to EUT. We also assume that likelihood insensitivity goes with pessimism. This assumption is also in line with the common findings from the literature (Wakker, 2010).

The principal knows that she contracts with a EUT agent with probability π_B and with a non-EUT agent with probability $1 - \pi_B$. We denote by t_B and t_{nB} the transfers to a EUT and to a non-EUT agent. The timing is as follows:

1. The agent learns his type: B or nB.
2. The principal offers a stochastic contract $t(q)$.
3. The agent accepts or rejects the contract.
4. If the contract is accepted, the agent exerts effort e, which translates into performance q.
5. The transfer specified by the contract is paid to the agent.

In what follows we assume that the principal finds it valuable to induce a high level of effort in both EUT and non-EUT types. Therefore, the moral hazard incentive constraint of the EUT agent is

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_B(q)\Big) f(q|\bar{e})\,\mathrm{d}q - c \geq \int_{\underline{q}}^{\bar{q}} u\Big(t_B(q)\Big) f(q|\underline{e})\,\mathrm{d}q, \tag{5}$$

and the moral hazard incentive constrain of the non-EUT agent is

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_{nB}(q)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})\,\mathrm{d}q - c \geq \int_{\underline{q}}^{\bar{q}} u\Big(t_{nB}(q)\Big) w'\Big(1 - F(q|\underline{e})\Big) f(q|\underline{e})\,\mathrm{d}q. \tag{6}$$

To distinguish between the two agents the contracts satisfy the adverse selection incentive-compatible constraint for the EUT agent

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_B(q)\Big) f(q|\bar{e})\,\mathrm{d}q - c \geq \max_{e\in\{\underline{e},\bar{e}\}} \left\{ \int_{\underline{q}}^{\bar{q}} u\Big(t_{nB}(q)\Big) f(q|\bar{e})\,\mathrm{d}q - c(e) \right\}, \tag{7}$$

and for the non-EUT agent

$$\begin{aligned} &\int_{\underline{q}}^{\bar{q}} u\Big(t_{nB}(q)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})\,\mathrm{d}q - c \\ &\geq \max_{e\in\{\underline{e},\bar{e}\}} \left\{ \int_{\underline{q}}^{\bar{q}} u\Big(t_B(q)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})\,\mathrm{d}q - c(e) \right\}. \end{aligned} \tag{8}$$

Finally, the participation constraint of the agents when their outside option is zero is

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_B(q)\Big) f(q|\bar{e})\,\mathrm{d}q - c \geq 0, \tag{9}$$

and

$$\int\limits_{\underline{q}}^{\bar{q}} u\Big(t_{nB}(q)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})\, \mathrm{d}q - c \geq 0. \tag{10}$$

The solution to the principal's program is presented in Proposition 6. As in standard adverse selection models the solution consists of offering a menu of contracts where each contract targets a particular type. In addition the contract for the most efficient type includes an informational rent that discourages him to mimic the less efficient type. In contrast to standard adverse selection models, however, the menu depends on how the agent's effort translates into probabilities.

Proposition 6. *The menu of optimal contracts in a setting of adverse selection followed by moral hazard, $\{t_{nB}, t_B\}$, exhibits the following properties:*

- *t_B satisfies $\mathbb{E}\left(u(t_B)|\bar{e}\right) = c$ while t_{nB} satisfies $\tilde{\mathbb{E}}\left(u(t_{nB})|\bar{e}\right) = \tilde{\mathbb{E}}\left(u(t_B)|\bar{e}\right)$ if $w'\left(1 - F(q|\bar{e})\right) > 1$.*
- *t_{nB} satisfies $\tilde{\mathbb{E}}(t_{nB}|\bar{e}) = c$ while t_B satisfies $\tilde{\mathbb{E}}(t_B|\bar{e}) = \tilde{\mathbb{E}}(t_{nB}|\bar{e})$ if $w'\left(1 - F(q|\bar{e})\right) \leq 1$.*

where $\tilde{\mathbb{E}}(t|e) := \int_{\underline{q}}^{\bar{q}} u(t)\, dw\Big(1 - F(q|e)\Big)$, a non-additive expectation.

An agent is more efficient when high effort yields to a larger increase in probability. For us, whether the EUT or the non-EUT agent is more efficient depends on the probability perception of the non-EUT agent. When the probabilities implied by the high and low effort are given large decision weights, the non-EUT agent is more efficient since the benefits of exerting high effort are inflated. The first menu of contracts in Proposition 6 corresponds to this case and is designed to disincentivize the non-EUT agent to mimic the EUT agent. This is accomplished by offering a contract that, according to his judgement, equalizes the benefits that would have been obtained from mimicking.

By contrast, when the non-EUT agent's actions yield probabilities located at the insensitivity region, exerting high or low effort leads to probabilities that are almost indistinguishable for him, reducing his expected gains from exerting high effort. The EUT agent is thus more efficient. The second menu of contracts in Proposition 6 is designed to offer the EUT agent a contract that makes him indifferent between working or not. Such a contract ensures that the EUT agent is disincentivized from mimicking the non-EUT agent.

The principal takes advantage of the biased expectations of the non-EUT agent to screen. For instance, a contract that makes the EUT agent indifferent is perceived by the non-EUT agent to be profitable even when, based on objective expectations, this contract is designed to make both agents indifferent. To make the non-EUT agent indifferent between mimicking or not, his contract must include what he considers rents, administered in accordance to his subjective expectation. In other words, the principal can offer a cost-equivalent contract but with incentives concentrated where it matters for the agent—at extreme performance realizations.

That the solution to this problems depends on how the agent's action translates to probabilities is novel and is at odds with standard solutions to the hidden action problem. This dependence emphasizes the importance of the difficulty of the delegated task, which determines how effort translates into probability. For tasks in which the high and low efforts correspond to probabilities that are located at one or both extremes of the performance interval, for example easy or difficult tasks, the principal needs to ensure that the non-EUT agent has no incentives to mimic. Instead, when efforts lead to intermediate probabilities the contracts need to ensure that the EUT agent has no incentives to mimic the non-EUT agent.

5.3 Additional extensions

Here we briefly discuss a few additional extensions and emphasize how they derive from our previous analyses.

Disappointment Aversion

According to Baillon et al. (2020), disappointment aversion models with linear consumption utility can be written as

$$\int_{\underline{q}}^{\bar{q}} t(q)\,\mathrm{d}F(q|e) + \int_{\underline{q}}^{\bar{q}} CPT(t,e,r)\,\mathrm{d}R(q|e), \tag{11}$$

where $R(q|e)$ is the probability measure corresponding to the reference point and there are no probability distortion, $w(p) = p$. When r is replaced by the agent's expectation of $t(q)$ we obtain the models of Bell (1985) and Loomes and Sugden (1986); when r is replaced by the agent's certainty-equivalent of $t(q)$ we obtain the model of Gul (1991).

Equation (11) shows that risk attitudes are entirely determined by the second expression. Therefore, diminishing sensitivity and loss aversion entirely determine the optimal incentives. The optimal contract must therefore have a similar shape as the one presented in Proposition 5 with the additional assumption that the agent exhibits no probability distortion. We obtain an option-like incentives scheme: it is performance-insensitive first and increasing afterwards. The linear consumption utility in the first expression of equation (11) shifts the location of the kink toward lower output levels.

Ambiguity

We can also relax the assumption that the principal and the agent know the probabilities. In this setting, both parties know that performance may take any value in the set $q \in [\underline{q}, \bar{q}]$, but they do not know the exact distribution. A first approach is to consider probability sophistication (Machina and Schmeidler, 1992): there exists a probability measure P on $[\underline{q}, \bar{q}]$ such that $t(q)$ is evaluated by the agent as a probability-contingent pay schedule. Probabilities from P may be subjective. The objective function of the RDU agent is now

$$RDU(t,e) = \int_{\bar{q}}^{\underline{q}} u\Big(t(q)\Big)\,\mathrm{d}w\Big(1 - P(q|e)\Big) - c(e). \tag{12}$$

Then, the solution to the maximization problem of the agent is equivalent to those presented in Propositions 2 to 4, with the difference that probabilities assigned to output are now subjective.

The condition of probabilistic sophistication is stringent and invalidated by robust empirical phenomena such as the "home bias" or the Ellsberg paradox. A less stringent condition made in non-EU models is that probabilistic sophistication holds within sources of uncertainty, but not necessarily between sources of uncertainty (Abdellaoui et al., 2011; Chew and Sagi, 2008). In our setting, this amounts to assuming a uniform degree of ambiguity for the source that determines performance. If one is ready to assume that the agent's effort does not alter the degree of ambiguity from which performance originates, then probabilistic sophistication holds within this source and the optimal contract is as in in Propositions 2 to 4.[12]

Adverse selection with more types

In Section 5.2 we have considered moral hazard preceded by adverse selection where the principal has to contract with two types: likelihood-insensitive non-EUT agents and EUT agents. A possible extension is to consider a setting in which the principal contracts with more types. For instance, there could be EUT agents, pessimists, optimists, and likelihood-insensitive agents.

As in Proposition 6, the solution to this problem consists of offering a menu of contracts wherein a contract targets each type. To perform screening, it is necessary to establish which agents are more efficient to include an informational rent in their contract. If it is assumed, as we did before, that the likelihood-insensitive agent is also pessimistic, there is an efficiency ranking at intermediate and large probabilities: the optimist is the most efficient agent, the EUT agent is the second-most efficient, and the pessimist is the least efficient. For small probabilities this ranking changes and, depending on how severe is the likelihood insensitivity, the likelihood-insensitive agent

[12]If we instead assume that effort changes how ambiguous sources are, one needs to specify different weighting functions for each source, which makes the problem very different from the one we have solved. For example it is necessary to specify which source is more ambiguous, make assumptions on how stronger ambiguity affects the weighting function, and solve for an incentive compatibility constraint that contains two weighting functions.

becomes the most efficient or the second-most efficient.

These rankings determine the magnitude of the informational rents included in the contracts. For instance, for intermediate and large probabilities, the contract targeting the pessimist must not include any rents, whereas the contract targeting the optimist includes the largest rent, and the one given to the likelihood insensitive, the second-largest. Thus, the solution to this problem, as the one presented in Proposition 4, crucially depends on how the agent's actions translate into probabilities.

Non-EUT principal

Assume now that the principal also evaluates probabilities non-linearly. We maintain the assumption that her utility is linear. These assumptions imply that, while she is able to pool large numbers of independent risks, she exhibits probability sensitivity.

This modified problem is equivalent to the one solved in Sections 3 and 4. For the principal, fully informed about the risk attitudes of the agent, the agent's probability weighting function exhibits either more pessimism or more optimism than her own in the sense of Definition 3, and more or less likelihood insensitivity in the sense of Definition 6. For example, when the principal is pessimistic, an EUT agent is, in her eyes, optimistic. Similarly, when the principal exhibits moderate likelihood insensitivity, an agent with sever likelihood insensitivity exhibits, for her, a moderate degree of insensitivity. These examples illustrate how the refinements to the optimal contracts presented in Corollaries 3, 4 and 5 solve this modified problem.

6 Conclusion

We show that the optimal implementation of incentives crucially depends on the agent's perception of probabilities. Strong motivational deviations from expected utility, i.e. optimism and pessimism, generate contracts that cannot or do not require incentive compatibility. This leads to solutions to the principal's program that are often observed in practice such as long-shot contracts and fixed salaries. Moderate motivational deviations from EUT lead to optimal contracts that concentrate incentives at extremes: option-like incentive schemes and contracts with "only punishments". Finally, cognitive deviations from EUT lead to optimal contracts that can include these two features, combining thus strong punishments and rewards at extremes.

This paper opens avenues for future research. An interesting extension would be to consider a dynamic setting. Typically, the solution to the principal's program in a dynamic setting requires that contracts exhibit the "Martingale Property". However, when agents aggregate risk using non-additive probability measures, it is unclear how this property must be adapted to ensure that the contract is optimal. Another interesting investigation would be to test empirically the validity of our findings. One could for example design a lab experiment in which preference is elicited and on that basis contracts are imparted, against a situation in which contracts are just given according to the standard solution.

References

Abdellaoui, Mohammed, "Parameter-Free Elicitation of Utility and Probability Weighting Functions," *Management Science*, 2000, *46* (11), 1497–1512.

_ **, Aurélien Baillon, Laetitia Placido, and Peter P Wakker**, "The Rich Domain of Uncertainty: Source Functions and Their Experimental Implementation," *American Economic Review*, 2011, *101* (2), 695–723.

Baillon, Aurélien, Han Bleichrodt, and Vitalie Spinu, "Searching for the reference point," *Management Science*, 2020, *66* (1), 93–112.

Barseghyan, Levon, Francesca Molinari, Ted O'Donoghue, and Joshua C Teitelbaum, "The nature of risk preferences: Evidence from insurance choices," *American Economic Review*, 2013, *103* (6), 2499–2529.

Bell, David E, "Disappointment in decision making under uncertainty," *Operations research*, 1985, *33* (1), 1–27.

Bolton, Patrick and Mathias Dewatripont, *Contract Theory*, MIT Press, 2005.

Bombardini, Matilde and Francesco Trebbi, "Risk aversion and expected utility theory: an experiment with large and small stakes," *Journal of the European Economic Association*, 2012, *10* (6), 1348–1399.

Borch, Karl, "The safety loading of reinsurance premiums," *Scandinavian Actuarial Journal*, 1960, *1960* (3-4), 163–184.

Bruhin, Adrian, Helga Fehr-Duda, and Thomas Epper, "Risk and Rationality: Uncovering Heterogeneity in Probability Distortion.," *Econometrica*, 2010, *78* (4), 1375–1412.

Chateauneuf, A, J Eichberger, and S Grant, "Choice under Uncertainty with the Best and Worst in Mind: Neo-additive Capacities," *Journal of Economic Theory*, 2007, *137* (1), 538–567.

Chew, Soo Hong and Jacob S Sagi, "Small worlds: Modeling attitudes toward sources of uncertainty," *Journal of Economic Theory*, 2008, *139* (1), 1–24.

De La Rosa, Enrique Leonidas, "Overconfidence and Moral Hazard," *Games and Economic Behavior*, 2011, *73* (2), 429–451.

Einhorn, Hillel J and Robin M Hogarth, "Ambiguity and uncertainty in probabilistic inference.," *Psychological review*, 1985, *92* (4), 433.

Fehr-Duda, Helga and Thomas Epper, "Probability and risk: Foundations and economic implications of probability-dependent risk preferences," *Annual Review of Economics*, 2011, *4*, 567–593.

Goldstein, W M and H J Einhorn, "Expression theory and the preference reversal phenomena," *Psychological Review*, 1987, *94* (2), 236.

Gonzalez-Jimenez, Victor, "Incentive contracts when agents distort probabilities," *mimeo*, 2020.

Gonzalez, Richard and George Wu, "On the Shape of the Probability Weighting Function a," *Cognitive Psychology*, 1999, *166*, 129–166.

Gul, Faruk, "A Theory of Dissapointment Aversion," *Econometrica*, 1991, *59*, 667–686.

Herweg, By Fabian, Daniel Müller, and Philipp Weinschenk, "Binary Payment Schemes : Moral Hazard and Loss Aversion," *The American Economic Review*, 2010, *100* (5), 2451–2477.

Holmstrom, Bengt, "Moral hazard and observability," *Bell Journal of Economics*, 1979, *10* (1), 74–91.

_ **and Paul Milgrom**, "Aggregation and linearity in the provision of intertemporal incentives," *Econometrica: Journal of the Econometric Society*, 1987, pp. 303–328.

Jullien, Bruno and Bernard Salanié, "Estimating preferences under risk: The case of racetrack bettors," *Journal of Political Economy*, 2000, *108* (3), 503–530.

Kahneman, Daniel and Amos Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 1979, *47* (2), 263–291.

_ **, Jack L Knetsch, and Richard H Thaler**, "Anomalies The Endowment Effect , Loss Aversion , and Status Quo Bias," *Journal of Economic Perspectives*, 1991, *5* (1), 193–206.

Koszegi, Botond, "Behavioral Contract Theory," *Journal of Economic Literature*, 2014, *52*, 1075–1118.

_ **and Matthew Rabin**, "A Model of Reference-Dependent Preferences," *The Quarterly Journal of Economics*, 2006, *121* (4), 1133–1165.

Laffont, Jean-Jacques and David Martimort, *The Theory of Incentives*, Princeton University Press, jul 2002.

Lazear, Edward P and Paul Oyer, "Personnel economics," Technical Report, National Bureau of economic research 2007.

Loomes, Graham and Robert Sugden, "Disappointment and dynamic consistency in choice under uncertainty," *The Review of Economic Studies*, 1986, *53* (2), 271–282.

Machina, Mark J and David Schmeidler, "A more robust definition of subjective probability," *Econometrica: Journal of the Econometric Society*, 1992, pp. 745–780.

Mirrlees, James A, "The optimal structure of incentives and authority within an organization," *The Bell Journal of Economics*, 1976, pp. 105–131.

Myerson, Roger B, "Optimal auction design," *Mathematics of operations research*, 1981, *6* (1), 58–73.

Prelec, Drazen, "The Probability Weighting Function," *Econometrica*, 1998, *66* (3), 497–527.

Prendergast, Canice, "The Provision of Incentives in Firms," *Journal of Economic Literature*, 1999, *37* (1), 7–63.

Quiggin, John, "A Theory of Anticipated Utility," *Journal of Economic Behavior and Organization*, 1982, *3* (4), 323–343.

Salanié, Bernard, "Testing contract theory," *CESifo Economic Studies*, 2003, *49* (3), 461–477.

Santos-Pinto, Luıs, "Positive Self-Image and Incentives in Organisations.," *Economic Journal*, 2008, *118* (531), 1315–1332.

Snowberg, Erik and Justin Wolfers, "Explaining the favorite-long shot bias: Is it risk-love or misperceptions?," *Journal of Political Economy*, 2010, *118* (4), 723–746.

Spalt, Oliver, "Probability Weighting and Employee Stock Options," *The Journal of Financial and Quantitative Analysis*, 2013, *48* (4), 1085–1118.

Spinnewijn, Johannes, "Insurance and Perceptions: How to Screen Optimisms and Pessimists.," *Economic Journal*, 2013, *123* (569), 606–633.

Starmer, Chris, "Developments in Non-Expected Utility Theory: The Hunt for a Descriptive Theory of Choice under Risk," *Journal of Economic Literature*, June 2000, *38* (2), 332–382.

Tversky, Amos and Craig R Fox, "Weighing risk and uncertainty.," *Psychological Review*, 1995, *102* (2), 269–283.

_ **and Daniel Kahneman**, "The Framing of Decisions and the Psychology of Choice," *Science*, 1981, *211* (4481), 453–458.

_ **and** _ , "Advances in prospect theory: Cumulative representation of uncertainty," *Journal of Risk and Uncertainty*, 1992, *5* (4), 297–323.

_ **and Peter P. Wakker**, "Risk Attitudes and Decision Weights," *Econometrica*, 1995, *63* (6), 1255–1280.

Wakker, Peter P., "Testing and Characterizing Properties of Nonadditive Measures through Violations of the Sure-Thing Principle," *Econometrica*, 2001, *69* (4), 1039–1059.

_ , *Prospect Theory for Risk and Ambiguity*, Cambridge University Press, 2010.

Wu, G. and R. Gonzalez, "Curvature of the Probability Weighting Function," *Management Science*, 1996, *42* (12), 1676–1690.

Appendix A: Proofs

Lemma 1

Proof. When the agent's action is observable or contractible, $t(q)$ can be made contingent on e. Hence, the optimal contract is the solution to the program:

$$\max_{\{t(q)\}} \int_{\underline{q}}^{\bar{q}} \Big(S(q) - t(q)\Big) f(q|\bar{e})\,\mathrm{d}q$$

$$\text{s.t.} \quad \int_{\underline{q}}^{\bar{q}} u(t(q))w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})\,\mathrm{d}q - c \geq \bar{U},$$

where $\bar{U}$ is the agent's reservation utility. Denoting by ν the Lagrange multiplier for the agent's participation constraint, the Lagrangian of the problem writes as:

$$\begin{aligned}\mathcal{L}(q,t) =& \Big(S(q) - t(q)\Big) f(q|\bar{e}) \\ &+ \nu\Big[u(t(q))w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e}) - \bar{U} - c\Big].\end{aligned}$$

Pointwise optimization with respect to $t(q)$ yields

$$f(q|\bar{e}) = \nu u'(t_{FB}(q))w'\left(1 - F(q|\bar{e})\right) f(q|\bar{e}) \tag{13}$$

and, after re-arranging the previous expression we get,

$$\frac{1}{u'(t_{FB}(q))w'\left(1 - F(q|\bar{e})\right)} = \nu. \tag{14}$$

Under assumptions $u' > 0$ and $w' > 0$ Equation (14) shows that $\nu > 0$ under assumptions $u' > 0$ and $w' > 0$. Hence, the first-best contract must satisfy equation (14).

To investigate the shape of $t_{FB}(q)$ we differentiate (13) with respect to q, giving us the expression

$$t'_{FB}(q) = \frac{u'(t_{FB}(q))}{u''(t_{FB}(q))} \frac{w''(1 - F(q|\bar{e}))}{w'(1 - F(q|\bar{e}))} f(q|\bar{e}). \tag{15}$$

When the agent is EUT, w displays $w' = 1$ and $w'' = 0$. These properties imply that the right hand side (RHS) of (15) is zero and the first-best contract offered to these agents must be everywhere constant.

When the agent is optimistic, w exhibits $w' > 0$ and $w'' < 0$. The RHS of equation (15) is strictly positive and $t_{FB}(q)$ exhibits $t'_{FB}(q) > 0$ everywhere. To further understand the shape of the incentives scheme given to the optimist agent, we investigate the behavior of equation (15) at the endpoints of the output space. Since $\lim_{q \to \bar{q}} w'(1 - F(q|e)) = \infty$, it must be, due to equation (15), that $\lim_{q \to \bar{q}} t'_{FB}(q) = 0$. Moreover, due to $\lim_{q \to \underline{q}} w'(1 - F(q|e)) = 0$, we obtain $\lim_{q \to \underline{q}} t'_{FB}(q) = \infty$.

For an agent with pessimism, w displays $w' > 0$ and $w'' > 0$. Implying that the RHS of equation (15) is strictly negative and $t'_{FB}(q) < 0$. However, offering such incentive scheme generates a crucial problem: it motivates the agent to destroy output in order to attain the highest payment. These incentives are obviously undesirable. To solve this problem the principal could offer a contract with $t'_{FB}(q) > 0$, but she would incur in an inefficient expenditure since she would be providing large rewards to outcomes that are perceived as unlikely. The other solution consists on eliminating risk by offering a contract with $t'_{FB}(q) = 0$. Such contract eliminates risk, and thus pessimism. Moreover, since we already established that the participation constraint is binding at the solution, this constant contract must ensure that the agent exactly attains his reservation utility.

■

Proposition 1

See Holmstrom (1979).

Remark 1

Proof. The W-MLRP can be written as

$$\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{w'\left(1-F(q|\underline{e})\right)f(q|\underline{e})}{w'\left(1-F(q|\bar{e})\right)f(q|\bar{e})}\right)=\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{f(q|\underline{e})}{f(q|\bar{e})}\right)\left(\frac{w'\left(1-F(q|\underline{e})\right)}{w'\left(1-F(q|\bar{e})\right)}\right)+\frac{f(q|\underline{e})}{f(q|\bar{e})}\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{w'\left(1-F(q|\underline{e})\right)}{w'\left(1-F(q|\bar{e})\right)}\right). \quad (16)$$

The first expression in the RHS of equation (16) states that the non-EUT agent weights the standard MLRP, $\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{f(q|\underline{e})}{f(q|\bar{e})}\right)$, with the ratio $\frac{w'(1-F(q|\underline{e}))}{w'(1-F(q|\bar{e}))}$. Since $w'>0$ for all q and e, this expression is negative despite the agent exhibiting pessimism or optimism.

Next, we examine the sign of the second expression in the RHS of equation (16). This expression can be written as:

$$\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{w'\left(1-F(q|\underline{e})\right)}{w'\left(1-F(q|\bar{e})\right)}\right)=-\frac{w''\left(1-F(q|\underline{e})\right)w'\left(1-F(q|\bar{e})\right)f(q|\underline{e})}{\left(w'\left(1-F(q|\bar{e})\right)\right)^2}+\frac{w''\left(1-F(q|\bar{e})\right)w'\left(1-F(q|\underline{e})\right)f(q|\bar{e})}{\left(w'\left(1-F(q|\bar{e})\right)\right)^2}. \quad (17)$$

Thus, for the W-MLRP to hold, it suffices that when $w''>0$:

$$\frac{w''(1-F(q|\bar{e}))f(q|\bar{e})}{w'(1-F(q|\bar{e}))}<\frac{w''(1-F(q|\underline{e}))f(q|\underline{e})}{w'(1-F(q|\underline{e}))}. \quad (18)$$

Equation (18) states that w must exhibit more relative curvature, that is more relative convexity, at probabilities generated by e_L than at probabilities generated by e_H. This is a property that can be fulfilled by strictly convex probability weighting functions, with higher values of the second derivative at higher probabilities, e.g. p^3.

Under $w''<0$, the W-MLPR holds as long as

$$\frac{w''(1-F(q|\bar{e}))f(q|\bar{e})}{w'(1-F(q|\bar{e}))}>\frac{w''(1-F(q|\underline{e}))f(q|\underline{e})}{w'(1-F(q|\underline{e}))}. \quad (19)$$

Equation (19) shows that the probability weighting function of the optimist should exhibit less relative curvature, that is less relative concavity, at the ranks generated by e_L, than at the ranks generated by e_H. This is a property that can be fulfilled by strictly concave probability weighting functions, which exhibit higher values in the second derivative at lower probabilities, e.g. p^1.

■

Proposition 2 & Proposition 3

Proof. The Principal's maximization program is:

$$\max_{t(q)} \quad \int_{\underline{q}}^{\bar{q}} \Big(S(q) - t(q)\Big) f(q|\bar{e}) \, \mathrm{d}q$$

$$\text{s.t.} \quad \int_{\underline{q}}^{\bar{q}} u(t(q)) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e}) \, \mathrm{d}q - c \geq \int_{\underline{q}}^{\bar{q}} u(t(q)) w'\Big(1 - F(q|\underline{e})\Big) f(q|\underline{e}) \, \mathrm{d}q,$$

$$\int_{\underline{q}}^{\bar{q}} u(t(q)) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e}) \, \mathrm{d}q - c \geq \bar{U}.$$

Denoting by ν and μ the Lagrange multipliers for the agent's participation constraint and the agent's incentive compatibility constraint, respectively, the Lagrangian of the problem writes as:

$$\begin{aligned} \mathcal{L}(q,t) = &\Big(S(q) - t(q)\Big) f(q|\bar{e}) \\ &+ \mu \Big[u(t(q)) \Big(w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e}) - w'\Big(1 - F(q|\underline{e})\Big) f(q|\underline{e}) \Big) - c \Big] \\ &+ \nu \Big[u(t(q)) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e}) - \bar{U} - c \Big]. \end{aligned}$$

Pointwise optimization with respect to $t(q)$ yields

$$\begin{aligned} -f(q|\bar{e}) + \mu \Big[u'(t_{SB}(q))(w'(1 - F(q|\bar{e}))f(q|\bar{e}) - w'(1 - F(q|\underline{e}))f(q|\underline{e})) \Big] \\ + \nu u'(t_{SB}(q)) w'(1 - F(q|\bar{e})) f(q|\bar{e}) = 0, \end{aligned} \tag{20}$$

and, after re-arranging,

$$\frac{1}{u'(t_{SB}(q))w'\left(1-F(q|\bar{e})\right)} = \nu + \mu\left(1 - \frac{w'\left(1-F(q|\underline{e})\right)f(q|\underline{e})}{w'\left(1-F(q|\bar{e})\right)f(q|\bar{e})}\right). \tag{21}$$

We first show the conditions under which the incentive compatibility (IC) constraint binds at the optimum. Suppose the IC does not bind, then the solution to the principal's problem is the contract $t(q)_{FB}$ presented in Lemma 1. Consider first the case in which w exhibits $w'' > 0$. Denote the first-best contract given to this pessimistic agent by $t(q)^{P}_{FB}$ and recall that this contract is constant in q. We have thus that:

$$\int_{\underline{q}}^{\bar{q}} u\left(t(q)^{P}_{FB}\right) w'(1-F(q|\bar{e}))f(q|\bar{e})\,\mathrm{d}q - c = \int_{\underline{q}}^{\bar{q}} u\left(t(q)^{P}_{FB}\right) w'(1-F(q|\underline{e}))f(q|\underline{e})\,\mathrm{d}q. \tag{22}$$

The above equation implies that the IC constraint becomes slack only if $c < 0$, which is a contradiction. It must be then that $\mu > 0$ for an agent with $w'' > 0$ and the second-best contract is given by $t_{SB}(q)$ satisfying (21).

Next, consider an agent with $w'' < 0$ and denote by $t(q)^{O}_{FB}$ his first-best contract. According to Lemma 1, $t(q)^{O}_{FB}$ strictly increases in q. Thus, the IC constraint is slack only if:

$$\int_{\underline{q}}^{\bar{q}} u\left(t(q)^{O}_{FB}\right) w'\left(1-F(q|\bar{e})\right)f(q|\bar{e})\mathrm{d}q - \int_{\underline{q}}^{\bar{q}} u\left(t(q)^{O}_{FB}\right) w'\left(1-F(q|\underline{e})\right)f(q|\underline{e})\mathrm{d}q > c. \tag{23}$$

That $t'(q)^{O}_{FB} > 0$ together with the W-MLRP, that high effort yields higher probability of better outcomes, imply that (23) can hold for small values of c. Specifically, since w' and u are continuous, there must exist a threshold cost, $\hat{c} > 0$ such that the inequality in equation (23) holds whenever $c < \hat{c}$. For those cost levels $\mu = 0$. Conversely, if $c \geq \hat{c}$ the IC does not slack and $\mu > 0$.

Next, we show that the threshold $\hat{c}$ increases on the agent's optimism, making $c < \hat{c}$ less stringent. The Definition 3 states that more optimism is represented by a stronger concavity of the weighting function. Hence, under higher optimism more weight is given to high performance realizations. This together with the W-MLPR and the fact that $t'(q)^O_{FB} > 0$ imply that more optimism enlarges the left-hand side of (23), increasing the values of c that make the IC slack and $\mu = 0$. All in all, severe optimism entails that the IC is less likely to bind at the optimum and that $t(q)^O_{SB} = t(q)^O_{FB}$. Instead, moderate optimism imply a contract $t(q)^O_{SB}$ satisfying equation (21).

The second part of the proof analyzes the shape of the second best contract when $\mu > 0$. To that end, we differentiate equation (21) with respect to q to obtain:

$$
\begin{aligned}
t'(q)_{SB} =& \frac{u'(t(q)_{SB})w''(1-F(q|\bar{e}))}{u''(t(q)_{SB})w'(1-F(q|\bar{e}))} f(q|\bar{e}) \\
&+ \mu \frac{w'(1-F(q|\bar{e}))\, u'(t(q)_{SB})^2}{u''(t(q)_{SB})} \frac{\mathrm{d}}{\mathrm{d}q} \left(\frac{f(q|\underline{e})w'(1-F(q|\underline{e}))}{f(q|\bar{e})w'(1-F(q|\bar{e}))} \right). \quad (24)
\end{aligned}
$$

The first expression in the RHS of (24) is identical to the RHS of (15) in Lemma 1, which determined the shape of $t(q)_{FB}$. Recall from the proof that Lemma that this expression is positive when $w'' < 0$ and negative if $w'' > 0$. The second expression on the RHS of equation (24) incorporates the W-MLPR, along with $u' > 0$,$w' > 0$, and $u'' < 0$. That expression is always positive regardless of whether the agent displays $w'' > 0$ or $w'' < 0$.

Hence, $t'(q)_{SB} > 0$ if $w'' < 0$, since both expressions in the RHS of (24) are positive. Instead, when $w'' > 0$ the sign of $t'(q)_{SB}$ depends on the relative magnitude of these two expressions, which, as it will be shown below, depends on the degree of pessimism of the agent and the probabilities implied by e.

To further understand the shape of $t(q)_{SB}$, we study the behavior of equation (24) at extremes of the output interval. Consider first the case $w'' < 0$. Under $c < \hat{c}$, it was established that $t^O_{SB} = t^O_{FB}$ and the properties of this contract are identical to those described in Lemma 1.Let instead $c > \hat{c}$. As $q \to \bar{q}$ then $\lim_{q\to\bar{q}} w'(1-F(q|e)) = \infty$, making the first expression at the RHS of equation (24) negligible while the second expression tends to infinity. Hence, $\lim_{q\to\bar{q}} t'(q)^O_{SB} = \infty$. Moreover, as $q \to \underline{q}$, $\lim_{q\to\underline{q}} w'(1-F(q|e)) = 0$, which implies that the second expression in equation (24) becomes negligible and the first expression determines the shape of the second-best contract. Therefore, $t(q)_{SB}$ when $w'' < 0$ increases everywhere, exhibits steep payment increments at the highest end of the output interval, and resembles the first-best contract at the lowest end of the output interval.

Consider now $w'' > 0$. As $q \to \underline{q}$ we have $\lim_{q \to \underline{q}} w'\left(1 - F(q|e)\right) = \infty$, making the first term in the RHS of equation (24) negligible while the second explodes, yielding $t'(q)^{P}_{SB} = \infty$. Moreover, equation (24) shows that if either μ or w' tend to zero, then $t'(q)^{P}_{SB} < 0$, since the first expression, which is negative, gains importance while the second expression, which is positive, losses importance. This is indeed what happens as $q \to \bar{q}$ where $\lim_{q \to \bar{q}} w'\left(1 - F(q|e)\right) = 0$, implying $t'(q)^{P}_{SB} < 0$. Hence, $t(q)^{P}_{SB} < 0$ must be first decreasing to then become increasing increasing at the high end of the output interval.

We next show how more pessimism leads to a larger interval at which the solution exhibits $t'(q)^{P}_{SB} < 0$. According to Definition 3, more pessimism entails stronger convexity, which translates into the agent assigning more weight to low outcomes at the expense of assigning less weight to large outcomes. Thus, more pessimism entails $w'\left(1 - F(q|e)\right) = 0$ being assigned at a larger segment at the higher end of performance levels.

Furthermore, we show that μ also depends positively on the agent's degree of pessimism. Rewrite equation (21) as

$$\frac{1}{u'(t(q)_{SB})} = \nu w'\left(1 - F(q|\bar{e})\right) + \mu \left(w'\left(1 - F(q|\bar{e})\right) - w'\left(1 - F(q|\underline{e})\right) \frac{f(q|\underline{e})}{f(q|\bar{e})} \right). \tag{25}$$

Multiplying both sides of equation (25) by $f(q|\bar{e})$, integrating both sides over $[\underline{q}, \bar{q}]$, and recognizing that

$$\int_{\underline{q}}^{\bar{q}} w'\left(1 - F(q|\bar{e}\right) f(q|\bar{e}) \mathrm{d}q = 1, \tag{26}$$

gives us

$$\int_{\underline{q}}^{\bar{q}} \frac{1}{u'(t(q)_{SB})} f(q|\bar{e})\,\mathrm{d}q := \mathbb{E}\left(\frac{1}{u'(t_{SB}(q))}\right) = \nu. \tag{27}$$

The above equation yields that ν is the same for EUTs, optimists, and pessimists. Thus, $\nu^B = \nu^O = \nu^P$. Plugging equation (27) into equation (25) and multiplying by $f(q|\bar{e})\left(u(t_{SB}(q)) - \bar{U}\right)$ we obtain

$$\begin{aligned}
&\mu\left(w'(1-F(q|\bar{e}))f(q|\bar{e}) - w'\left(1-F(q|\underline{e})\right)f(q|\underline{e})\right)\left(u(t(q)_{SB}) - \bar{U}\right) \\
&= f(q|\bar{e})\left(u(t(q)_{SB}) - \bar{U}\right)\left(\frac{1}{u'(t(q)_{SB})} - \mathbb{E}\left(\frac{1}{u'(t(q)_{SB})}\right)w'(1-F(q|\bar{e}))\right).
\end{aligned} \tag{28}$$

If $t(q)_{SB}$ is a solution, we know from the Kuhn-Tucker complementary slackness conditions that

$$\begin{aligned}
\mu\Bigg(&\int_{\underline{q}}^{\bar{q}} u(t(q)_{SB})w'\left(1-F(q|\bar{e})\right)f(q|\bar{e})\,\mathrm{d}q \\
&- \int_{\underline{q}}^{\bar{q}} u(t(q)_{SB}))w'\left(1-F(q|\underline{e})\right)f(q|\underline{e})\,\mathrm{d}q - c\Bigg) = 0.
\end{aligned} \tag{29}$$

Hence, we rewrite (28), after integrating over $[\underline{q}, \bar{q}]$, as follows

$$\begin{aligned}
\mu c &= \int_{\underline{q}}^{\bar{q}} \left(u(t(q)_{SB}) - \bar{U}\right)\left(\frac{1}{u'(t(q)_{SB})} - \mathbb{E}\left(\frac{1}{u'(t(q)_{SB})}\right)w'(1-F(q|\bar{e}))\right)f(q|\bar{e})\,\mathrm{d}q \\
&= \int_{\underline{q}}^{\bar{q}} \frac{u(t(q)_{SB}) - \bar{U}}{u'(t(q)_{SB})} f(q|\bar{e})\,\mathrm{d}q - \mathbb{E}\left(\frac{1}{u'(t(q)_{SB})}\right)\int_{\underline{q}}^{\bar{q}} (u(t(q)_{SB}) - \bar{U})w'(1-F(q|\bar{e}))f(q|\bar{e})\,\mathrm{d}q \\
&= \mathbb{E}\left(\frac{u(t(q)_{SB}) - \bar{U}}{u'(t(q)_{SB})}\right) - \mathbb{E}\left(\frac{1}{u'(t(q)_{SB})}\right)\tilde{\mathbb{E}}\left(u(t(q)_{SB}) - \bar{U}\right)
\end{aligned} \tag{30}$$

where

$$\tilde{\mathbb{E}}\left(u(t(q)_{SB}) - \bar{U}\right) := \int\limits_{\underline{q}}^{\bar{q}} \left(u(t(q)_{SB}) - \bar{U}\right) w'(1 - F(q|\bar{e}))f(q|\bar{e})\,\mathrm{d}q. \quad (31)$$

Under EUT $w(p) = p$, so that $\tilde{\mathbb{E}}\left(u(t(q)_{SB}) - \bar{U}\right) = \mathbb{E}\left(u(t(q)_{SB}) - \bar{U}\right)$ and the third equality in (30) is equivalent to $cov\left(u(t(q)_{SB}, \frac{1}{u'(t(q)_{SB})}\right)$ which is strictly positive since due to the assumptions $u' > 0$ and $u'' < 0$.

Under pessimism $\tilde{\mathbb{E}}\left(u(t(q)_{SB}) - \bar{U}\right) < \mathbb{E}\left(u(t(q)_{SB}) - \bar{U}\right)$, since all probabilities are underweighted, and under optimism $\tilde{\mathbb{E}}\left(u(t(q)_{SB}) - \bar{U}\right) > \mathbb{E}\left(u(t(q)_{SB}) - \bar{U}\right)$, since all probabilities are overweighted. Therefore, equation (30) allows us to establish that $\mu^P > \mu^B > \mu^O \geq 0$. The fact that $\mu^O \geq 0$ holds with weak inequality confirms the previously mentioned result that the IC for this agent does not always bind and the degree of optimism determines if the IC binds.

Equations (30) and (31) show that $\mu^P > 0$ and that the stronger degree of pessimism of an agent is, i.e. more weight being given to low performance realizations, the larger μ^P becomes. As mentioned before, more pessimism also implies that $w' = 0$ is assigned to more values at the highest end of the output interval. Thus, according to (24) more pessimism implies that, $t(q)^P_{SB} > 0$ for low output levels, since μ has a large value and large values of w' are assigned to those levels, but also exhibits a larger performance interval whereby $t'(q)^P_{SB} < 0$, since $w' = 0$ is assigned to more values at the highest end of the output interval.

Finally, we modify the solution given by the first-order approach to guarantee monotonicity. Thus far we established that under pessimism, $t(q)^P_{SB}$, can be non-monotonic; increasing in q at low realizations and decreasing in q at high realizations. This shape of the second-best contract yields the problem that it generates an incentive for the agent to destroy output to obtain the largest possible payment. To avoid this problem, we modify the solution found by the first-best approach using "ironing"(Myerson, 1981).

Let $q_h \in (\underline{q}, \bar{q})$ be the output realization satisfying $q_h := max\{t^P_{SB}(q)\}$. The transfer implied by $t(q)^P_{SB}$ increases in q for the interval $q \in [\underline{q}, q_h]$ and decreases in q for the interval $q \in (q_h, \bar{q}]$. To implement ironing, find $q_{\mathcal{I}}$

satisfying:

$$\int_{q_{\mathcal{I}}}^{q_h} t(q)_{SB}^P \mathrm{d}q - \int_{q_h}^{\bar{q}} t(q)_{SB}^P \mathrm{d}q = 0. \tag{32}$$

There are two cases. If $\int_{\underline{q}}^{q_h} t(q)_{SB}^P \mathrm{d}q > \int_{q_h}^{\bar{q}} t(q)_{SB}^P \mathrm{d}q = 0$, there exists a performance level $q_{\mathcal{I}} \in [\underline{q}, q_h)$ such that equation (32) holds. For this case, ironing can be implemented and the modified solution is:

$$\tilde{t}(q)_{SB}^P = \begin{cases} t(q)_{SB}^P & \text{if } q \in [\underline{q}, q_{\mathcal{I}}), \\ t(q_{\mathcal{I}})_{SB}^P & \text{if } q \in [q_{\mathcal{I}}, \bar{q}]. \end{cases} \tag{33}$$

For this case, the modified solution $\tilde{t}(q)_{SB}^P$ is first strictly increasing in q, following the solution given by the first-order approach, and becomes flat after reaching the performance level $q_{\mathcal{I}}$.

Instead, when $\int_{\underline{q}}^{q_h} t(q)_{SB}^P \mathrm{d}q \leq \int_{q_h}^{\bar{q}} t(q)_{SB}^P \mathrm{d}q$ ironing cannot be implemented and the resulting incentives scheme, provided that the principal would like to avoid incentives that motivate output destruction, is constant everywhere, only satisfying the participation constraint. For this case, the principal faces an impossibility as no incentive-compatible scheme exists that also avoids destroying output.

Note that we already established that more pessimistic attitudes toward risk entail that the weighting function exhibits $w'(p) = 0$ for a larger interval at the highest-end of the output interval, which yields, according to (24), a larger interval where $t(q)_{SB}^P < 0$. This implies that stronger pessimism increases the likelihood that the resulting incentives scheme flat.

■

Corollary 1 & Corollary 2

Proof. The first part of the proof contrasts the second-best contract to the first-best contract to establish how the the principal is required to include rewards and punishments.

Equation (21) shows that rewards are given when:

$$\frac{w'\left(1-F(q|\underline{e})\right)}{w'\left(1-F(q|\bar{e})\right)}\frac{f(q|\underline{e})}{f(q|\bar{e})}<1, \tag{34}$$

which holds only if $\frac{f(q|\underline{e})}{f(q|\bar{e})}<1$ and $\frac{w'(1-F(q|\underline{e}))}{w'(1-F(q|\bar{e}))}<1$ are satisfied. These two conditions are met when large performance levels realize and under pessimism, respectively. Under optimism, which entails $\frac{w'(1-F(q|\underline{e}))}{w'(1-F(q|\bar{e}))}>1$, the inequality in (34) holds as long as $\frac{f(q|\underline{e})}{f(q|\bar{e})}$ is considerably smaller than one. Thus, higher performance realizations are required for optimists to receive rewards as compared to EUT agents and pessimists.

Equation (21) also shows that punishments are given when

$$\frac{w'\left(1-F(q|\underline{e})\right)}{w'\left(1-F(q|\bar{e})\right)}\frac{f(q|\underline{e})}{f(q|\bar{e})}>1, \tag{35}$$

which holds only if $\frac{f(q|\underline{e})}{f(q|\bar{e})}>1$ and $\frac{w'(1-F(q|\underline{e}))}{w'(1-F(q|\bar{e}))}>1$ are satisfied. These two conditions are met for low output realizations and under optimism, respectively. Under pessimism, which entails $\frac{w'(1-F(q|\underline{e}))}{w'(1-F(q|\bar{e}))}<1$, the inequality in (35) also holds as long as the ratio $\frac{f(q|\underline{e})}{f(q|\bar{e})}$ is considerably larger than one. Hence, punishments are given to optimists in the contingency of higher output realizations as compared to the EUT agents and pessimists.

The second part of the proof compares the incentives included in second-best contracts $t(q)^O_{SB}$,$t(q)^P_{SB}$, and the second-best contract given to the EUT agent. From the proof of Propositions 2 and 3 we know that that the Lagrangian multiplier ν corresponding to the participation constraint is the same across agents with different probability perception. We also established that the Lagrangian multiplier corresponding to the participation constraint exhibits the ordering $\mu^P>\mu^B>\mu^0\geq 0$. We rewrite the first-order condition presented in equation (21) as:

$$\frac{1}{u'(t(q)^{nB}_{SB})}=\nu w'\left(1-F(q|\bar{e})\right)+\mu^{nB}w'\left(1-F(q|\bar{e})\right)\left(1-\frac{w'\left(1-F(q|\underline{e})\right)f(q|\underline{e})}{w'\left(1-F(q|\bar{e})\right)f(q|\bar{e})}\right), \tag{36}$$

where $t_{SB}^{nB} \in \{t_{SB}^{O}, t_{SB}^{P}\}$ and $\mu^{nB} \in \{\mu^{O}, \mu^{P}\}$, the second-best contract and IC multiplier, respectively, corresponding optimists and pessimists. Instead, for the EUT agent the second-best contract satisfies:

$$\frac{1}{w'(t(q)_{SB}^{B})} = \nu + \mu^{B}\left(1 - \frac{f(q|\underline{e})}{f(q|\bar{e})}\right). \tag{37}$$

Comparison of equations (36) and (37) yields that $t(q)_{SB}^{nB} > t(q)_{SB}^{B}$ only if $w'(1 - F(q|\bar{e})) < 1$, $w'(1 - F(q|\bar{e})) < w'(1 - F(q|\underline{e}))$, and $\mu^{B} > \mu^{nB}$. These conditions hold for the optimist in the interval $q \in [\underline{q}, q^{*}]$, where q^{*} is the output level satisfying $w'(1 - F(q^{*}|\bar{e})) = 1$ for the optimist.

For $q \in [q^{*}, \bar{q}]$, where the optimist displays $w'(1 - F(q|\bar{e})) > 1$, $t(q)_{SB}^{O} > t(q)_{SB}^{B}$ can hold when $w'(1 - F(q|\bar{e})) - w'(1 - F(q|\underline{e})) < 0$ is negative enough, facilitating the following condition:

$$\begin{aligned}\mu^{B} - \mu^{O}w'(1 - F(q|\bar{e})) >& \nu(w'(1 - F(q|\bar{e}) - 1) + \\ & \left(\mu^{B} - \mu^{O}w'(1 - F(q|\underline{e}))\right)\frac{f(q|\underline{e})}{f(q|\bar{e})}.\end{aligned} \tag{38}$$

The difference $w'(1 - F(q|\bar{e})) - w'(1 - F(q|\underline{e})) < 0$ is more negative the more optimistic the worker is. If instead, this difference is not too negative, due to optimism being moderate, then $t_{SB}^{O} < t_{SB}^{B}$. Since we are considering a situation where $\mu^{O} > 0$ and the second-best contract is optimal, the latter situation is more likely.

Comparison of (36) and (37) yields that $t(q)_{SB}^{B} > t(q)_{SB}^{nB}$ only if $w'(1 - F(q|\bar{e})) > 1$, $w'(1 - F(q|\underline{e})) < w'(1 - F(q|\bar{e}))$, and $\mu^{NB} \geq \mu^{B}$. These conditions hold for the pessimist in the interval $q \in [\underline{q}, q^{*}]$ where q^{*} is the production level that guarantees $w'(1 - F(q^{*}|\bar{e})) = 1$ for the pessimist.

Alternatively, for $q \in [q^{*}, \bar{q}]$, where $w'(1 - F(q|\bar{e})) < 1$ for the pessimist, $t_{SB}^{B} > t_{SB}^{P}$ can hold. Specifically, if pessimism is severe enough so that the difference $w'(1 - F(q|\bar{e})) - w'(1 - F(q|\underline{e})) > 0$ is sufficiently large, the following condition is less stringent:

$$\mu^P w'(1-F(q|\bar{e}))-\mu^B > \nu(1-w'(1-F(q|\bar{e}))+\left(\mu^P w'(1-F(q|\underline{e})-\mu^B)\right)\frac{f(q|\underline{e})}{f(q|\bar{e})}. \tag{39}$$

Instead, when $w'(1-F(q|\bar{e}))-w'(1-F(q|\underline{e}))>0$ is small, it must be that $t(q)^B_{SB} < t(q)^P_{SB}$. Since we are considering a situation where incentives can be implemented for the pessimist, the latter situation is more likely.

■

Lemma 2

Proof. The problem is similar to that solved in the proof Lemma 1 with the difference that w is now an inverse-S function. Hence, we use the first the first-order condition in equation (13). Denote by $t(q)^{LI}_{FB}$, the contract satisfying (13) when w has an inverse S-shape.

First note from equation (14), which results from rearranging (13), that $\nu > 0$, the participation constraint binds at the optimum, due to the assumptions $u' > 0$ and $w' > 0$.

Equation (15), obtained by differentiating (13) with respect to q, shows that $t(q)^{LI}_{FB}$ decreases in $q \in [\underline{q}, \tilde{q})$, where $\tilde{q} \in [\underline{q}, \bar{q}]$ is a performance level satisfying $w(1-F(q|e))'' = 0$. That is because in that interval w exhibits $w'' < 0$ making (15) negative. While for the interval $q \in [\tilde{q}, \bar{q}]$ the first-best contract exhibits $t'(q)^{LI}_{FB} > 0$ inasmuch as w displays $w'' > 0$ in that interval.

It is informative to investigate how $t(q)^{LI}_{FB}$ behaves at the extremes of the output interval as well as at the inflection point $\tilde{q}$. Since, $\lim_{q\to\bar{q}} w' = \infty$ and $\lim_{q\to\underline{q}} w' = \infty$ equation (13) allows us to establish that $t'(q)^{LI}_{FB} = 0$ as $q \to \underline{q}$ and as $q \to \bar{q}$. Also, when the weighting function is evaluated at $\tilde{q}$ we obtain $t'(q)_{FB} = 0$ since at that point w displays $w'' = 0$.

That t^{LI}_{FB} is decreasing in $q \in [\underline{q}, \tilde{q}]$ is undesirable. To overcome the difficulties associated with a decreasing incentives scheme we iron the solution resulting from the first-best approach as in Proposition 3. To that end, find

$q_I \in (\tilde{q}, \bar{q}]$ such that

$$\int_{\underline{q}}^{\tilde{q}} t(q)_{FB}^{LI} \mathrm{d}q - \int_{\tilde{q}}^{q_I} t(q)_{FB}^{LI} \mathrm{d}q = 0. \tag{40}$$

There are two cases. When $\int_{\underline{q}}^{\tilde{q}} t(q)_{FB}^{LI} \mathrm{d}q < \int_{\tilde{q}}^{q_I} t(q)_{FB}^{LI} \mathrm{d}q$, there exists $q_I \in (\tilde{q}, \bar{q}]$ such that equation (40) holds. In such case the modified solution is:

$$\tilde{t}(q)_{FB}^{LI} = \begin{cases} t(q_I)_{FB}^{LI} & \text{if } q \in [\underline{q}, q_I]), \\ t(q)_{FB}^{LI} & \text{if } q \in [q_I, \bar{q}]. \end{cases} \tag{41}$$

In contrast, when $\int_{\underline{q}}^{\tilde{q}} t(q)_{FB}^{LI} \mathrm{d}q \geq \int_{\tilde{q}}^{q_I} t(q)_{FB}^{LI} \mathrm{d}q$, the level q_I does not exist. A possible solution in such case is to eradicate risk by paying a contract offering a fixed amount that makes the agents' utility exactly equal to his outside option $\bar{U}$.

Equation (40) shows that the existence of $q_I \in (\tilde{q}, \bar{q}]$ that irons the solution determined by the first-order approach is less stringent as $q \in [\underline{q}, \tilde{q}]$ is smaller, which amounts to a smaller region whereby the weighting function exhibits $w'' > 0$. Also, when w' is considerably larger over the segment $q \in [\tilde{q}, \bar{q}]$ than over $q \in [\underline{q}, \tilde{q}]$, the existence of a solution given by (41) is more likely.

■

Proposition 4

Proof. The principal's program is similar to that solved in Propositions 3 and 2, with the difference that w has now an inverse-s shape. Hence, the optimal contract offered t_{SB}^{LI} must satisfy the first-order condition presented in equation (21).

We first show that the incentive compatibility constraint (IC) binds unless likelihood insensitivity is accompanied by strong optimism. From equation (27) we conclude that ν is the same to the EUT agent and the likelihood insensitive agent. Furthermore, equations (30) and (31) demonstrate that $\mu^P > \mu^B > \mu^O$ and that $\mu^B \geq 0$. Since $\hat{p} \in (0, 1)$, then, either

$\mu^B > \mu^{LI}$ or $\mu^{LI} > \mu^B$, depending on whether optimism or pessimism accompanies likelihood insensitivity. To understand how pessimism and optimism determine the relationship between μ^{LI} and μ^B note that under likelihood insensitivity and pessimism:

$$||\int_{\underline{q}}^{\hat{q}} \left(u(t_{SB}(q)) - \bar{U}\right) w' \left(1 - F(q|\bar{e})\right) f(q|\bar{e})\mathrm{d}q - \int_{\underline{q}}^{\hat{q}} \left(u(t_{SB}(q)) - \bar{U}\right) f(q|\bar{e})\mathrm{d}q|| >$$
$$||\int_{\hat{q}}^{\bar{q}} \left(u(t_{SB}(q)) - \bar{U}\right) w' \left(1 - F(q|\bar{e})\right) f(q|\bar{e})\mathrm{d}q - \int_{\hat{q}}^{\bar{q}} \left(u(t_{SB}(q)) - \bar{U}\right) f(q|\bar{e})\mathrm{d}q||. \tag{42}$$

Hence, pessimism accompanying likelihood insensitivity implies $\tilde{\mathbb{E}}\left(u(t_{SB}(q)) - \bar{U}\right) < \mathbb{E}\left(u(t_{SB}(q)) - \bar{U}\right)$, and, due to equation (30), that $\mu^{LI} > \mu^B$.Since $\mu^B \geq 0$, as shown in the proof of Propositions 2 and 3, when (42) holds, the IC constraint of the pessimistic and likelihood insensitive agent binds at the optimum.

Instead, when optimism accompanies likelihood insensitivity the inequality in (42) cannot hold and $\mathbb{E}\left(u(t_{SB}(q)) - \bar{U}\right) \geq \tilde{\mathbb{E}}\left(u(t_{SB}(q)) - \bar{U}\right)$. Equation (30) establishes that for this case $\mu^B \geq \mu^{LI}$. Thus, with sufficient optimism, such that the non-additive expectation $\tilde{\mathbb{E}}\left(u(t_{SB}(q)) - \bar{U}\right)$ makes the RHS of (30) at most equal to zero, the IC slacks at the optimum. If that is the case, then $t_{SB}^{LI} = t_{FB}^{LI}$.

The second part of the proof studies the shape of the solution obtained by the first order approach, t_{SB}^{LI}. Under sufficiently strong optimism, we established that $t_{SB}^{LI} = t_{FB}^{LI}$ and the shape of the contract is described in Lemma 2. We instead focus on the shape of the solution satisfying (21) with $\mu > 0$.

Equation (24) presents the derivative of (21) with respect to q. From (24) it is evident that for the interval $q \in (\tilde{q}, \bar{q}]$, the solution exhibits $t'(q)_{SB}^{LI} > 0$ due to the $u'' < 0$,$u' > 0$, $w' > 0$, $w'' < 0$, together with the WMLRP. Additionally, $t(q)_{SB}^{LI} > 0$ becomes steeper as $q \to \bar{q}$, where, due to$\lim_{q \to \bar{q}} w' =$, the solution exhibits $t'(q)_{SB}^{LI} = \infty$.

For $q \in [\underline{q}\tilde{q})$ the sign of $t'(q)_{SB}^{LI}$ depends on the considered segment. This is evident from the two expressions in equation (24): the first expression is negative everywhere due to $w'' > 0$, while the second expression is positive due to the W-MLRP. The relative importance of these two expressions depends on w' which takes different values over the considered interval. When w' is sufficiently large the first expression losses importance, while the second gains importance. This is indeed what happens as $q \to \underline{q}$ where $\lim_{q \to \underline{q}} w' = \infty$ generating $t'(q)^{LI}_{SB} = \infty$. Instead, as $q \to \tilde{q}$, w' becomes smaller and the second expression in equation (24) losses importance with respect to the first expression. When $w' \to 0$, that is when insensitivity is strong, the first expression of (24) becomes relatively important yielding $t'(q)^{LI}_{SB} < 0$ when performance levels $q \to \tilde{q}$ from the left.

Thus far, we have found three possible shapes of the second-best contract. First, when optimism is strong the IC constraint is slack at the optimum and $t(q)^{LI}_{SB} = t(q)^{LI}_{FB}$. Second, when the IC binds and the agent exhibits moderate likelihood insensitivity, the first-order approach yields a solution that is always increasing and that exhibits steep payment increments at both extremes of the output interval. Third, when the IC binds and the agent exhibits severe insensitivity, so that $w' \to 0$ as $q \to \tilde{q}$, the incentives scheme has a decreasing slope for intermediate performance realizations but displays an increasing slope before and after this interval.

That $t(q)^{LI}_{SB}$ can exhibits a decreasing pay segment yields incentives to destroy output. To avoid these problematic incentives we iron the solution. Let $q_{\mathcal{M}} := max\left(t(q)^{LI}_{SB}\right)$ in $q \in [\underline{q}, \tilde{q}]$ and $q_{\mathcal{S}} := min\left(t(q)^{LI}_{SB}\right)$ in $q \in (q_{\mathcal{M}}, \tilde{q})$. Ironing $t(q)^{LI}_{SB}$ consists on finding $q_{\mathcal{I}1} \in [\underline{q}, q_{\mathcal{M}})$ and $q_{\mathcal{I}2} \in (q_{\mathcal{S}}, \bar{q}]$ such that

$$\int_{q_{\mathcal{I}1}}^{q_{\mathcal{M}}} t(q)^{LI}_{SB}\mathrm{d}q - \int_{q_{\mathcal{M}}}^{q_{\mathcal{S}}} t(q)^{LI}_{SB}\mathrm{d}q + \int_{q_{\mathcal{S}}}^{q_{\mathcal{I}2}} t(q)^{LI}_{SB}\mathrm{d}q = 0, \tag{43}$$

and

$$t^{LI}_{SB}(q_{\mathcal{I}1}) = t^{LI}_{SB}(q_{\mathcal{I}2}). \tag{44}$$

If such $q_{\mathcal{I}l}$ and $q_{\mathcal{I}2}$ exist, the ironed incentives scheme becomes:

$$\tilde{t}_{SB}^{LI}(q) = \begin{cases} t_{SB}^{LI}(q) & \text{if } q \in [\underline{q}, q_{\mathcal{I}1}) \cup [q_{\mathcal{I}2}, \bar{q}], \\ t_{SB}^{LI}(q_{\mathcal{I}1}) & \text{if } q \in [q_{\mathcal{I}1}, q_{\mathcal{I}2}]. \end{cases} \tag{45}$$

There are two boundary cases worth discussion. First, under strong optimism, but not sufficiently severe as to yield $\mu = 0$, most of the weighting function is concave and $\tilde{q}$ is thus located at low output levels. For this case it is possible that the ironing solution yields $q_{\mathcal{I}l} = \underline{q}$. Thus, the resulting incentives scheme becomes:

$$\tilde{t}_{SB}^{LI}(q) = \begin{cases} t_{SB}^{LI}(q) & \text{if } q \in [q_{\mathcal{I}2}, \bar{q}], \\ t_{SB}^{LI}(q_{\mathcal{I}1}) & \text{if } q \in [\underline{q}, q_{\mathcal{I}2}]. \end{cases} \tag{46}$$

Second, under severe pessimism, yielding that most of the weighting function is convex, $\tilde{q}$ is located at high output levels and the segment at which the solution implied by the first-best approach decreases is large. In such case, it is possible that the ironing solution yields $q_{\mathcal{I}2} = \bar{q}$. The resulting incentives scheme in such case is:

$$\tilde{t}_{SB}^{LI}(q) = \begin{cases} t_{SB}^{LI}(q) & \text{if } q \in [q_{\mathcal{I}1}, \bar{q}], \\ t_{SB}^{LI}(q_{\mathcal{I}1}) & \text{if } q \in [\underline{q}, q_{\mathcal{I}1}]. \end{cases} \tag{47}$$

Finally, when $q_{\mathcal{I}l}$ and $q_{\mathcal{I}2}$ do not exist, because $t(q)_{SB}^{LI}$ exhibits a sizeable interval with $t'(q)_{SB}^{LI} < 0$, the first-order solution cannot be ironed and incentive compatibility cannot be implemented without including perverse incentives. This case is more likely to emerge under severe likelihood insensitivity along with strong pessimism. In such case the principal can also implement an incentive scheme that is constant everywhere and that yields a welfare level such that the contract is accepted by the agent, but such contract is not incentive compatible.

■

Corollary 5

Proof. We first establish how $t(q)^{LI}_{SB}$ and $t(q)^{LI}_{FB}$ compare. In $q \in (\tilde{q}, \bar{q}]$, where $\tilde{q}$ is such that $w''(1 - F(\tilde{q}|\bar{e})) = 0$, the weighting function exhibits $w'' < 0$. Thus $\frac{w'(1-F(q|\underline{e}))}{w'(1-F(q|\bar{e}))} > 1$ in $q \in (\tilde{q}, \bar{q}]$. According to equation (21), rewards with respect to the first-best are given as long as $\frac{f(q|\underline{e})}{f(q|\bar{e})} < \frac{w'(1-F(q|\bar{e}))}{w'(1-F(q|\underline{e}))}$, which due to $\frac{w'(1-F(q|\underline{e}))}{w'(1-F(q|\bar{e}))} > 1$ happens at higher performance levels as compared to the performance levels at which the EUT agent receives rewards.

Moreover, the second-best contract specifies punishments with respect to the first-best at very low output levels. For $q \in [\underline{q}, \tilde{q}]$, the weighting function displays $w'' > 0$ which implies $\frac{w'(1-F(q|\underline{e}))}{w'(1-F(q|\bar{e}))} < 1$. Hence, punishments, emerge only if $\frac{f(q|\underline{e})}{f(q|\bar{e})} > \frac{w'(1-F(q|\bar{e}))}{w'(1-F(q|\underline{e}))}$, which, due to $\frac{w'(1-F(q|\underline{e}))}{w'(1-F(q|\bar{e}))} < 1$, happens at lower output realizations as compared to the performance levels at which the EUT agent receives punishments.

Next, we compare the strength of the incentives implied by t^{LI}_{SB} and t^{B}_{SB}. We established in Proposition 4 that $\nu = \nu^{LI} = \nu^{B}$. Furthermore, we also know that either $\mu^{LI} \geq \mu^{B}$, under pessimism , or $\mu^{LI} < \mu^{B}$ otherwise. We consider each of this cases separately.

Let $\mu^{B} > \mu^{LI}$. Comparison of equations (36) with (37) entails that $t(q)^{LI}_{SB} > t(q)^{B}_{SB}$ if $w' < 1$ and $w'(1 - F(q|\underline{e})) > w'(1 - F(q|\bar{e}))$. These conditions hold for the interval $q \in [\tilde{q}, q^{**}_h]$ where is a performance level $q^{**}_h \in (\tilde{q}, \bar{q})$ satisfying $w'(1 - F(q^{**}_h|\bar{e})) = 1$. Alternatively, when, $w' > 1$ and $w'(1 - F(q|\underline{e})) > w'(1 - F(q|\bar{e}))$, the ordering $t^{LI}_{SB} > t^{B}_{SB}$ can hold as long as the difference $w'(1 - F(q|\underline{e})) - w'(1 - F(q|\bar{e}))$ is sufficiently large. That is, when the concavity in the interval $q \in [\tilde{q}, \bar{q}]$ yields a sufficiently large difference between the ranks generated by high and low effort. This happens under strong likelihood insensitivity and optimism, both of which are consistent with the assumption $\mu^{B} > \mu^{LI}$.

Moreover, $t^{LI}_{SB} < t^{B}_{SB}$ can be obtained if $w' > 1$, $w'(1 - F(q|\underline{e})) < w'(1 - F(q|\bar{e}))$, and the difference $w'(1 - F(q|\underline{e})) - w'(1 - F(q|\bar{e}))$ is very negative. These conditions hold for $q \in [\underline{q}, q^{**}_l]$, where $q^{**}_l \in (\underline{q}, \tilde{q})$ is a performance level $w'(1 - F(q_l^{**}|\bar{e})) = 1$, and when $w'' > 0$. The latter condition can be odds with the assumption $\mu^B > \mu^{LI}$.

Consider now $\mu^{LI} > \mu^B$. Comparison of equations (36) with (37) entails that $t(q)_{SB}^{LI} < t(q)_{SB}^{B}$ if $w' > 1$ and $w'(1 - F(q\underline{e})) < w'(1 - F(q\bar{e}))$. These conditions hold in the interval $q \in [\underline{q}, q_l^{**}]$, where $q_l^{**} \in (\underline{q}, \tilde{q})$ is a performance level satisfying $w'(1 - F(q_l^{**}|\bar{e})) = 1$. When $w' < 1$ $t_{SB}^{L}I < t_{SB}^{B}$ holds as long as the difference $w'(1 - F(q|\underline{e})) - w'(1 - F(q|\bar{e}))$ is sufficiently negative, which happens under strong likelihood insensitivity and pessimism, both of which are consistent with the departing assumption $\mu^{LI} > \mu^B$.

Finally, $t_{SB}^{LI} > t_{SB}^{B}$ holds as long as $w' > 1$, $w'(1 - F(q|\underline{e})) > w'(1 - F(q|\bar{e}))$, and when the weighting function exhibits a sufficiently large degree of concavity. These conditions hold for the interval $q \in [q_h^{**}, \bar{q}]$ where $q_h^{**} \in (\tilde{q}, \bar{q})$ is such that $w'(1 - F(q^{**}|\bar{e})) = 1$. However, that the weighting function is too concave can be at odds with the initial assumption $\mu^{LI} > \mu^B$.

To summarize, we find that under $\mu^B > \mu^{LI}$, $t(q)_{SB}^{LI} > t(q)_{SB}^{B}$ in $q \in [\tilde{q}, q_h^{**}]$ as well as in the interval $[q_h^{**}, \bar{q}]$ when optimism is sufficiently strong. Also, $t_{SB}^{LI} < t_{SB}^{B}$ in $q \in [\underline{q}, q_l^{**}]$ when w exhibits strong convexity in that interval. Also we find that under $\mu^B < \mu^{LI}$, $t(q)_{SB}^{B} > t(q)_{SB}^{LI}$ in $q \in [\underline{q}, q_l^{**}]$ as well as in the interval $q \in [q_l^{**}, \tilde{q}]$ if pessimism is sufficiently strong. Moreover, $t_{SB}^{LI} > t_{SB}^{B}$ in $q \in [q_h^{**}, \bar{q}]$ when concavity is strong at that interval.

■

Proposition 5

Proof. Rewrite equation (4) using Assumption 5 as

$$CPT(t, e, r) = \int_{q_r}^{\bar{q}} v\Big(t(q) - r\Big) w'\Big(1 - F(q|e)\Big) f(q|e) \mathrm{d}q + \lambda \int_{\underline{q}}^{q_r} v\Big(r - t(q)\Big) w'\Big(1 - F(q|e)\Big) f(q|e) \mathrm{d}q - c(e) \tag{48}$$

At the global optimum either the first expression or the second expression of equation (48) are included in the participation and the incentive compatibility constraints. This depends on whether the considered segment of the solution is located in the domain of gains or in the domain of losses, respectively.

Consider first the case in which the solution is in the domain of gains, $q \in [q_r, \bar{q}]$. In this case, the principal's problem is solved with the second-best contract presented in Proposition 2, 3, or 4. The reason is that in the domain of gains CPT is identical to RDU; the agent's sensitivity to probabilities is represented by w and the weighting function and his value function exhibits $v' > 0$ and $v'' < 0$. Moreover, which of this solutions apply depends on the shape of w, e.g. when the CPT agent has a w with $w'' > 0$ everywhere, then the contract from Proposition 2applies.

Suppose now that the solution is in the domain of losses $q \in [q_r, \bar{q}]$. To analyze the shape of the second-best contract we solve the principal's program when the preferences of the agent reflect that he is in the domain of losses. Denoting by ν and μ the multipliers associated to the participation and the incentive compatibility constraints, the Lagrangian of the principal's problem can be written as

$$\begin{aligned}\mathcal{L}(q,t) =& \Big(S(q) - t(q)\Big) f(q|\bar{e}) \\ &+ \mu\Bigg(\lambda v\Big(r - t(q)\Big)\Big(w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e}) - w'\Big(1 - F(q|\underline{e})\Big) f(q|\underline{e})\Big) - c\Bigg) \\ &+ \nu\Bigg(\lambda v\Big(r - t(q)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e}) - c - \bar{U}\Bigg).\end{aligned} \tag{49}$$

Pointwise optimization with respect to $t(q)$, and some re-arrangements yield:

$$\begin{aligned}\frac{f(q|\bar{e})}{-\lambda v'\Big(r - t_{SB}\Big)} =& \nu\Bigg(w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})\Bigg) \\ &+ \mu\Bigg(w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e}) - w'\Big(1 - F(q|\underline{e})\Big) f(q|\underline{e})\Bigg).\end{aligned} \tag{50}$$

Integrating both sides over $[\underline{q}, \bar{q}]$ and recognizing that $\int_{\underline{q}}^{\bar{q}} w'\Big(1 - F(q|\bar{e})\Big) \mathrm{d}q = 1$, gives us:

$$\int_{\underline{q}}^{\bar{q}} \frac{f(q|\bar{e})}{-\lambda v'\Big(r - t_{SB}\Big)} = \mathbb{E}\left(\frac{1}{-\lambda v'\Big(r - t_{SB}\Big)}\right) = \nu \tag{51}$$

Equation (51) shows that $\nu \leq 0$ and the participation constraint is slack in the domain of losses. This result implies that the agent is reluctant to accept a contract that locates him in that domain.

Next, we show that a contract specifying $t(q) < r$ can be incentive compatible. We include the last equality in equation (51) in (50) and multiply both sides of that equation by $\lambda v\left(r - t_{SB}\right)$ to obtain:

$$\begin{aligned}
&\mu\Big(\lambda v\left(r - t_{SB}\right)\Big)\left(w'\left(1 - F(q|\bar{e})\right) f(q|\bar{e}) - w'\left(1 - F(q|\underline{e})\right) f(q|\underline{e})\right) = \\
&\Big(\lambda v\left(r - t_{SB}\right)\Big)\left(\frac{1}{-\lambda v'\Big(r - t_{SB}\Big)}\right) f(q|\bar{e}) \\
&- \Big(\lambda v\left(r - t_{SB}\right)\Big)\mathbb{E}\left(\frac{1}{-\lambda v'\Big(r - t_{SB}\Big)}\right) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e}).
\end{aligned} \tag{52}$$

If t_{SB} is a solution we know from the Kuhn-Tucker complementary slackness conditions that

$$\begin{aligned}
&\mu\Bigg(\int_{\bar{q}}^{\bar{q}} \Big(\lambda v\left(r - t_{SB}\right)\Big) w'\left(1 - F(q|\bar{e})\right) f(q|\bar{e}) \mathrm{d}q - c \\
&- \int_{\underline{q}}^{\bar{q}} \Big(\lambda v\left(r - t_{SB}\right)\Big) w'\left(1 - F(q|\underline{e})\right) f(q|\bar{e}) \mathrm{d}q\Bigg) = 0.
\end{aligned} \tag{53}$$

This allows us to rewrite equation (52), after integrating both sides over

$[\underline{q}, \bar{q}]$, as:

$$\mu c = \int_{\underline{q}}^{\bar{q}} \Big(\lambda v\left(r - t_{SB}\right)\Big) \left(\frac{1}{-\lambda v'\Big(r - t_{SB}\Big)}\right) f(q|\bar{e})\mathrm{d}q$$
$$- \int_{\underline{q}}^{\bar{q}} \mathbb{E}\left(\frac{1}{-\lambda v'\Big(r - t_{SB}\Big)}\right) v\left(r - t_{SB}\right) w'\Big(1 - (F(q|\bar{e}))\Big) f(q|\bar{e})\mathrm{d}q. \quad (54)$$

Rearranging leads to

$$\mu c =$$
$$\mathbb{E}\left(\frac{\lambda v\left(r - t_{SB}\right)}{-\lambda v'\Big(r - t_{SB}\Big)}\right) - \tilde{\mathbb{E}}\Big(\lambda v\left(r - t_{SB}\right)\Big)\mathbb{E}\left(\frac{1}{-\lambda v'\Big(r - t_{SB}\Big)}\right), \quad (55)$$

where

$$\tilde{\mathbb{E}}\Big(\lambda v\left(r - t_{SB}\right)\Big) := \int_{\underline{q}}^{\bar{q}} \Big(\lambda v\left(r - t_{SB}\right)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})\mathrm{d}q. \quad (56)$$

Consider first a CPT agent with no probability distortion, $w(p) = p$. Then, $\tilde{\mathbb{E}}\Big(\lambda v\left(r - t_{SB}\right)\Big) = \mathbb{E}\Big(\lambda v\left(r - t_{SB}\right)\Big)$ and the right hand side of (55) can be interpreted as a covariance operator. According to assumption 4 v in this domain exhibits $v' > 0$ and $v'' > 0$, both of which imply that $\lambda v\left(r - t_{SB}(q)\right)$ and $-\frac{1}{\lambda v'\left(r - t_{SB}(q)\right)}$ covary positively. Hence, the right-hand side (RHS) of (55) is positive, i.e. is a covariance operator, and that $\mu \geq 0$.

Let now w exhibit pessimism, $w'' > 0$. Then, $\tilde{\mathbb{E}}\Big(\lambda v\left(r - t_{SB}\right)\Big) < \mathbb{E}\Big(\lambda v\left(r - t_{SB}\right)\Big)$, which, together with the aforementioned result that $\mu \geq 0$ when there are no probability distortions, guarantees that the RHS of (55) is positive and that $\mu > 0$. Finally, the IC does not always bind when w exhibits $w'' < 0$. To see how, note that with this shape of w we obtain $\tilde{\mathbb{E}}\Big(\lambda v\left(r - t_{SB}\right)\Big) > \mathbb{E}\Big(\lambda v\left(r - t_{SB}\right)\Big)$. Thus, the RHS (55) can be equal to zero when $\tilde{\mathbb{E}}\Big(\lambda v\left(r - t_{SB}\right)\Big)$ is large enough, that is with more sever optimism according to Definition 3. We thus conclude that for severe optimism, the IC slacks and we obtain $\mu = 0$, but for moderate optimists we have that $\mu \geq 0$.

All in all an incentive feasible contract cannot be implemented in the domain of losses. That is because the participation constraint does not bind at the optimum. This result implies that the loss averse agent does not accept contracts that expose him to losses. To avoid that the agent rejects a contract, the principal needs to ensure that the agent is in the domain of gains. The cheapest way to accomplish this is by offering $t_{SB} = r$ in the segments whereby a second-best contract that takes into account the CPT agent probability perception but ignores his relativistic perceptions of outcomes, i.e. any of the contracts presented in Propositions 2, 3, or 4, locates him in the domain of losses.

Thus, when r is sufficiently large so that $t_{SB}(\bar{q}) \leq r$, the second-best contract given to the CPT agent must be constant everywhere paying $t_{SB}(q) = r$. If r is low enough such that $t_{SB}(\underline{q}) \geq r$ the contract in Propositions 2, 3, or 4 that best fits the agent's perceptions of probabilities suffice to motivate him. More importantly, if $t_{SB}(\underline{q}) < r < t_{SB}(\bar{q})$ then the optimal incentives scheme pays $t_{SB}(q) = r$ up to the performance level in which the contract that better reflects the agent's sensitivity to probabilities offers payments larger than r, after which the optimal incentives scheme follows the contract in Propositions 2, 3, or 4, depending on which of these contracts best fits the agent's perception of probabilities.

■

Proposition 6

Proof. We first show that whether one agent type is more efficient than another depends on the magnitude of the probabilities generated by e. The efficiency of an agent is the impact of exerting $\bar{e}$ rather $\underline{e}$ on $1 - F(q|e)$.

Formally, for the non-EUT agent efficiency efficiency amounts to:

$$\int_{\underline{q}}^{\bar{q}} w'\left(1-F(q|\bar{e})\right) f(q|\bar{e})\mathrm{d}q - \int_{\underline{q}}^{\bar{q}} w'\left(1-F(q|\underline{e})\right) f(q|\underline{e})\mathrm{d}q = w\left(1-F(q|\bar{e})\right) - w\left(1-F(q|\underline{e})\right). \quad (57)$$

Instead, for the non-EUT, efficiency amounts to:

$$\int_{\underline{q}}^{\bar{q}} f(q|\bar{e})\mathrm{d}q - \int_{\underline{q}}^{\bar{q}} f(q|\underline{e})\mathrm{d}q = \left(1-F(q|\bar{e})\right) - \left(1-F(q|\underline{e})\right). \quad (58)$$

Note that the W-MLRP implies both $F(q|\bar{e}) < F(q|\underline{e})$ and $w\left(1-F(q|\bar{e})\right) > w\left(1-F(q|\underline{e})\right)$.

A sufficient condition for the magnitude of the difference in equation (57) to be larger than that in equation (58) is $w'(1-F(q|e)) > 1$ for any e. Under likelihood insensitivity this happens in the interval $q \in [\underline{q}, q_l^{**})$, where q_l^{**} satisfies $w'\left(1-F(q_l^{**}|e)\right) = 1$ and $w''\left(1-F(q_l^{**}|e)\right) > 0$, and at the interval $q \in (q_h^{**}, \bar{q}]$, where q_h^{**} is such that $w'\left(1-F(q_h^{**}|e)\right) = 1$ and $w''\left(1-F(q_h**|e)\right) < 0$. In words, the likelihood insensitive agent is more efficient if the probabilities generated by $\bar{e}$ and $\underline{e}$ are located at one or at both extremes of the output interval. Instead, when these two probabilities lie in $q \in [q_l^{**}, q_h^{**}]$, the interval where probabilities are given similar weight, the EUT agent is more efficient.

There are two cases in which only one of the probabilities implied by the agent's possible choice lie in the insensitivity region. First, choosing $\bar{e}$ can generate a probability in $q \in [q_l^{**}, q_h^{**}]$ while the probability implied by $\underline{e}$ lies in $q \in [\underline{q}, q_l^{**}]$. Second, choosing $\bar{e}$ can generate a probability located in the interval $q \in (q_H^{**}, \bar{q}]$ while choosing $\underline{e}$ generates a probability located in $q \in [q_l^{**}, q_h^{**}]$. In these two cases, a sufficient condition for the non-EUT agent to be more efficient results from a direct comparison of the right hand sides of (57) and (58):

$$w\left(1-F(q|\bar{e})\right) - w\left(1-F(q|\underline{e})\right) > F(q|\underline{e}) - F(q|\bar{e}). \quad (59)$$

The above equation shows that the non-EUT agent is more efficient when $w\,(1 - F(q|\bar{e}))$ is sufficiently large, which is more likely to happen in the second case, i.e. when high effort generates a probability located in the interval of output realizations $q \in (q_h^{**}, \bar{q}]$. Also, (59) is more likely to hold when $\underline{e}$ generates a small probability, which happens in the first case when low effort generates a probability in $q \in [q_l^{**}, q_h^{**}]$. Finally, the non-EUT agent is more likely to be efficient when $F(q|\underline{e}) - F(q|\bar{e})$ is small, which happens when the difference in likelihoods generated by high and low effort is small.

Let us first consider the case in which the non-EUT agent is more efficient. As shown above this mainly happens when the agent's possible actions generate probabilities that are located at extremes of the output interval. To solve the principal's problem, we first reduce the number of constraints. Equations (9) and (8) immediately imply (10). Intuitively, when a menu of contracts allows the inefficient agent to attain his reservation utility while exerting high effort, the efficient agent should also attain this utility level. Hence, at the optimum the participation constraint in (9) binds, while the participation constraint in (10) slacks.

From equation (7) and the constraint in (9), which we already established binds at the optimum, we obtain:

$$0 \geq \max_{e \in \{\underline{e}, \bar{e}\}} \left\{ \int_{\underline{q}}^{\bar{q}} u\Big(t_{nB}(q)\Big) f(q|\bar{e})\,\mathrm{d}q - c(e) \right\}, \tag{60}$$

which implies that EUT agents cannot afford to mimic non-EUT agents. Hence, the relevant equation must be the adverse selection constrain in (8), which states that the non-EUT agent derives rents from mimicking the EUT agent. In contrast, equation (7) slacks at the optimum.

A direct implication that (8) binds is $t_{nB}(q) \geq t_B(q)$, leading to:

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_B(q)\Big) f(q|\bar{e})\,\mathrm{d}q - c = 0 > \int_{\underline{q}}^{\bar{q}} u\Big(t_B(q)\Big) f(q|\underline{e})\,\mathrm{d}q. \tag{61}$$

Hence, the moral hazard constraint in (5) slacks at the optimum.

Next, from the inequality in (10), which slacks at the optimum, along with equation (60), which holds with strict inequality, we obtain:

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_{nB}(q)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})\, \mathrm{d}q - c > 0 \geq \max_{e \in \{\underline{e}, \bar{e}\}} \left\{ \int_{\underline{q}}^{\bar{q}} u\Big(t_{nB}(q)\Big) f(q|\bar{e})\, \mathrm{d}q - c(e) \right\}. \tag{62}$$

The above equation, together with the assumption that pessimism coexists with likelihood insensitivity, implies that the agent's perception of probabilities yield:

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_{nB}(q)\Big) f(q|\underline{e})\, \mathrm{d}q > \int_{\underline{q}}^{\bar{q}} u\Big(t_{nB}(q)\Big) w'\left(1 - F(q|\underline{e})\right) f(q|\underline{e})\, \mathrm{d}q, \tag{63}$$

Equations (62) and (63) imply

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_{nB}(q)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})\, \mathrm{d}q - c > \int_{\underline{q}}^{\bar{q}} u\Big(t_{nB}(q)\Big) w'\left(1 - F(q|\underline{e})\right) f(q|\underline{e})\, \mathrm{d}q. \tag{64}$$

and equation (6) is implied by other constraints in the principal's program.

Hence, at the solution only equations (8) and (9) bind. The optimal transfer given to the EUT, t_B, must guarantee $\mathbb{E}\left(u(t_B)|\bar{e}\right) := \int_{\underline{q}}^{\bar{q}} u(t_B) f(q|\bar{e}) \mathrm{d}q = c$, satisfying the binding constraint in (9). Also, the transfer offered to the non-EUT, t_{nB}, should satisfy $\tilde{\mathbb{E}}\left(u(t_{nB})|\bar{e}\right) := \int_{\underline{q}}^{\bar{q}} u(t_{nB}) w'\left(1 - F(q|\bar{e})\right) f(q|\bar{e}) \mathrm{d}q = \tilde{\mathbb{E}}\left(u(t_B)|\bar{e}\right)$ as implied by equation (8).

At output intervals where the EUT is more efficient, the proof follows a similar logic. Then, it can be shown that the participation constraint of the non-EUT agent binds and the adverse selection incentive compatibility constraint for the EUT binds. Together these two binding constraints lead to a solution whereby t_{nB} guarantees $\tilde{\mathbb{E}}\left(u(t_{nB})|\bar{e}\right) = c$ and t_B guarantees $\mathbb{E}\left(u(t_B)|\bar{e}\right) = \mathbb{E}\left(u(t_{nB})|\bar{e}\right)$, at those output intervals.

■