TI 2022-090/I
Tinbergen Institute Discussion Paper

Optimal Incentives without Expected Utility

Víctor González-Jiménez[1]

1 Erasmus University Rotterdam

Optimal Incentives without Expected Utility*

Víctor González-Jiménez†

November 29, 2022

Abstract

This paper investigates the optimal design of incentives when agents distort probabilities. 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, Probability Weighting, Incentives, Mechanism Design, Rank-Dependent Utility.

*I thank Paul Heidhues, Mats Köster, Aurelien Baillon, Geoffrey Castillo, Gyula Seres, Daniel Garcia, and Cole Williams for valuable comments. I also thank participants in seminars at Erasmus University, Dusseldorf Institute for Competition Economics, and the University of Vienna.

†Erasmus School of Economics, Erasmus University Rotterdam. E-mail: `victor.gonzalez@ese.eur.nl`

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 many 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 et al., 2011, 2007; Bruhin et al., 2010; Fehr-Duda and Epper, 2011; Kahneman and Tversky, 1979; l'Haridon and Vieider, 2019; Tversky and Kahneman, 1992).[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 in the theory of incentives (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

[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 on probability weighting and reference dependence throughout this paper.

[3]See also 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).

stochastic dominance.

Our main contribution is to show how the principal can take advantage of an agent who distorts probabilities. We consider different types of probability distortion. With these we find that optimal contracts mimic those observed in real-life, such as salaries, high-performance commissions, and option-like contracts.

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 that the agent perceives to be more likely. For example, when facing an overly optimistic agent, the principal offers a contract that provides large payments only if the highest performance levels realize—in other words, a high-performance commission or long-shot.

We further show that, when optimism is moderate or pessimism is severe, incentive-compatible contracts in the standard sense are either not needed or cannot be implemented. Under moderate optimism, the first-best contract suffices to induce high effort; the agent's confidence that high performance levels realize is on its own sufficient to generate strong incentives. By contrast, the incentive-compatible contract under severe pessimism concentrates incentives at lowest performance levels. To avoid perverse incentives, such as agents wanting to destroy output, the principal needs to provide a high and fixed payment for all other performance realizations; a property that makes this contract excessively costly. The principal gives up incentive compatibility and ends up offering a contract with a constant payment for all performance levels—a salary.

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 (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 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 resembles an incentive scheme with two bonuses, one at low performance levels— an entry bonus— and one at high performance levels—a high performance bonus.

Using our framework, we consider a number of extensions. For example, we look at agents who also 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 biases (see Abdellaoui et al., 2007; Baillon et al., 2020a; Kahneman and Tversky, 1979; Kahneman et al., 1991; Tversky and Kahneman, 1992).[5] Depending on the circumstances, reference dependence gives rise to richer contracts, such as an incentive scheme featuring multiple bonuses ; or simpler contracts, such as a fixed wage with a lump-sum bonus.

We also discuss how our model can be adapted to incorporate ambiguity attitudes. Accounting thus for another well-documented deviation from expected utility (Ellsberg, 1961; Halevy, 2007). Our framework not only captures ambiguity-averse attitudes, but also ambiguity-seeking attitudes and overweighting of rare events. As such, it provides a more general and descriptive characterization of attitudes toward ambiguity (Trautmann and Van De Kuilen, 2015). This feature enables us to generate novel findings as well as to reconcile results in the literature that seemed hitherto scattered.

Broadly speaking our paper contributes to the behavioral contract theory literature. This literature incorporates biases into contract theory 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 when contracting with agents with cumulative prospect theory preferences it is first-best optimal to use stock options. We find a similar result in which the first-best contract given to likelihood-insensitive agents exhibits an option-like shape. But importantly, we go beyond since

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

we also look at what happens when effort is *not* contractible, for different possible shapes of probability weighting, do not commit to functional forms of utility, and do not restrict our analyses to one type of compensation scheme. Additionally, Baillon et al. (2020b) demonstrates that probability distortion can lead to lower motivation. That paper is however silent about the optimal design of incentives. We show how the principal must design a contract to optimally motivate the agent while accounting for probability distortions and other irrationalities.

We also contribute to the contract theory literature. We speak to a well-known paradox put forward by Salanié (2003) stating that the complex solutions predicted by contract theory do not match the simplicity of contracts observed in practice. 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 reaching high performance levels. These two contracts are among the most popular compensation practices. Generally speaking, we show that introducing descriptively valid theories of risk in the principal-agent problem leads to contracts that are often observed in practice.

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 on the task $e \in \{\underline{e}, \bar{e}\}$. 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 $c > 0$. In Section 5, we relax this assumption of binary effort and constant effort costs and show that our results hold in a more general setting in which effort is assumed to be a continuous variable.

To incentivize the agent to exert high effort, 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, about which we make the following assumptions.

Assumption 1. *The basic utility function $u : \mathbb{R}_+ \to \mathbb{R}_+$ is $\mathcal{C}^2$, and exhibits $u(0) = 0$, $u' > 0$, $u'' < 0$, and $-\frac{u''}{u'} < B$ for $|B| < +\infty$.*

The basic utility, also known as von Neumann-Morgenstern utility function, exhibits the standard property of diminishing returns, i.e. $u' > 0$ and $u'' < 0$. A property that generates risk-averse attitudes in an expected utility framework.

The agent's action, e, cannot be directly observed by the principal. Furthermore, output on the task q, throughout also referred as performance, is a random variable that takes values in the compact interval $[\underline{q}, \bar{q}]$. Hence, by observing a realization of output, the principal 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 $f(q|e)$. Furthermore, we assume that the relationship between output and effort is governed by the monotone likelihood ratio property

Assumption 2. *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, it implies that high output realizations are more likely to be drawn from a distribution of output conditional on high effort. Thus, the agent can influence the likelihood of obtaining higher performance levels on the task.

Throughout, we assume that 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 $S : \mathbb{R}_+ \to \mathbb{R}_+$ exhibits $S' > 0$, $S'' \leq 0$ for all q, and $S(\underline{q}) = 0$. This objective function together with Assumption 2 imply that the principal is interested in implementing high effort. It generates output levels that dominate those generated by low effort—in the sense of first-order stochastic dominance—, which, in turn, boosts profits.

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

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

To relate to standard notation in the literature, we use decumulative probabilities. That is, we refer to a probability, p, as the likelihood that a realization better than an output level $Q \in [\underline{q}, \bar{q}]$ for a given e takes place. Formally, let $p := 1 - F(Q|e)$. 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]

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

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. The following assumptions are imposed on w:

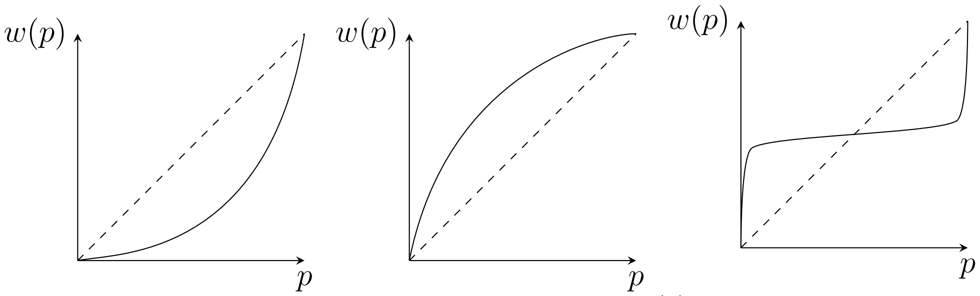
Assumption 3. *Let $p := 1 - F(q|e)$ for any $q \in [\underline{q}, \bar{q}]$. The probability weighting function $w : [0,1] \to [0,1]$ is $\mathcal{C}^2$ and exhibits:*

- *$w(0) = 0$ and $w(1) = 1$;*
- *$w'(p) > 0\ \forall p \in (0,1)$;*

[6]For further clarification, let $q_1, q_2 \in [\underline{q}, \bar{q}]$ with $q_2 > q_1$. Notice that,

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

Figure 1: Examples of probability weighting functions



(a) Pessimism (b) Optimism (c) Likelihood insensitivity

Note: Dashed lines represent accurate perception of probability.

- *For some* $\tilde{p} \in [0, 1]$, $w''(p) < 0$ *if* $p < \tilde{p}$ *and* $w''(p) > 0$ *if* $p > \tilde{p}$*;*
- *If* $\tilde{p} = 1$, $\lim_{p \to 0} w'(p) = +\infty$ *and* $\lim_{p \to 1} w'(p) = 0$*;*
- *If* $\tilde{p} = 0$, $\lim_{p \to 0} w'(p) = 0$ *and* $\lim_{p \to 1} w'(p) = +\infty$*;*
- *If* $\tilde{p} \in (0, 1)$, $\lim_{p \to 0} w'(p) = +\infty$ *and* $\lim_{p \to 1} w'(p) = +\infty$

In words, the probability weighting function is an strictly increasing and continuous function that maps the unitary interval into 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 large weights to extreme performance levels while assigning similar weights to intermediate output levels. An implication of this latter shape is the existence of an interior fixed-point, $\hat{p} \in (0, 1)$ such that $w(\hat{p}) = \hat{p}$.

The seemingly drastic assumptions of extreme sensitivity to rare and almost-certain events, i.e. $\lim_{p\to 1} w'(p) = \infty$ and $\lim_{p\to 0} w'(p) = \infty$, are incorporated in the most prominent proposals of probability weighting functions. For instance, in the parametric form proposed by Prelec (1998). There, the behavioral foundation of compound invariance implies that "the slope tends to infinity at zero" and that "the picture at the other endpoint, is almost the same [...], the slope $\frac{dw}{dp}$, tends again to infinity" (Prelec (1998, p. 505)). Moreover, Dierkes and Sejdiu (2019) show that these assumptions are also implied by the parametric forms of Tversky and Kahneman (1992) and Goldstein and Einhorn (1987).[7]

Furthermore, these assumptions have relevant implications that are formally presented next. We relegate all formal proofs to Appendix A.

Lemma 1. *If* $\lim_{p\to 0} w'(p) = +\infty$*, then* $\lim_{p\to 0} w''(p) = -\infty$ *and* $\lim_{p\to 0} \frac{w''(p)}{w'(p)} = -\infty$*.*

Lemma 2. *If* $\lim_{p\to 1} w'(p) = +\infty$*, then* $\lim_{p\to 1} w''(p) = +\infty$ *and* $\lim_{p\to 1} \frac{w''(p)}{w'(p)} = +\infty$*.*

In words, the second derivative of the weighting function and the analog of the Arrow-Pratt measure in probabilities, $\frac{w''(p)}{w'(p)}$, can be unbounded at small and large probabilities. These implications are not only useful for proving our main results, but also formalize and generalize the findings of Dierkes and Sejdiu (2019). In Appendix A, we also present similar results for the cases of insensitivity to extreme events, relevant to optimistic agents, and insensitivity to almost-certain events, relevant to pessimistic agents.

All in all, the preferences of the agent who exhibits probability distortions are characterized by rank-dependent utility (RDU):

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

[7]Notably, 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 extreme sensitivity to rare and almost-certain probability events.

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.[8] We assume that the principal can contract with either EUT or non-EUT agents and, as is standard in the literature, that she is fully informed about the agent's risk preferences. In Section 5, we relax this assumption by considering a framework in which the principal screens agents according to their risk preference before solving the moral hazard problem. That extension shows that optimal contracts resulting from the plain moral hazard framework are crucial to solve that more complicated problem.

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: optimists and pessimists. These agents deviate from expected utility due to motivational factors reflecting a proneness or a dislike for risk. Optimists like risk and assign large weights to the best outcomes—they irrationally believe the best outcomes realize more often. Optimism is captured with a concave probability weighting function:

Definition 1. *Optimism is characterized by a probability weighting function $w(p)$, defined in Assumption 3, with the additional restriction $\tilde{p} = 1$. Therefore, $\lim_{p\to 0} w'(p) = +\infty$ and $\lim_{p\to 1} w'(p) = 0$.*

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

Definition 2. *Pessimism is characterized by a probability weighting function $w(p)$, defined in Assumption 3, with the additional restriction $\tilde{p} = 0$. Therefore, $\lim_{p\to 0} w'(p) = 0$ and $\lim_{p\to 1} w'(p) = +\infty$.*

[8] It is relevant to emphasize that probability distortion is not the only departure from expected utility that we consider. In Section 5 we also consider reference-dependence and discuss ambiguity attitudes.

3.1 First best

We first consider a situation in which effort is contractible. This is done to build intuition about the role of probability weighting on contracting aside from incentive compatibility. If effort is contractible, the principal only needs to ensure participation. Formally:

$$\begin{aligned} \max_{t(q)} \quad & \int_{\underline{q}}^{\bar{q}} \big(S(q) - t(q)\big) f(q|\bar{e})\, \mathrm{d}q \\ \text{s.t.} \quad & \int_{\underline{q}}^{\bar{q}} u(t) w'(1 - F(q|\bar{e})) f(q|\bar{e})\, \mathrm{d}q - c \geq \bar{U}. \end{aligned}$$

The following proposition describes the optimal contract for each agent type. Interestingly, the principal can, sometimes, take advantage of the irrationalities exhibited by non-EUT agents with this first-best contract.

Proposition 1. *Let Assumptions 1 and 3 hold. The first-best contract, $t^{fb}(q)$, exhibits three possible shapes, all continuous:*

(i) If the agent is EUT, the contract t^{fb}_{EU} is constant in q;

(ii) if the agent exhibits optimism, $t^{fb}_{O}(q)$ is everywhere increasing in q;

(iii) if the agent exhibits pessimism, t^{fb}_{P} is constant in q and is equal to t^{fb}_{EU}.

The first part of Proposition 1 establishes the standard risk-sharing argument of Borch (1960). When the agent is EUT and exhibits risk aversion, the principal fully insures him with a contract that transfers a fixed amount regardless of the output realization. The magnitude of that constant transfer ensures that the contract will be accepted by the agent.

When facing an optimist, the first-best contract increases in performance. From the agent's point-of-view, this risky contract offers full insurance; it provides larger transfers for realizations that are perceived to be more likely and lower transfers for realizations that are perceived to be less likely. However, this contract is in fact taking advantage of his misperception of probabilities. As compared to the actual full insurance contract, described in Proposition 1 (i), the principal is underpaying more likely events and overpaying unlikely events.

The principal would like to take advantage of a pessimistic agent in a similar way. A strategy that 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. To avoid such sabotage and ensure participation, the principal instead offers a constant transfer that yields utility equal to the agent's outside option.[9] This solution is nonetheless far from ideal; it renders impossible the exploitation of the agent's irrationality because it eliminates risk. Hence, under pessimism, participation and monotonicity requirements come at the cost of efficiency.

To investigate how the aforementioned first-best contracts change as pessimism/optimism become stronger, we will also talk about agents who are more optimistic or more pessimistic than others. We use the following definition from Yaari (1987):

Definition 3. *Agent i is more pessimistic (optimistic) than agent j if $w_i = \theta \circ w_j$, where w_i and w_j are the probability weighting functions corresponding to agent i and j, respectively, and $\theta : [0,1] \to [0,1]$ is continuous, strictly increasing, and convex (concave).*

A probability weighting function that is more concave than another generates more optimism because it makes the agent assign larger probability weights to high performance levels and smaller weights to low performance levels. The reasoning is mirrored for convex probability weighting functions.

The following corollary shows that when contracting with a more optimistic agent, it can be first-best optimal to offer a contract with stronger incentives. By concentrating larger transfers at highest performance levels, the principal takes advantage of the agent's stronger confidence that high output levels realize. This result is nevertheless conditional on the agent's coefficient of absolute risk aversion not becoming larger with the contract's change.[10]

[9] This solution also avoids the excessive expenditures that would result from offering the pessimistic agent a contract that increases in performance. There, the principal would need to provide extremely large payments to cover for the implied exposure to risk.

[10] In a straightforward extension of Corollary 1, it can be shown that a necessary and

Corollary 1. *Assume* $-\frac{u''(t^{fb}_{O,i}(q))}{u'(t^{fb}_{O,i}(q))} < -\frac{u''(t^{fb}_{O,j}(q))}{u'(t^{fb}_{O,j}(q))}$. *If agents* i *and* j *are optimistic and agent* i *is more optimistic than agent* j*, then the first-best contract offered to agent* i *exhibits* $\frac{dt^{fb}_{O,i}(q)}{dq} > \frac{dt^{fb}_{O,j}(q)}{dq}$.

Furthermore, the efficiency loss borne by the principal from not being able to exploit pessimism (Proposition 1 (iii)) becomes more pronounced as the agent becomes more pessimistic. Recall that this efficiency loss emerges from the impossibility of offering contracts that decrease in performance.

Corollary 2. *Under the first-best contract, the Principal cannot exploit pessimism, she offers* t^{fb}_{EU} *to all pessimistic agents regardless of their degree of pessimism. The efficiency loss of offering* t^{fb}_{EU} *instead of a schedule that decreases in* q *increases with pessimism.*

3.2 Second best

We now consider the more interesting setting in which the agent's action is not contractible. The principal now seeks to maximize her objective function by choosing a transfer that is accepted by the agent and also elicits high effort. Therefore, the maximization problem of the principal is:

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

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 everywhere in performance.

sufficient condition for $\frac{\mathrm{d}t^{fb}_{O,i}(q)}{\mathrm{d}q} > \frac{\mathrm{d}t^{fb}_{O,j}(q)}{\mathrm{d}q}$ is that θ, the transformation of the probability weighting function from Definition 3, is sufficiently concave.

We present this solution in the next Proposition.

Proposition 2. *Under Assumptions 1, 2, and $w(p) = p$, the optimal incentive scheme, $t^{sb}_{EU}(q)$, is continuous and everywhere increasing in q.*

Before presenting the second-best contract for non-EUT agents, we introduce an assumption that is crucial to our analysis. We strengthen the MLRP (Assumption 2) to ensure that, from the agent's point-of-view, output realizations are sufficiently informative about his chosen action.

Assumption 4 (W-MLRP)**.** *The modified monotone likelihood ratio property (W-MLRP) is*

$$\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{w'\Big(1-F(q|\underline{e})\Big)f(q|\underline{e})}{w'\Big(1-F(q|\bar{e})\Big)f(q|\bar{e})}\right) < 0.$$

The W-MLRP implies that the principal is sophisticated. She anticipates how the agent's probability distortions affect his perception about the informativeness of a chosen action and implements incentives accordingly. The agent, on the other hand, is naive and does not evaluate the informativeness of output realizations using mathematical expectations, which would be equivalent to anticipating the way in which the principal evaluates those realizations. Consequently, solutions under Assumption 4 constitute exploitative contracts: they are designed to take advantage of that naïveté.

The following Lemmas are not only important to prove our main propositions but also provide an intuition about the strength of the W-MLRP vis-à-vis the MLRP.

Lemma 3. *The W-MLRP implies:*

(i) $w\Big(1-F(q|\bar{e})\Big) \geq w\Big(1-F(q|\underline{e})\Big)$*;*

(ii) the MLRP.

Lemma 4. *If the MLRP holds and*

$$\frac{w''\Big(1-F(q|\bar{e})\Big)}{w'\Big(1-F(q|\bar{e})\Big)}f(q|\bar{e}) - \frac{w''\Big(1-F(q|\underline{e})\Big)}{w'\Big(1-F(q|\underline{e})\Big)}f(q|\underline{e}) \leq 0,$$

then the W-MLRP holds.

The W-MLRP is more stringent than the standard MLRP (Lemma 3). Moreover, the merit of Lemma 4 is to show that under pessimism and optimism, the W-MLRP can be attributed to the MLRP along with reasonable properties of the probability weighting function.[11] Throughout, we will assume the W-MLRP holds; note, however, that thanks to this Lemma that the results presented below can be obtained under more standard assumptions along with restrictions on the concavity or convexity of the weighting function.

The next proposition describes the properties of the second-best contract that solves the principal's program when she faces an optimist.

Proposition 3. *Let Assumptions 1, 3, 4 hold. Under optimism, there exists a threshold cost level $\hat{c}_O$, such that the second-best contract, $t_O^{sb}(q)$, is:*

(i) Identical to the first-best contract, $t_O^{fb}(q)$, in Proposition 1 if $c < \hat{c}_O$;

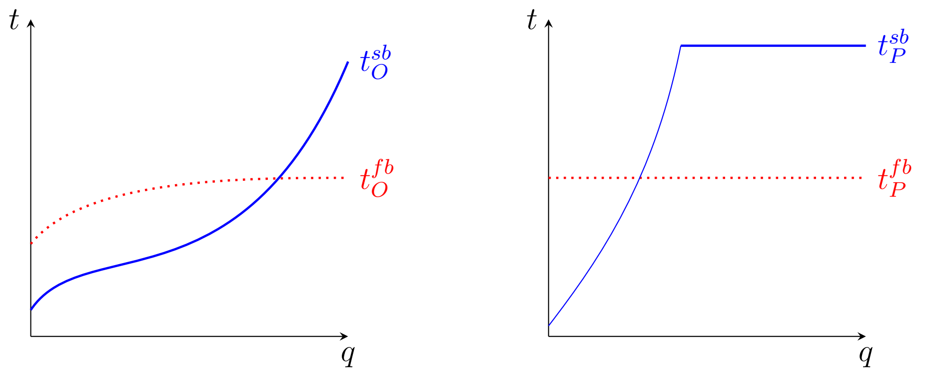
(ii) strictly increasing in q everywhere and introducing rewards and punishments with respect to $t_O^{fb}(q)$ if $c \geq \hat{c}_O$.

For an optimistic agent, the first-best contract from Proposition 1 might suffice to incentivize high effort. Recall that such contract specifies transfers that increase in performance. Since the optimist erroneously believes that high performance levels are more likely to realize, the higher transfers specified by the first-best contract at those performance levels convince him that high effort is profitable. Therefore, optimism can make the incentive compatibility constraint slack at the optimum.

When the cost of exerting high effort is pronounced, the optimistic agent requires a contract with higher-powered incentives as compared to the first-best. That is, a contract that specifies lower transfers at low output levels, generating punishments with respect to the first-best, and larger transfers at the high-end of the output space, generating rewards. This contract motivates the agent because the combination of high transfers at high performance levels and optimism inflates the agent's perceived benefit of exerting high effort. An example of these contracts is presented in Figure 2a.

[11] Specifically, if $w(p)$ exhibits more convexity at probabilities generated by low effort, $\underline{e}$, than at probabilities generated by high effort, $\bar{e}$, and the MLRP is assumed, then the W-MRLP holds.

Figure 2: Illustration of Proposition 1 and Proposition 3.



(a) Optimal contracts under optimism (b) Optimal contracts under pessimism

Note: Red dotted lines represent first-best contracts. Blue solid lines represent second-best contracts.

We use Definition 3 to provide a comparative static that demonstrates how the contracts in Proposition 3 are chosen according to the agent's degree of optimism.

Corollary 3. *Let* $q^* \in (\underline{q}, \bar{q})$ *be such that* $w'\Big(1 - F(q^*|\bar{e})\Big) = 1$*. Stronger optimism in the sense of Definition 3*

(i) Shortens the set $c \in (0, \hat{c}_O]$ *in which* $t_O^{sb}(q) = t_O^{fb}(q)$ *if* $q < q^*$*;*

(ii) enlarges the set $q < q^*$*.*

Stronger optimism makes it less likely that the first-best contract, $t_O^{fb}(q)$, incentivizes high effort in a moral hazard setting. The agent's greater confidence that high output levels realize leads him to the erroneous belief that such outcomes can be achieved with low effort. That conviction makes the implementation of more powerful incentives crucial; they enhance the perceived benefits from choosing high effort. Put differently, the agent needs to be incentivized not to "rest on his laurels."

In a related paper, De La Rosa (2011) also finds that the first-best contract increases in performance and that the incentive compatibility constraint might slack. His approach differs from ours because the agent is assumed to be EUT while exhibiting overconfident beliefs. Our results can be understood as complementary since they require less stringent assumptions.[12] Moreover, our results generalize those of De La Rosa (2011). We consider a rich output space, which enables us to fully characterize the shape of optimal contracts. Also, we show that these findings are not an artifact of violations of first-order stochastic dominance; a normative condition that does not hold in his setting.[13]

We turn to study the optimal contract given to pessimistic agents. The following proposition characterizes the solution.

Proposition 4. *Let Assumptions 1, 3, 4 hold. Under pessimism, the second-best contract, $t_P^{sb}(q)$, is:*

(i) Smoothly increasing in q up to some threshold $q_{\mathcal{I}} \in (\underline{q}, \bar{q})$ after which pay is constant in q;

(ii) constant in q everywhere.

When the agent is pessimistic, the optimal contract can take two possible shapes. First, the principal concentrates incentives at low performance levels; this amounts to a contract that specifies larger transfers in exchange of higher performance levels at the lower-end of the output interval while being performance-insensitive everywhere else. These incentives along with the agent's irrational perception that low performance levels realize more often motivate high effort.[14] An example of this contract shape is presented in Figure 2b. Second, the principal cannot implement incentives, for reasons that will be explained below, and has no other option than to offer a constant transfer. i.e. a salary.

The following comparative static is useful to understand the circumstances under which the principal chooses one of the contract-shapes in Proposition 4

[12]Specifically, the assumption of disagreement in beliefs, justified in his setting through different priors or lack of revision, is not necessary. Misperceptions of probabilities occur in situations in which agents are given the same priors or objective probabilities.

[13]Additionally, our characterization of optimism has an axiomatic foundation. This property allows us to trace back our results to violations of axioms of choice. In our case, the independence axiom.

[14]It would be inefficient to, say, implement a schedule that is increasing everywhere in performance since the pessimistic agent perceives high output realizations to be unlikely. To motivate the agent with such a contract shape, the principal would need to incur in excessively large transfers, yielding wasteful expenditures.

over the other.

Corollary 4. *Assume that incentives can be implemented. Stronger pessimism implies a larger segment* $q \in [q_{\mathcal{I}}, \bar{q}]$ *for which* $t_P^{sb}(q)$ *(Proposition 4) is flat.*

Stronger pessimism leads the principal to concentrate incentives at lower performance levels. A property that makes incentive schemes increasingly performance-insensitive.

Corollary 4 implies that to motivate an agent with acute pessimism, rewards and punishments must be concentrated in a thin output subset in the neighborhood of $\underline{y}$. However, by doing so, the principal incurs in inefficient expenditures. That is because a flat amount equal to the largest reward is paid in the remainder of the output space; considerable transfers that do not generate incentives and make the contract overly expensive. As a result, incentive compatibility is given up and participation is ensured with a salary (Proposition 4 (ii)). For more moderate pessimism, the implementation of incentives is affordable. In that case incentives are concentrated at low, but not at the lowest, output levels rendering incentive compatibility affordable (Proposition 4 (i)).

Proposition 4 and Corollary 4 justify the prevalence of performance-insensitive contracts in organizations (Lazear and Oyer, 2007; Salanié, 2003) by virtue of pessimism. In fact, excessive pessimism about performance explains the widespread usage of salaries.[15] Our framework also rationalizes "put contracts." A financial instrument in which the contract-writer commits to buy a commodity at a fixed price during a specified time period. The profits of this contract exhibit the shape described in Proposition 4 (ii). Hence, they can be thought as emerging from the contract-writer having moderate pessimism about the prize of the commodity.

To conclude this section, we compare the transfers of the second-best contract presented in Proposition 2 with those of the contracts presented in Propositions 3 (ii) and 4 (i). We thus focus on the contracts that emerge when the incentive compatibility constraint binds and when incentives can be

[15] In Section 5, we discuss how this result can be further reinforced by ambiguity aversion, a form of pessimism toward ambiguity.

implemented. The following corollary formalizes these comparisons.

Corollary 5. *Assume that the incentive compatibility constraint binds,*

(i) $t_O^{sb}(q)$ *(Proposition 3) offers lower transfers in* $q \in [\underline{q}, q^*)$ *and higher transfers for some* $q \in (q^*, \bar{q}]$ *as compared to* t_{EU}^{sb} *(Proposition 2).*

(ii) $t_P^{sb}(q)$ *(Proposition 4) offers higher transfers in* $q \in [\underline{q}, q^*)$ *and lower transfers for some* $q \in (q^*, \bar{q}]$ *as compared to* $t_{EU}^{sb}(q)$ *(Proposition 2).*

As compared to the contract given to the EUT agent, the contract targeting the optimist offers lower transfers at low performance levels and higher transfers at high performance levels. This result shows, once again, how the principal exploits optimism. She offers modest transfers at output levels that the biased agent does not consider to be likely, but that, under accurate perception of probabilities, require larger transfers. Furthermore, the principal is overpaying high and thus unlikely realizations of output. This latter feature of the contract motivates the optimist.

In contrast, the pessimist receives a contract that specifies larger transfers at low performance levels and lower transfers at high performance levels. The principal underpays high output realizations, which under accurate perception of probabilities require larger transfers, and overpays low output realizations. Incentives are thus concentrated where it matters to this agent.

4 Likelihood insensitivity and inverse S-shaped probability weighting functions

So 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 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 probability weights to very small or very large probabilities—associated to near-certain and near-impossible events—and assigns probability weights that are similar to intermediate probabilities, thus yielding an inverse S-shaped probability weighting function.

Definition 4. *Likelihood insensitivity is characterized by a probability weighting function $w(p)$, defined in Assumption 3 with the additional restriction $\tilde{p} = \hat{p} = 0.5$. Therefore, $\lim_{p \to 0} w'(p) = +\infty$ and $\lim_{p \to 1} w'(p) = +\infty$.*

Recall that $\hat{p}$ is the interior fixed-point that results from $\tilde{p} \in (0, 1)$.

4.1 First best

As in the previous section, we first characterize the optimal contract when effort is contractible. The following proposition shows that when agents are likelihood insensitive, the optimal first-best contract is an "option-like" incentive scheme.[16]

Proposition 5. *Let Assumptions 1,3, 4 hold. Under likelihood insensitivity, the first-best contract, $t_L^{fb}(q)$, is constant up to threshold $q_I \in (\underline{q}, \bar{q})$ after which pay strictly increases in q.*

A consequence of low probabilities being overweighted is that the likelihood-insensitive agent exhibits risk-seeking attitudes for unlikely events. The increasing segment of the first-best contract in Proposition 5 is designed to match this proneness to risk. From the agent's point-of-view this feature of the contract delivers full insurance: it awards larger transfers for events that are given higher probability weight. Similarly, the flat part of the contract reflects underweighting of large probabilities. The agent's tendency to give too much weight to worse events realizing generates risk aversion. Therefore,

[16]We refer to a contract as being option-like when it exhibits a performance-insensitive segment at low output levels but increases in performance at high performance levels. A shape that resembles a "*call option*." In doing so we follow the convention in the literature (See Proposition 1 in de Meza and Webb (2007)). To avoid confusion with *put options* we have referred to those instruments as "put contracts."

insurance requires that the agent is fully protected from risk at those output levels with a fixed transfer.

Proposition 5 echoes the result of Spalt (2013); an option-like contract is first-best optimal under likelihood insensitivity. Our result however emerges in a setting that is free of assumptions on the contract structure and that does not impose parametric forms of utility and probability weighting functions. Furthermore, we go beyond and investigate whether this results holds in a more general moral hazard setting.

To investigate how the first-best contract changes as likelihood insensitivity becomes stronger, we will talk about agents who are more likelihood insensitive than others. To that end, we 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 $\mathcal{C}^2$ with $\phi' > 0$, and there exist 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 are now in a position to provide the more-likelihood-insensitive-than relation also due to Tversky and Wakker (1995).

Definition 6. *Agent i is more likelihood insensitive than agent j if $w_i = \phi \circ w_j$ where w_i and w_j are their respective probability weighting functions and $\phi : [0,1] \to [0,1]$ is subadditive.*

An agent is more likelihood insensitive than another when he assigns more probability weight to extreme probability events– highest and lowest performances realizing— while assigning less weight to intermediate probabilities. In other words, his weighting function exhibits a more pronounced inverse-S shape. The following corollary shows how the contract of Proposition 5 changes as likelihood insensitivity becomes more severe.

Corollary 6. *If agents i and j are likelihood insensitive and agent i is more likelihood insensitive than j, the first-best contract offered to agent i exhibits a larger segment $[\underline{q}, q_I)$ in which pay is constant.*

Stronger likelihood insensitivity must be matched by the principal with a contract that exhibits a larger performance-insensitive segment. This contract property acknowledges the agent's stronger risk aversion; a consequence of assigning larger probability weights to worst events. Moreover, a flatter contract also responds to the agent's increasing insensitivity to probabilities. The increasing part of the first-best contract becomes less effective in insuring an agent who becomes worse at distinguishing (small) probabilities. Consequently, the principal can save in costs by making the contract flatter.

4.2 Second best

Likelihood insensitivity makes the moral hazard problem more restrictive because the W-MLRP cannot be attributed to properties of the probability weighting function, as it was the case for optimists and pessimists (Lemma 4). Recall that around $\tilde{p}$ probabilities are almost indistinguishable to the likelihood insensitive agent. Therefore, the requirement of Lemma 4—that $w(p)$ must be more convex at the probability weight implied by high effort as compared to that implied by low effort— cannot hold in the neighborhood of $\tilde{p}$. The W-MLRP must be now explicitly assumed.

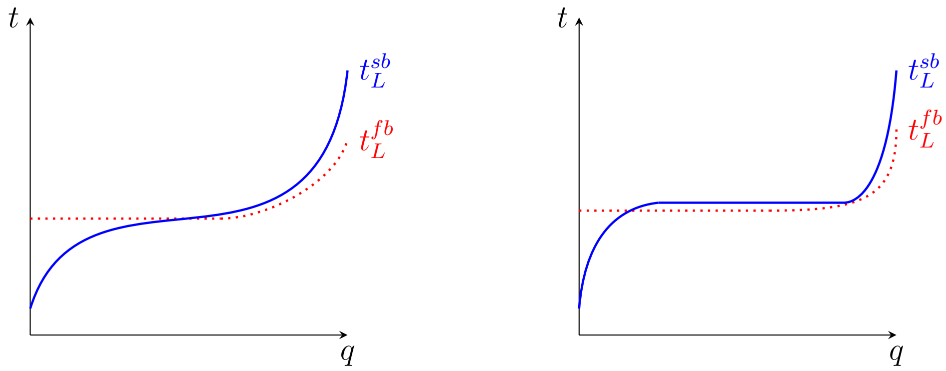
The next proposition presents the second-best contract given to the likelihood-insensitive agent.

Proposition 6. *Let Assumptions 1, 3, 4, and likelihood insensitivity hold. There exists a threshold cost level $\hat{c}_L > 0$, such that the second-best contract, $t_O^{sb}(q)$:*

(i) is identical to the first-best contract, $t_L^{fb}(q)$ from Proposition 5 if $c < \hat{c}_L$;

(ii) is strictly increasing in q everywhere and exhibits steep payment increments at extreme performance levels if $c \geq \hat{c}_L$;

Figure 3: Illustration of Proposition 5 and Proposition 6.



(a) Optimal contracts under moderate likelihood insensitivity

(b) Optimal contracts under strong likelihood insensitivity

Note: Red dotted lines represent first-best contracts. Blue solid lines represent second-best contracts.

(iii) pays a constant amount for some finite, fixed, compact interval, but above and/or below this interval pay steeply increases in q if $c \geq \hat{c}_L$.

The first part of the proposition shows that the incentive compatibility constraint might not bind at the optimum. The rationale for this result is analogous to that given for Proposition 3 (i). The agent's inaccurate perception that high output levels are more likely to realize, than they actually are, combined with the first-best contract offering higher transfers for those output levels convince him that choosing high effort is profitable.

When the cost of high effort is sufficiently pronounced, i.e. $c \geq \hat{c}_L$, the principal must implement a contract with higher-powered incentives as compared to those included in the first-best. These incentives have to be concentrated at extreme output levels, i.e. those output realizations that are assigned the largest probability weights. The resulting contract can be either everywhere increasing, as the contract described by Proposition 6 (ii) and exemplified in Figure 4a, or performance-insensitive at intermediate performance levels, as the contract described by Proposition 6 (iii) and exemplified in Figure 4b. Whether one of these contracts is chosen over the other depends on the agent's degree of likelihood insensitivity, as it will be explained below.

Next, we study the consequences of stronger degrees of likelihood insensitivity on optimal contracting. We focus on the case in which the incentivie compatibility constraint holds.

Corollary 7. *Assume that the incentive compatibility constraint binds. If agents i and j are likelihood insensitive and i is more likelihood insensitive than j, the second-best contract, t_L^{sb} (Proposition 6), given to i exhibits a larger segment in which pay is performance insensitive.*

The principal responds to a poorer capacity to discriminate intermediate probabilities by increasingly concentrating incentives at extreme output levels. An adjustment that generates contracts that are performance-insensitive to a greater extent. Consequently, an agent with acute likelihood insensitivity, who barely discriminates intermediate probabilities, renders the implementation of incentives for middle-ranged outcomes wasteful. Making the contract in Proposition 6 (iii) ideal for him.

Furthermore, Corollary 7 entails that the performance-insensitive segment of the second-best contract must shrink as the agent becomes less likelihood insensitive. Thus, at the limit, for an agent with slight to modest likelihood insensitivity, an incentive scheme that is everywhere increasing is effective; his slight distortion of probabilities does not impede him from responding to effectives over the entire output space. Consequently, he obtains the contract described by Proposition 6 (ii).

To conclude this section, we compare the transfers included in the optimal contracts of Proposition 6 (ii) and (iii) to those included in the contract given to the EUT agent.

Corollary 8. *Assume that the incentive compatibility constraint binds. Let $q^*, q^{**}, \tilde{q} \in (\underline{q}, \bar{q})$ satisfy $w''\big(1 - F(\tilde{q}|e)\big) = 0$, $w'\big(1 - F(q^*|\bar{e})\big) = w'\big(1 - F(q^{**}|\bar{e})\big) = 1$, $w''\big(1 - F(q^*|\bar{e})\big) > 0$, and $w''\big(1 - F(q^{**}|\bar{e})\big) < 0$. As compared to t_{EU}^{sb} (Proposition 2), contract $t_L^{sb}(q)$ (Proposition 6) offers lower transfers in $q \in [\tilde{q}, q^{**})$ and higher transfers in $q \in [\underline{q}, q^*)$, at the lowest output levels in $q \in [q^*, \tilde{q})$, and at the highest output levels in $q \in [q^{**}, \bar{q}]$.*

Corollary 8 shows how the agent's irrationality is exploited. As compared to the contract given to the EUT agent, the contract offered to the likelihood-insensitive agent underpays events with intermediate likelihood and ovepays extreme events, i.e. high and low performance levels. Again, incentives are concentrated where it matters to this biased agent.

Proposition 6 and Corollary 8 provide another justification for the extensive usage of performance-insensitive contracts. Namely, the incapacity of individuals to accurately discriminate probabilities which renders strong incentives ineffective. We note that likelihood insensitivity and pessimism are prevalent in decision-making under risk (Wakker, 2010).[17] Therefore, our model yields the sharp prediction that insensitivities in incentive schemes arise due to combination of these factors.

5 Extensions

In this section, we briefly discuss some extensions and emphasize how they can be derived from our previous analyses.

5.1 Preferences

5.1.1 Loss Aversion and Diminishing Sensitivity

We enrich the agent's risk preferences by considering Cumulative Prospect Theory (CPT from here onward, Tversky and Kahneman, 1992). Agents with these preferences not only distort probabilities but also evaluate potential transfers relative to a reference point $r > 0$. Transfers below the reference point count as *losses*, while transfers above it count as *gains*. The main departure of CPT with respect to RDU and EUT is that the agent can exhibit different risk preferences for gains and losses. A consequence of loss aversion and diminishing sensitivity.

We ask ourselves whether including these richer risk preferences lead to considerably different solutions. In the interest of space, we present the results of our analyses in Appendix B. It turns out that the solutions to the principal's problem crucially incorporate the properties of the contracts

[17]Their conjunction generates inverse-S probability weighting functions with an interior fixed-point at low probabilities.

presented in Sections 3 and 4. Furthermore, loss aversion and diminishing sensitivity augment the segments in which the optimal contract is performance insensitive. A feature that generates more realistic contracts. For example, when the agent is loss averse and pessimistic, the optimal contract consists of a bonus and a wage.

5.1.2 Ambiguity

Our framework can also capture deviations from EUT due to attitudes toward ambiguity. It turns our that the model can incorporate a wide variety of ambiguity attitudes, namely ambiguity aversion, ambiguity seeking, and rare-event overweighting. Thus, it generates novel findings and reconciles results in the literature that seemed unrelated.

Consider a setting in which agent and principal does not know $F(q|e)$, i.e. the probability associated to obtaining some output level $Q \in [\underline{q}, \bar{q}]$. Moreover, assume that the principal exhibits subjective expected utility (Savage, 1954). So, she quantifies the uncertainties associated to obtaining any output level in $[\underline{q}, \bar{q}]$ using subjective probabilities, and evaluates the desirability of the event-contingent wage schedule $t(q)$ using mathematical expectations.

On the other hand, the agent is non-EUT. He exhibits probabilistic sophistication, i.e. is also able to quantify uncertainties using probabilities, but evaluates the resulting probabilities using RDU. Formally, there exists a probability measure μ on $[\underline{q}, \bar{q}]$ such that $t(q)$ is evaluated by the agent using the following objective function

$$RDU(t,e) = \int_{\bar{q}}^{\underline{q}} u\Big(t(q)\Big)\, \mathrm{d}W_s\Big(1 - \mu(q|e)\Big) - c(e). \tag{4}$$

Probabilistic sophistication for non-EUT preferences and when probabilities are conditional has been defined by Machina and Schmeidler (1992).

We compare the considered situation of ambiguity, i.e. when $F(q|e)$ is not known, to the standard situation of risk that was studied in the main body of the paper, i.e. when $F(q|e)$ is known. We say, that in each situation the agent is facing a different *source of uncertainty*. In our framework, this amounts to the worker facing different types of tasks; a monotone and repetitive task that enables the calculation of frequencies, leading to decision-making under risk, and a task that involves many more irregularities and external factors, rendering probability estimations difficult.

Suppose that probabilistic sophistication holds within sources of uncertainty, but not necessarily between sources of uncertainty (Chew and Sagi, 2006, 2008). Accordingly, probability weighting depends on the considered source. Denote by W_s the weighting function under ambiguity and assume that it adopts the properties of Assumption 3. Moreover, let w, as it has been done through the paper, be the probability weighting function under risk.

When W_s is convex and it is so to a greater extent than w, the agent is ambiguity averse. This ambiguity attitude is consistent with robust empirical phenomena such as the "home bias" or the "Ellsberg paradox" (Ellsberg, 1961; French and Poterba, 1991). In that case, Proposition 4 captures the optimal contract. A result that echoes the finding in Lang (2017) that a constant wage can be optimal. However, we further show that a strong degree of pessimism toward ambiguity is required for this result to emerge.

When the non-sabotage constraint is ignored (See the following sections), the optimal contract given to an ambiguity averse agent is non-monotonic, i.e. first-increasing and then decreasing in performance, capturing the result in Kellner (2017). We go beyond by showing that non-monotonicities also emerge under likelihood insensitivity. Therefore, they are not uniquely due to ambiguity aversion but appear when large weights are given to worse events. Moreover, Corollary 4, demonstrates that stronger degrees of ambiguity aversion, i.e. as W_s becomes more convex while keeping the convexity of w fixed, leads to flatter contracts. Thus accommodating Ghirardato (1994)'s result that under stronger ambiguity aversion an action can be implemented for a "uniformly lower incentive scheme."

Recent research shows that ambiguity aversion is not universal. In fact, individuals tend to overweight the likelihood of rare events (Abdellaoui et al., 2011; Baillon and Emirmahmutoglu, 2018; Baillon et al., 2018; Trautmann and Van De Kuilen, 2015). This can be captured by $W_s(p)$ being subadditive in the sense of Definition 6. Under those ambiguity attitudes, the contract in

Proposition 6 applies. When subadditivity of W_s is stronger than that of w, the optimal contract under ambiguity compares to that under risk as explained in Corollary 7. Therefore, our framework also accommodates the result in Vierø (2014); indexing the performance measure to some ambiguous source of information is desirable as it takes advantage of ambiguity-seeking attitudes. In that case, the principal would benefit from the incentive compatibility constraint slacking, or from paying less for likely events while paying more for unlikely ones.[18]

5.2 Robustness

5.2.1 Continuous effort

In our model, the decision maker's action was assumed to be binary. This assumption is convenient inasmuch as it guarantees a solution to the principal's problem under relatively mild conditions. This tractability might be at the cost of generalizability if our results fail to hold when richer action spaces are considered.

In Appendix C, we solve the model in a setting in which effort is continuous. We find that our results emerge under well-known regularity conditions that guarantee a solution, i.e. the convexity of the density function (Mirrlees, 1999; Rogerson, 1985), plus some additional requirements on the cost function. Therefore, our findings are not an artifact of action space assumed in the model.

5.2.2 Adverse Selection

We considered a setting in which the principal perfectly knows the agent's risk attitude. While this assumption is typically made in moral hazard models, its limitation becomes more prominent in our framework because risk attitudes are richer. In Appendix D, we consider a model in which this assumption is relaxed. Consequently, the principal's goal is to first screen agents according

[18] Notice how this result emerges in our setting without requiring ambiguity seeking for all events, amounting to W_s being everywhere concave.

to their risk preference to then incentivize high effort. In other words, we study a framework of adverse selection followed by moral hazard.

Our analyses demonstrate that the solutions presented in Sections 3 and 4 are crucial to solve this more involved problem. These contracts are included in the optimal solution because they incentivize high effort provided that screening is successful. Moreover, to ensure screening, these contracts are enriched with information rents. An enhancement that ensures that the most efficient types are disincentivized from engaging in a strategy of mimicking the least efficient types. We note that these rents enter as lump-sums and do not change incentives, i.e. shapes of optimal contracts are left unchanged.

5.2.3 General Performance Measure

Our model considered a setting in which higher realizations of performance lead to higher profits. Consequently, pay schedules that weakly increase in performance are in place. In this extension, we study a framework in which the principal faces a general performance measure. This measure does not have a specific link to output on the task, and thus might render unnecessary the requirement that pay weakly increases in performance.

As in the standard model, we assume that the principal's objective is to minimize the agent's compensation while implementing high effort. Moreover, we focus on the interesting case in which the agent's compensation does not need to be monotonically increasing in output on the task. In the following we characterize optimal contracts resulting from our analyses without that restriction.

Under optimism, the contracts presented in Propositions 1 and Proposition 3 prevail as solutions to this modified problem. Even if monotonicity is not directly imposed, this agent is best incentivized with contracts that offer higher payments in exchange for higher performance.

A more interesting result emerges when the agent is pessimistic. The following Corollary presents the optimal contracts given to that agent.

Corollary 9. *Let Assumptions 1 and 3 hold. Under pessimism and for a general measure of performance:*

(i) The first-best, t_o^{fb}, is everywhere decreasing in q.

(ii) The second-best, t_o^{sb}, increases in q up to a threshold $q_h \in (\underline{q}, \bar{q})$ after which it decreases in q.

Pessimism is exploited with contracts that include transfers that can decrease in performance. When effort is observable, the pessimist considers to be fully insured with a contract that offers the highest transfers at low realizations and the lowest payments at high realizations. Such an insurance contract reflects the agent's perception that worse outcomes are more likely to realize. Moreover, in a situation of moral hazard, incentives are concentrated at realizations that the pessimistic agent deems to be more likely. As a result, transfers increase in performance at the low-end of the output space and decrease in performance thereafter.

Next, we characterize the optimal contracts for a likelihood insensitive agent.

Corollary 10. *Let Assumptions 1 and 3 hold. Under likelihood insensitivity and for a general measure of performance:*

(i) The first-best, t_L^{fb}, decreases in q up to a threshold $\tilde{q} \in (\underline{q}, \bar{q})$ after which it increases in q.

(ii) The second-best, t_L^{sb}, decreases in q in the compact interval $q_{\in}(q_{\mathcal{M}}, q_{\mathcal{S}})$ where $q_{\mathcal{M}}, q_{\mathcal{M}} \in (\underline{q}, \bar{q})$ but after and before this interval it increases in q.

As in the main body of the paper, likelihood insensitivity is targeted with a contract that implements incentives at both ends of the output space. At these extremes, the contract transfers increase in performance. However, the contract also decreases in performance for intermediate outcome levels. This is due to a combination of incentives not working in that segment and the principal wanting to save on compensation costs.

6 Conclusion

In this paper we show how the optimal implementation of incentives crucially depends on the agent's perception of probabilities. Motivational and cognitive deviations from expected utility can lead to contracts that do not require or cannot implement incentive compatibility. These solutions can resemble payment schemes observed in practice. For example, performance-insensitive salaries under strong pessimism, long-shot contracts under moderate optimism, and option-like contracts under likelihood insensitivity. We thus provide a foundation for simple contracts based on preference.

Appendix A: Proofs

Lemma 1

Proof. Suppose $\lim_{p\to 0} w'(p) = +\infty$ but, to set up the contradiction, that $\lim_{p\to 0} w''(p) \neq -\infty$. Hence, there exists $\bar{p} \in (0,1)$ such that, for $p \in [0,\bar{p}]$ and $B > 0$, $w''(p) > -B$. Integrating both sides of this inequality over $[p_0, p_1] \subseteq [0,\bar{p}]$ yields $w'(p_1) - w'(p_0) > -(p_1 - p_0)B$, and looking at the limit as p_0 goes to 0 gives $\lim_{p_0\to 0} w'(p_0) < Bp_1 + w'(p_1)$, which contradicts $\lim_{p\to 0} w'(p) = +\infty$. Hence, it must be that $\lim_{p\to 0} w''(p) = -\infty$.

Similarly, suppose $\lim_{p\to 0} w'(p) = +\infty$ but $\lim_{p\to 0} \frac{w''(p)}{w'(p)} \neq -\infty$. So for $p \in [0,\bar{p}]$ and $B > 0$, $\frac{w''(p)}{w'(p)} > -B$. Integrating over $[p_0, p_1] \subseteq [0,\bar{p}]$ yields

$$\begin{aligned} \ln w'(p_1) - \ln w'(p_0) = \ln \frac{w'(p_1)}{w'(p_0)} &> -B(p_1 - p_0) \\ \Leftrightarrow w'(p_0) &< \frac{w'(p_1)}{\exp\Big(-B(p_1 - p_0)\Big)} \end{aligned}$$

and looking at the limit as p_0 goes to 0 yields $\lim_{p_0\to 0} w'(p_0) < \frac{w'(p_1)}{\exp(-Bp_1)}$. Therefore, $w'(p)$ must be bounded as well as p approaches 0, which contradicts $\lim_{p\to 0} w'(p) = +\infty$. So it must be that $\lim_{p\to 0} \frac{w''(p)}{w'(p)} = -\infty$. ■

Lemma 2

Proof. Suppose $\lim_{p\to 1} w'(p) = +\infty$ but $\lim_{p\to 1} w''(p) \neq +\infty$. So there exists $\underline{p} \in (0,1)$ such that, for $p \in \Big[\underline{p}, 1\Big]$ and $B > 0$, $w''(p) < B$. Integrating both sides over $[p_0, p_1] \subseteq \Big[\underline{p}, 1\Big]$ and taking the limit as p_1 goes to 1 yields $\lim_{p_1\to 1} w'(p_1) < w'(p_0) + B - p_0 B$, contradicting $\lim_{p\to 1} w'(p) = +\infty$, so $\lim_{p\to 1} w''(p) = +\infty$.

Next, suppose $\lim_{p\to 1} w'(p) = +\infty$ but $\lim_{p\to 1} \frac{w''(p)}{w'(p)} \neq +\infty$. So for $p \in [\underline{p}, 1]$ and $B > 0$, $\frac{w''(p)}{w'(p)} < B$. Integrating over $[p_0, p_1]$ and taking the limit as p_1 goes to 1 yields $\lim_{p_1\to 1} w'(p_1) < \exp\Big(B(1 - p_0)\Big) \cdot w'(p_0)$, contradicting $\lim_{p\to 1} w'(p) = +\infty$, so $\lim_{p\to 1} \frac{w''(p)}{w'(p)} = +\infty$. ■

Lemma 5. *If* $\lim_{p\to 1} w'(p) = 0$*, then* $\lim_{p\to 1} w''(p) < 0$ *and* $\lim_{p\to 1} \frac{w''(p)}{w'(p)} = -\infty$.

Proof. Suppose $\lim_{p\to 1} w'(p) = 0$ but $\lim_{p\to 1} w''(p) \geq 0$. So, for $p \in [\underline{p}, 1]$ and $B \geq 0$, $w''(p) \geq B$. Integrating over $[p_0, p_1] \subseteq \left[\underline{p}, 1\right]$ and taking the limit as p_1 goes to 1 yields $\lim_{p_1\to 1} w'(p_1) > w'(p_0) + B - p_0 B > 0$, contradicting $\lim_{p\to 1} w'(p) = 0$. Therefore, $\lim_{p\to 1} w''(p) < 0$.

Next, suppose $\lim_{p\to 1} w'(p) = 0$ but $\lim_{p\to 1} \frac{w''(p)}{w'(p)} \neq -\infty$. So for $p \in \left[\underline{p}, 1\right]$, $\frac{w''(p)}{w'(p)} > -B$. Integrating over $[p_0, p_1] \subseteq \left[\underline{p}, 1\right]$ and taking the limit as p_1 goes to 1 yields $\lim_{p_1\to 1} w'(p_1) > \exp\Big(-B(1-p_0)\Big) \cdot w'(p_0) > 0$. This contradicts $\lim_{p\to 1} w'(p) = 0$, so $\lim_{p\to 1} \frac{w''(p)}{w'(p)} = -\infty$. ■

Lemma 6. *If* $\lim_{p\to 0} w'(p) = 0$*, then* $\lim_{p\to 0} w''(p) > 0$ *and* $\lim_{p\to 0} \frac{w''(p)}{w'(p)} = +\infty$.

Proof. Suppose $\lim_{p\to 0} w'(p) = 0$ but $\lim_{p\to 0} w''(p) \leq 0$. So, for $p \in \left[0, \bar{p}\right]$ and $B \geq 0$, $w''(p) \leq -B$. Integrating over $[p_0, p_1] \subseteq [0, \bar{p}]$ and taking the limit as p_0 goes to 0 yields $\lim_{p_0\to 0} w'(p_0) \geq w'(p_1) + p_1 B > 0$, contradicting $\lim_{p\to 0} w'(p) = 0$. Hence, $\lim_{p\to 0} w''(p) > 0$.

Next, suppose $\lim_{p\to 0} w'(p) = 0$ but $\lim_{p\to 0} \frac{w''(p)}{w'(p)} \neq +\infty$. So, for $p \in \left[0, \bar{p}\right]$ and $B > 0$, $\frac{w''(p)}{w'(p)} \leq B$. Again integrating over $[p_0, p_1] \subseteq [0, \bar{p}]$ and taking the limit as p_0 goes to 0 yields $\lim_{p_0\to 0} w'(p_0) > \frac{w'(p_1)}{\exp(Bp_1) > 0}$. This contradicts $\lim_{p\to 0} w'(p) = 0$, so $\lim_{p\to 0} \frac{w''(p)}{w'(p)} = +\infty$. ■

Proposition 1

Proof. Denoting the Lagrange multiplier of the agent's participation constraint by ν, the Lagrangian of the principal's problem writes as:

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

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

$$-f(q|\bar{e}) + \nu u'\Big(t^{fb}(q)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e}) = 0 \tag{5}$$

and, after re-arranging, we get

$$\frac{1}{u'\Big(t^{fb}(q)\Big)w'\Big(1-F(q|\bar{e})\Big)}=\nu. \tag{6}$$

By assumption, $u'(t) > 0$ and $w'(p) > 0$, so $\nu > 0$. The participation constraint binds at the optimum.

To investigate the shape of $t^{fb}(q)$ we differentiate (5) with respect to q, giving us

$$t^{fb\prime}(q)=\frac{u'\Big(t^{fb}(q)\Big)}{u''\Big(t^{fb}(q)\Big)}\frac{w''\Big(1-F(q|\bar{e})\Big)}{w'\Big(1-F(q|\bar{e})\Big)}f(q|\bar{e}). \tag{7}$$

Expected utility. Under expected utility, $w(p)=p$, $w'(p)=1$ and $w''(p)=0$, so the right-hand side of (7) is 0. Hence, the first-best contract given to the EU agent, $t_{EU}^{fb}(q)$, is everywhere constant and satisfies

$$\frac{1}{u'\left(t_{EU}^{fb}(q)\right)}=\nu.$$

Optimism. If the agent exhibits optimism, we have $w'(p)>0$ and $w''(p)<0$ (Assumption 3 and Definition 1). Under these conditions the right-hand side of (7) is positive, implying that the first-best contract given to the optimist, $t_O^{fb}(q)$, is everywhere increasing in q.

To better understand the shape of $t_O^{fb}(q)$ we look at its behavior at extremes. From Definition 1 and Lemma 1 we know that $\lim_{p\to 0}\frac{w''(p)}{w'(p)}=-\infty$. Since $u''<0$, it follows from (7) that $\lim_{q\to\bar{q}} t_O^{fb\prime}(q)=+\infty$. Moreover, from Definition 1 and Lemma 5 we have $\lim_{p\to 1}\frac{w''(p)}{w'(p)}=-\infty$, implying $\lim_{q\to\underline{q}} t_O^{fb\prime}(q)=+\infty$.

Pessimism. If instead the agent exhibits pessimism, we have $w'(p)>0$ and $w''(p)>0$ (Assumption 3 and Definition 2). Now, the right-hand side of (7) is strictly negative, implying that the first-best contract given to the pessimist, $t_P^{fb}(q)$, is strictly decreasing in q.

A contract decreasing in output is undesirable. It leads to sabotage; the agent wanting to destroy effort. We apply Myerson (1981)'s ironing. Since $t_P^{fb}(q)$ is decreasing everywhere, the modified solution is the ironed solution for all q:

$$\tilde{t}_P^{fb}(q) = \frac{\int_{\underline{q}}^{\bar{q}} t_P^{fb}(q)\,\mathrm{d}q}{\bar{q} - \underline{q}},$$

which means that $\tilde{t}_P^{fb}(q)$ is everywhere constant.

Next, we show that $\tilde{t}_P^{fb}(q) < t_{EU}^{fb}(q)$. Using integration by parts, we obtain

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_P^{fb}(q)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})\,\mathrm{d}q = u\Big(t_P^{fb}(\underline{q})\Big) + \int_{\underline{q}}^{\bar{q}} u'\Big(t_P^{fb}(q)\Big) t_P^{fb\prime}(q) w\Big(1 - F(q|\bar{e})\Big)\,\mathrm{d}q. \quad (8)$$

If $w(p) = p \Leftrightarrow w'(p) = 1$, (8) becomes

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_P^{fb}(q)\Big) f(q|\bar{e})\,\mathrm{d}q = u\Big(t_P^{fb}(\underline{q})\Big) + \int_{\underline{q}}^{\bar{q}} u'\Big(t_P^{fb}(q)\Big) t_P^{fb\prime}(q) \Big(1 - F(q|\bar{e})\Big)\,\mathrm{d}q. \quad (9)$$

Subtracting (9) from (8) gives

$$\int_{\underline{q}}^{\bar{q}} u'\Big(t_P^{fb}(q)\Big) t_P^{fb\prime}(q) \Big(w\Big(1 - F(q|\bar{e})\Big) - \Big(1 - F(q|\bar{e})\Big)\Big)\,\mathrm{d}q. \quad (10)$$

Because $t_P^{fb\prime}(q) < 0$ and under pessimism $w\Big(1 - F(q|\bar{e})\Big) - \Big(1 - F(q|\bar{e})\Big) < 0$, (10) is positive, so

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_P^{fb}(q)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})\,\mathrm{d}q > \int_{\underline{q}}^{\bar{q}} u\Big(t_P^{fb}(q)\Big) f(q|\bar{e})\,\mathrm{d}q. \quad (11)$$

Moreover, the first-best contract t_P^{fb} must satisfy the participation constraint:

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_P^{fb}(q)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e}) \, \mathrm{d}q = \bar{U}$$

so from (11) we get

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_P^{fb}(q)\Big) f(q|\bar{e}) \, \mathrm{d}q < \bar{U}.$$

Since the first-best contract offered to the expected-utility agent, t_{EU}^{fb}, for whom $w(p) = p$, also satisfies the participation constraint:

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_{EU}^{fb}(q)\Big) f(q|\bar{e}) \, \mathrm{d}q = \bar{U},$$

Hence,

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_P^{fb}(q)\Big) f(q|\bar{e}) \, \mathrm{d}q < \int_{\underline{q}}^{\bar{q}} u\Big(t_{EU}^{fb}(q)\Big) f(q|\bar{e}) \, \mathrm{d}q,$$

which is implied by $t_P^{fb}(q) < t_{EU}^{fb}(q)$. Because t_{EU}^{fb} is also everywhere constant in output, it must be that $\tilde{t}_P^{fb}(q) < t_{EU}^{fb}(q)$.

Therefore, offering $\tilde{t}_P^{fb}(q)$ which is, by construction, constant in performance yields utility lower than $\bar{U}$, the utility generated by the constant contract $t_{EU}^{fb}(q)$. Consequently, contract $\tilde{t}_P^{fb}(q)$ would be rejected. To ensure participation with a constant contract, the principal must offer the pessimist the contract $t_{EU}^{fb}(q)$. ∎

Lemma 7. *If agent i is more optimistic than agent j, then:*

1. $-\frac{w_i''(p)}{w_i'(p)} > -\frac{w_j''(p)}{w_j'(p)} \ \forall p \in (0,1)$;
2. $w_i(p) > w_j(p) \ \forall p \in (0,1)$;
3. *There exists a unique $p_k \in (0,1)$ such that $w_i'(p_k) = w_j'(p_k)$, this point becomes smaller the more optimistic i is with respect to j.*

If agent i is more pessimistic that agent j, the inequalities in 1. and 2. are reversed, and the unique point in 3. becomes larger.

Proof. **Part 1.** If agent i is more optimistic than agent j, $w_i(p) = \theta\Big(w_j(p)\Big)$. Note that

$$\frac{w_i''(p)}{w_i'(p)} = \frac{\theta''\Big(w_j(p)\Big)}{\theta'\Big(w_j(p)\Big)} w_j'(p) + \frac{w_j''(p)}{w_j'(p)}. \tag{12}$$

Because $\theta'' < 0$, it must be that

$$-\frac{w_i''(p)}{w_i'(p)} > -\frac{w_j''(p)}{w_j'(p)}. \tag{13}$$

If instead i is more pessimistic than j, similar steps lead to $\frac{w_i''(p)}{w_i'(p)} > \frac{w_j''(p)}{w_j'(p)}$.

Part 2. Let $p_0, p_1 \in [0,1]$ such that $p_1 > p_0$. Integrate the equation in (13) over $[p_0, p_1]$ to get

$$\begin{aligned}
\int_{p_0}^{p_1} -\frac{w_i''(s)}{w_i'(s)}\,\mathrm{d}s &> \int_{p_0}^{p_1} -\frac{w_j''(s)}{w_j'(s)}\,\mathrm{d}s \\
\Leftrightarrow -\ln w_i'(p_1) + \ln w_i'(p_0) &> -\ln w_j'(p_1) + \ln w_j'(p_0) \\
\Leftrightarrow \ln\left(\frac{w_j'(p_1)}{w_j'(p_0)}\right) &> \ln\left(\frac{w_i'(p_1)}{w_i'(p_0)}\right) \\
\Leftrightarrow \frac{w_j'(p_1)}{w_j'(p_0)} &> \frac{w_i'(p_1)}{w_i'(p_0)}.
\end{aligned}$$

Integrating the resulting expression over the range of p_0 gives

$$\begin{aligned}
w_i'(p_1)\int_0^{p_1} w_j'(s)\,\mathrm{d}s &< w_j'(p_1)\int_0^{p_1} w_i'(s)\,\mathrm{d}s \\
\Leftrightarrow w_i'(p_1)w_j(p_1) &< w_j'(p_1)w_i(p_1) \\
\Leftrightarrow \frac{w_j'(p_1)}{w_j(p_1)} &> \frac{w_i'(p_1)}{w_i(p_1)}.
\end{aligned}$$

Integrating again but now over the range of p_1 gives

$$\int_{p_0}^{1} \frac{w_j'(s)}{w_j(s)}\,\mathrm{d}s > \int_{p_0}^{1} \frac{w_i'(s)}{w_i(s)}\,\mathrm{d}s$$
$$\Leftrightarrow \ln w_i(1) - \ln w_i(p_0) < \ln w_j(1) - \ln w_j(p_0)$$
$$\Leftrightarrow w_i(p) > w_j(p)$$

since p_0 can be any $p \in [0,1)$. Similar steps lead to $w_i(p) < w_j(p)$ when i is more pessimistic than j.

Part 3. Suppose that $w_i'(p) < w_j'(p)$ for all $p \in (0,1)$. From Assumption 3, $w_i(0) = w_j(0)$ and $w_i(1) = w_j(1)$. Hence, $\int_0^1 w_j'(p)\mathrm{d}p = w_j(1) - w_j(0) = 1 > \int_0^1 w_j'(p)\mathrm{d}p$. Contradicting the assumption that $w_i(1) = 1$. A similar rationale disregards $w_i'(p) > w_j'(p)$ for all $p \in (0,1)$. Hence, if $w_i'(p) \leq w_j(p)$ holds, it must do so for some segment in $p \in (0,1)$.

Let $w_J(p) := \eta\big(w_j(p)\big)$ where η is a concave, increasing, and continuous probability weighting function. Lemma 1 and Lemma 5 show that $\lim_{p\to 0} w'(p) = +\infty$ and $\lim_{p\to 1} w'(p) = 0$ for generic weighting function w. The first part of this Lemma implies that $-\frac{w_J''(p)}{w_J'(p)} > -\frac{w_j''(p)}{w_j'(p)}$ $\forall p \in (0,1)$. Therefore, $w_J'(p)$ tends to infinity faster than $w_j'(p)$ as $p \to 0^+$.

Assumption 3 states that, under optimism, $w'(p)$ is decreasing and continuous. These properties together with $-\frac{w_J''(p)}{w_J'(p)} > -\frac{w_j''(p)}{w_j'(p)}$ $\forall p \in (0,1)$, that $w_J'(p)$ tends to infinity faster than $w_j(p)$ as $p \to 0^+$, and the fact that $\lim_{p\to 1} w'(p) = 0$, imply that there exists a unique point $p_k \in (0,1)$ such that $w_J'(p_k) = w_j'(p_k)$. For $p < p_k$ then $w_J'(p) > w_j'(p)$ but instead $w_J'(p) < w_j'(p)$ if $p > p_k$.

Next, let $w_i := \theta\big(w_J(p)\big)$ where θ is a concave, increasing, and continuous function. Thus $-\frac{w_i''(p)}{w_i'(p)} > -\frac{w_J''(p)}{w_J'(p)}$ $\forall p \in (0,1)$ and, using the reasoning given above, $w_i'(p)$ tends to infinity faster than $w_J'(p)$ as $p \to 0^+$. Hence, the point $p_l \in (0,1)$ such that $w_i'(p_l) = w_J'(p_l)$ is such that $p_l < p_k$. ■

Corollary 1

Proof. Using $\bar{p} := 1 - F(q|\bar{e})$, agent i receives a higher-powered first-best contract when

$$\frac{\mathrm{d}t_{O,i}^{fb}(q)}{\mathrm{d}q} > \frac{\mathrm{d}t_{O,j}^{fb}(q)}{\mathrm{d}q}$$
$$\Leftrightarrow \frac{u'\Big(t_{O,i}^{fb}(q)\Big)}{u''\Big(t_{O,i}^{fb}(q)\Big)} \frac{w_i''(\bar{p})}{w_i'(\bar{p})} f(q|\bar{e}) > \frac{u'\Big(t_{O,j}^{fb}(q)\Big)}{u''\Big(t_{O,j}^{fb}(q)\Big)} \frac{w_j''(\bar{p})}{w_j'(\bar{p})} f(q|\bar{e}).$$

After cancelling out the $f(q|\bar{e})$ on both sides and using (12), the inequality given above becomes

$$\frac{u'\Big(t_{O,i}^{fb}(q)\Big)}{u''\Big(t_{O,i}^{fb}(q)\Big)} \left\{ \frac{\theta''\Big(w_j(\bar{p})\Big)}{\theta'\Big(w_j(\bar{p})\Big)} w_j'(\bar{p}) + \frac{w_j''(\bar{p})}{w_j'(\bar{p})} \right\} > \frac{u'\Big(t_{O,j}^{fb}(q)\Big)}{u''\Big(t_{O,j}^{fb}(q)\Big)} \frac{w_j''(\bar{p})}{w_j'(\bar{p})}$$
$$\Leftrightarrow \frac{w_j''(\bar{p})}{w_j'(\bar{p})} \left\{ \frac{u'\Big(t_{O,i}^{fb}(q)\Big)}{u''\Big(t_{O,i}^{fb}(q)\Big)} - \frac{u'\Big(t_{O,j}^{fb}(q)\Big)}{u''\Big(t_{O,j}^{fb}(q)\Big)} \right\}$$
$$+ \frac{u'\Big(t_{O,i}^{fb}(q)\Big)}{u''\Big(t_{O,i}^{fb}(q)\Big)} \frac{\theta''\Big(w_j(\bar{p})\Big)}{\theta'\Big(w_j(\bar{p})\Big)} w_j'(\bar{p}) > 0, \tag{14}$$

which in turn can be rewritten as

$$-\frac{\theta''\Big(w_j(\bar{p})\Big)}{\theta'\Big(w_j(\bar{p})\Big)} w_j'(\bar{p}) > \frac{w_j''(\bar{p})}{w_j'(\bar{p})} \left(1 - \frac{\frac{u''(t_{O,i}^{fb}(q))}{u'(t_{O,i}^{fb}(q))}}{\frac{u''(t_{O,j}^{fb}(q))}{u'(t_{O,j}^{fb}(q))}} \right). \tag{15}$$

We know that $u'' < 0$. Further, if agent j is optimistic, $w_j'' < 0$, and if agent i is more optimistic than agent j, $\theta'' < 0$ from Definition 3. Therefore, (15) holds if

$$-\frac{u''(t_{O,j}^{fb}(q))}{u'(t_{O,j}^{fb}(q))} > -\frac{u''(t_{O,i}^{fb}(q))}{u'(t_{O,i}^{fb}(q))}. \tag{16}$$

Furthermore, (15) shows that under $-\frac{u''(t_{O,j}^{fb}(q))}{u'(t_{O,j}^{fb}(q))} < -\frac{u''(t_{O,i}^{fb}(q))}{u'(t_{O,i}^{fb}(q))}$ the concavity of θ needs to be sufficiently large to guarantee that inequality. ■

Corollary 2

Proof. Denote by w_i and w_j the probability weighting functions of agents i and j. Assume that i is more pessimistic than j. We follow similar steps at for the proof of Proposition 1 in the case of pessimism.

Using integration by parts, we have for j

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_{P,j}^{fb}(q)\Big) w_j'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e}) \, \mathrm{d}q = u\Big(t_{P,j}^{fb}(\underline{q})\Big) + \int_{\underline{q}}^{\bar{q}} u'\Big(t_{P,j}^{fb}(q)\Big) t_{P,j}^{fb\,\prime}(q) w_j\Big(1 - F(q|\bar{e})\Big) \, \mathrm{d}q. \quad (17)$$

If i were to get the same contract, he would derive utility

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_{P,j}^{fb}(q)\Big) w_i'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e}) \, \mathrm{d}q = u\Big(t_{P,j}^{fb}(\underline{q})\Big) + \int_{\underline{q}}^{\bar{q}} u'\Big(t_{P,j}^{fb}(q)\Big) t_{P,j}^{fb\,\prime}(q) w_i\Big(1 - F(q|\bar{e})\Big) \, \mathrm{d}q. \quad (18)$$

Subtracting the right-hand side of (17) from that of (18) gives

$$\int_{\underline{q}}^{\bar{q}} u'\Big(t_{P,j}^{fb}(q)\Big) t_{P,j}^{fb\,\prime}(q) \Big(w_i\Big(1 - F(q|\bar{e})\Big) - w_j\Big(1 - F(q|\bar{e})\Big)\Big) f(q|e) \, \mathrm{d}q. \quad (19)$$

Because $t_{P,j}^{fb\,\prime}(q) < 0$ (Proposition 1) and $w_i(p) < w_j(p)$ (Lemma 7), the expression in (19) is positive, so

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_{P,j}^{fb}(q)\Big) w_i'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e}) \, \mathrm{d}q > \int_{\underline{q}}^{\bar{q}} u\Big(t_{P,j}^{fb}(q)\Big) w_j'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e}) \, \mathrm{d}q. \quad (20)$$

For j, the candidate solution from the first-order approach, $t_{P,j}^{fb}(q)$, ensures that the participation constraint binds

$$\int_{\underline{q}}^{\bar{q}} u\Big(t^{fb}_{P,j}(q)\Big)w'_j\Big(1-F(q|\bar{e})\Big)f(q|\bar{e})\,\mathrm{d}q = \bar{U},$$

so from (20)

$$\int_{\underline{q}}^{\bar{q}} u\Big(t^{fb}_{P,j}(q)\Big)w'_i\Big(1-F(q|\bar{e})\Big)f(q|\bar{e})\,\mathrm{d}q > \bar{U}.$$

Similarly, for i, the candidate solution form he first-order approach makes the participation constraint binds:

$$\int_{\underline{q}}^{\bar{q}} u\Big(t^{fb}_{P,i}(q)\Big)w'_i\Big(1-F(q|\bar{e})\Big)f(q|\bar{e})\,\mathrm{d}q = \bar{U}.$$

Therefore,

$$\int_{\underline{q}}^{\bar{q}} u\Big(t^{fb}_{P,j}(q)\Big)w'_i\Big(1-F(q|\bar{e})\Big)f(q|\bar{e})\,\mathrm{d}q > \int_{\underline{q}}^{\bar{q}} u\Big(t^{fb}_{P,i}(q)\Big)w'_i\Big(1-F(q|\bar{e})\Big)f(q|\bar{e})\,\mathrm{d}q.$$

which is implied by $t^{fb}_{P,j}(q) > t^{fb}_{P,i}(q)$. Thus, the ironed solutions exhibit, $\tilde{t}^{fb}_{P,j} > \tilde{t}^{fb}_{P,i}$. Proposition 1 shows that to ensure participation t^{fb}_{EU} is given to both i and j. Since $t^{fb}_{EU} > \tilde{t}^{fb}_{P,j} > \tilde{t}^{fb}_{P,i}$ the cost borne by the principal of not being able to implement $\tilde{t}^{fb}_{P,j}$ nor $\tilde{t}^{fb}_{P,i}$ is higher for i, the more pessimistic agent. ■

Proposition 2

Proof. This standard result comes from Holmstrom (1979). ■

Lemma 3

Proof. **Part 1.** From the definition of the W-MLRP, for all $q_0, q_1 \in [\underline{q}, \bar{q}]$ such that $q_1 \geq q_0$, we have

$$\frac{w'\Big(1-F(q_1|\underline{e})\Big)f(q_1|\underline{e})}{w'\Big(1-F(q_1|\bar{e})\Big)f(q_1|\bar{e})} \leq \frac{w'\Big(1-F(q_0|\underline{e})\Big)f(q_0|\underline{e})}{w'\Big(1-F(q_0|\bar{e})\Big)f(q_0|\bar{e})}$$

$$\Leftrightarrow w'\Big(1-F(q_1|\underline{e})\Big)f(q_1|\underline{e})w'\Big(1-F(q_0|\bar{e})\Big)f(q_0|\bar{e}) \leq w'\Big(1-F(q_0|\underline{e})\Big)f(q_0|\underline{e})w'\Big(1-F(q_1|\bar{e})\Big)f(q_1|\bar{e}). \tag{21}$$

Integrating both sides of the inequality with respect to q_0 from $\underline{q}$ to q_1 gives

$$w'\Big(1-F(q_1|\underline{e})\Big)f(q_1|\underline{e})\int_{\underline{q}}^{q_1} w'\Big(1-F(q_0|\bar{e})\Big)f(q_0|\bar{e})\,\mathrm{d}q_0 \leq w'\Big(1-F(q_1|\bar{e})\Big)f(q_1|\bar{e})\int_{\underline{q}}^{q_1} w'\Big(1-F(q_0|\underline{e})\Big)f(q_0|\underline{e})\,\mathrm{d}q_0$$

and, after rearranging and using $\int_{\underline{q}}^{q_1} w'\Big(1-F(q_0|e)\Big)f(q_0|e)\,\mathrm{d}q_0 = 1 - w\Big(1-F(q_1|e)\Big)$,

$$\frac{w'\Big(1-F(q_1|\underline{e})\Big)f(q_1|\underline{e})}{w'\Big(1-F(q_1|\bar{e})\Big)f(q_1|\bar{e})} \leq \frac{1-w\Big(1-F(q_1|\underline{e})\Big)}{1-w\Big(1-F(q_1|\bar{e})\Big)}. \tag{22}$$

Integrating (21) again, but now with respect to q_1 from q_0 to $\bar{q}$, gives

$$\frac{w\Big(1-F(q_0|\underline{e})\Big)}{w\Big(1-F(q_0|\bar{e})\Big)} \leq \frac{w'\Big(1-F(q_0|\underline{e})\Big)f(q_0|\underline{e})}{w'\Big(1-F(q_0|\bar{e})\Big)f(q_0|\bar{e})}. \tag{23}$$

Letting $q_0 = q_1 = q$ and combining (22) and (23) gives

$$\frac{w\Big(1-F(q|\underline{e})\Big)}{w\Big(1-F(q|\bar{e})\Big)} \leq \frac{1-w\Big(1-F(q|\underline{e})\Big)}{1-w\Big(1-F(q|\bar{e})\Big)}$$

$$\Leftrightarrow w\Big(1-F(q|\bar{e})\Big) \geq w\Big(1-F(q|\underline{e})\Big)$$

which proves the first part of the Lemma.

Part 2. Let $w(p) = p \Leftrightarrow w'(p) = 1$. The W-MLRP becomes

$$\frac{\mathrm{d}}{\mathrm{d}q}\frac{f(q|\underline{e})}{f(q|\bar{e})} \leq 0,$$

which is the MLRP. ■

Lemma 4

Proof. We have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{w'\big(1-F(q|\underline{e})\big)f(q|\underline{e})}{w'\big(1-F(q|\bar{e})\big)f(q|\bar{e})}\right) &= \frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{w'\big(1-F(q|\underline{e})\big)}{w'\big(1-F(q|\bar{e})\big)}\right)\frac{f(q|\underline{e})}{f(q|\bar{e})} \\
&\quad + \frac{w'\big(1-F(q|\underline{e})\big)}{w'\big(1-F(q|\bar{e})\big)}\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{f(q|\underline{e})}{f(q|\bar{e})}\right) \\
&= \frac{w'\big(1-F(q|\underline{e})}{w'\big(1-F(q|\bar{e})}\Bigg[\left(\frac{w''\big(1-F(q|\bar{e})\big)}{w'\big(1-F(q|\bar{e})\big)}f(q|\bar{e})\right. \\
&\quad \left. - \frac{w''\big(1-F(q|\underline{e})\big)}{w'\big(1-F(q|\underline{e})\big)}f(q|\underline{e})\right)\frac{f(q|\underline{e})}{f(q|\bar{e})} \\
&\quad + \frac{\mathrm{d}}{\mathrm{d}q}\frac{f(q|\underline{e})}{f(q|\bar{e})}\Bigg].
\end{aligned} \tag{24}$$

Notice that $\frac{\mathrm{d}}{\mathrm{d}q}\frac{f(q|\underline{e})}{f(q|\bar{e})} \leq 0$ due to the MLRP. So, according to (24) the WMLRP holds if

$$\frac{w''\big(1-F(q|\bar{e})\big)}{w'\big(1-F(q|\bar{e})\big)}f(q|\bar{e}) - \frac{w''\big(1-F(q|\underline{e})\big)}{w'\big(1-F(q|\underline{e})\big)}f(q|\underline{e}) \leq 0 \tag{25}$$

The Lemma follows immediately. ■

Lemma 8. *the W-MLRP holds if and only if*

$$\left(\frac{w''\big(1-F(q|\bar{e})\big)}{w'\big(1-F(q|\bar{e})\big)}f(q|\bar{e}) - \frac{w''\big(1-F(q|\underline{e})\big)}{w'\big(1-F(q|\underline{e})\big)}f(q|\underline{e})\right)\frac{f(q|\underline{e})}{f(q|\bar{e})} \leq -\frac{\mathrm{d}}{\mathrm{d}q}\frac{f(q|\underline{e})}{f(q|\bar{e})}.$$

Proof. The Lemma follows from the last equality in equation (24). ■

Proposition 3 & Proposition 4

Proof. Denote by ν the Lagrange multiplier of the agent's participation constraint, and μ, of the incentive compatibility constraint. The Lagrangian of the principal's maximization 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

$$\begin{aligned}-f(q|\bar{e}) + \mu\big[u'(t^{sb}(q))(w'(1 - F(q|\bar{e}))f(q|\bar{e}) - w'(1 - F(q|\underline{e}))f(q|\underline{e}))\big] \\ + \nu u'(t^{sb}(q))w'(1 - F(q|\bar{e}))f(q|\bar{e}) = 0,\end{aligned} \tag{26}$$

and, after re-arranging,

$$\frac{1}{u'\Big(t^{sb}(q)\Big)w'\Big(1 - F(q|\bar{e})\Big)} = \nu + \mu\left(1 - \frac{w'\Big(1 - F(q|\underline{e})\Big)f(q|\underline{e})}{w'\Big(1 - F(q|\bar{e})\Big)f(q|\bar{e})}\right). \tag{27}$$

Incentive constraint is binding We first show that $\mu > 0$ might not always hold the optimum. Suppose that $\mu = 0$. Then $t^{sb}(q) = t^{fb}(q)$, where $t^{fb}(q)$ is the first-best contract presented in Proposition 1.

Optimism Consider the case of an agent with optimism in the sense of Definition 1. From the complementary slackness condition from μ we get

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_O^{fb}(q)\Big) w'(1 - F(q|\bar{e})) f(q|\bar{e}) \, \mathrm{d}q - c$$
$$> \int_{\underline{q}}^{\bar{q}} u\Big(t_O^{fb}(q)\Big) w'(1 - F(q|\underline{e})) f(q|\underline{e}) \, \mathrm{d}q. \quad (28)$$

Integration by parts of (3) yields

$$u(t_O^{fb}(\underline{q})) + \int_{\underline{q}}^{\bar{q}} u'\Big(t_O^{fb}(q)\Big) t_O^{fb\prime}(q) w\Big(1 - F(q|e)\Big) \, \mathrm{d}q - c(e)$$

which we use to rewrite (28) as

$$\int_{\underline{q}}^{\bar{q}} u'\Big(t_O^{fb}(q)\Big) t_O^{fb\prime}(q) \bigg(w\Big(1 - F(q|\bar{e})\Big) - w\Big(1 - F(q|\underline{e})\Big) \bigg) \, \mathrm{d}q > c. \quad (29)$$

According to Lemma 3, Assumption 4 implies $w\left(1 - F(q|\bar{e})\right) \geq w\left(1 - F(q|\underline{e})\right)$ which, together with $t_O^{fb\prime}(q) > 0$ (Proposition 1) and $u'(t) > 0$ (Assumption 1), imply that the left-hand side of (29) is weakly positive. Since $w(p)$ and $u(t)$ are $\mathcal{C}^2$, and since c is a constant unbounded from above, there exists $\hat{c}_O > 0$ such that, for a given $t_O^{fb}(q)$,

- if $c \leq \hat{c}_O$, (29) holds: $\mu = 0$ and $t_O^{sb}(q) = t_O^{fb}(q)$; on the other hand,
- if $c > \hat{c}_O$, (29) does not hold: $\mu > 0$ and $t_O^{sb}(q)$ satisfies (27).

Pessimism Now consider the case of an agent with pessimism in the sense of Definition 2. From the complementary slackness condition corresponding to $\mu = 0$ we get

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_P^{fb}(q)\Big) w'(1 - F(q|\bar{e})) f(q|\bar{e}) \,\mathrm{d}q - c > \int_{\underline{q}}^{\bar{q}} u\Big(t_P^{fb}(q)\Big) w'(1 - F(q|\underline{e})) f(q|\underline{e}) \,\mathrm{d}q$$

$$\Leftrightarrow u\Big(t_P^{fb}(q)\Big) - c > u\Big(t_P^{fb}(q)\Big)$$

$$\Leftrightarrow -c > 0.$$

The second inequality is due to t_P^{fb} being constant in q (Proposition 1). The last inequality contradicts the assumption $c > 0$, so it must be that $\mu > 0$ for the pessimistic agent.

Shape of $t^{sb}(q)$ The second part of the proof analyzes the shape of $t^{sb}(q)$. We assume throughout that $\mu > 0$. Differentiate (26) with respect to q to obtain:

$$\begin{aligned} t^{sb\prime}(q) = & \frac{u'(t^{sb}(q)) w''(1 - F(q|\bar{e}))}{u''(t^{sb}(q)) w'\left(1 - F(q|\bar{e})\right)} f(q|\bar{e}) \\ & + \mu \frac{w'\left(1 - F(q|\bar{e})\right) u'(t^{sb}(q))^2}{u''(t^{sb}(q))} \frac{\mathrm{d}}{\mathrm{d}q} \left(\frac{w'\Big(1 - F(q|\underline{e})\Big) f(q|\underline{e})}{w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})} \right). \end{aligned} \tag{30}$$

We know that $\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{w'(1-F(q|\underline{e}))f(q|\underline{e})}{w'(1-F(q|\bar{e}))f(q|\bar{e})}\right) < 0$ (Assumption 4), $u'(t^{sb}(q)) > 0$, $u''(t^{sb}(q)) < 0$ (Assumption 1), and $w'(p) > 0$ (Assumption 3), so the second term on the right-hand side of (30) is always positive. The first term on the right-hand side of (30) is identical to the right-hand side of (7) in Proposition 1, which determined the shape of $t^{fb}(q)$.

Optimism When the agent exhibits optimism (Definition 1), $w''(p) < 0$ for all $p \in (0, 1)$, so the two terms on the right-hand side of (30) are positive. Hence, $t_O^{sb\prime}(q) > 0$ for all q.

We also study (30) at the extremes. From Definition 1 and Lemma 1 we know that $\lim_{p \to 0} \frac{w''(p)}{w'(p)} = -\infty$, so $\lim_{q \to \bar{q}} t_O^{sb\prime}(q) = +\infty$. Furthermore, Definition 1 and Lemma 5 give us $\lim_{p \to 1} \frac{w''(p)}{w'(p)} = -\infty$, so $\lim_{q \to \underline{q}} t_O^{sb\prime}(q) = +\infty$. Contract t_O^{sb} is high-powered at extremes.

Pessimism When the agent exhibits pessimism (Definition 2), $w''(p) > 0$ for all $p \in (0,1)$, so the first term on the right-hand side of (30) is negative while the second one is positive. Hence, the sign of $t_P^{sb\prime}(q)$ depends on which of these terms dominates the other, which in turn depends on the size of $w'(1-F(q|\bar{e}))$. When q decreases, $w'(1-F(q|\bar{e}))$ increases; the second term becomes larger and $t_P^{sb\prime}(q)$ is more likely to be positive. The opposite happens when q increases and $w'(1-F(q|\bar{e}))$ decreases.

To further formalize that contract shape, we study (30) at the extremes and under pessimism. We start with $q \to \bar{q}$. From Definition 2 and Lemma 6, we know that $\lim_{p\to 0} w'(p) = 0$ and $\lim_{p\to 0} \frac{w''(p)}{w'(p)} = +\infty$. Since $u'' < 0$, as q goes to $\bar{q}$ the first term on the right-hand side of (30) goes to $-\infty$ while the second goes to 0. Therefore, $\lim_{q\to\bar{q}} t_P^{sb\prime}(q) = -\infty$.

We ask whether t_P^{sb} ever increases with output; that is, whether $t_P^{sb\prime}(q) > 0$ for any segment in $[\underline{q}, \bar{q}]$, or equivalently, using (30),

$$
\begin{aligned}
&\frac{w''\Big(1-F(q|\bar{e})\Big)f(q|\bar{e})}{w'\Big(1-F(q|\bar{e})\Big)} \frac{1}{\mu u'(t^{sb}(q))w'\Big(1-F(q|\bar{e})\Big)} \\
&\qquad\qquad \leq -\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{w'\Big(1-F(q|\underline{e})\Big)f(q|\underline{e})}{w'\Big(1-F(q|\bar{e})\Big)f(q|\bar{e})}\right)
\end{aligned}
$$

for some q. We use

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{w'\Big(1-F(q|\underline{e})\Big)f(q|\underline{e})}{w'\Big(1-F(q|\bar{e})\Big)f(q|\bar{e})}\right) = &\frac{w'\Big(1-F(q|\underline{e})\Big)}{w'\Big(1-F(q|\bar{e})\Big)}\Bigg[\Bigg(\frac{w''\Big(1-F(q|\bar{e})\Big)}{w'\Big(1-F(q|\bar{e})\Big)}f(q|\bar{e}) \\
&- \frac{w''\Big(1-F(q|\underline{e})\Big)}{w'\Big(1-F(q|\underline{e})\Big)}f(q|\underline{e})\Bigg)\frac{f(q|\underline{e})}{f(q|\bar{e})} \\
&+ \frac{\mathrm{d}}{\mathrm{d}q}\frac{f(q|\underline{e})}{f(q|\bar{e})}\Bigg]
\end{aligned}
$$

to rewrite the inequality above as

$$\frac{f(q|\underline{e})}{f(q|\bar{e})}\left[\frac{w''\Big(1-F(q|\bar{e})\Big)}{w'\Big(1-F(q|\bar{e})\Big)}f(q|\bar{e})\left(\frac{1}{\mu u'(t^{sb}(q))w'\Big(1-F(q|\underline{e})\Big)\frac{f(q|\underline{e})}{f(q|\bar{e})}}+1\right)\right.$$
$$\left.-\frac{w''\Big(1-F(q|\underline{e})\Big)}{w'\Big(1-F(q|\underline{e})\Big)}f(q|\underline{e})\right]\leq-\frac{\mathrm{d}}{\mathrm{d}q}\frac{f(q|\underline{e})}{f(q|\bar{e})}. \quad (31)$$

From Assumption 3 we know that under pessimism $\lim_{q\to\underline{q}} w'\Big(1-F(q|\underline{e})\Big)=+\infty$. Further, the MLRP states that $\frac{f(q|\underline{e})}{f(q|\bar{e})}$ increases as q decreases. Therefore, the quantity

$$\frac{1}{\mu u'(t^{sb}(q))w'\Big(1-F(q|\underline{e})\Big)\frac{f(q|\underline{e})}{f(q|\bar{e})}} \quad (32)$$

goes to 0 as q goes to $\underline{q}$. All that is left is

$$\left(\frac{w''\Big(1-F(q|\bar{e})\Big)}{w'\Big(1-F(q|\bar{e})\Big)}f(q|\bar{e})-\frac{w''\Big(1-F(q|\underline{e})\Big)}{w'\Big(1-F(q|\underline{e})\Big)}f(q|\underline{e})\right)\frac{f(q|\underline{e})}{f(q|\bar{e})}\leq-\frac{\mathrm{d}}{\mathrm{d}q}\frac{f(q|\underline{e})}{f(q|\bar{e})},$$

which we know to hold from Lemma 8. Therefore, there exists an output level $q_h\in(\underline{q},\bar{q})$ such that ${t_P^{sb}}'(q)\geq 0$ if $q<q_h$ and ${t_P^{sb}}'(q)<0$ otherwise.

We implement ironing to avoid having ${t_P^{sb}}'(q)<0$ in $q\in(q_h,\bar{q}]$. To that end, find $q_{\mathcal{I}}$ that satisfies:

$$\int_{q_{\mathcal{I}}}^{q_h} t^{sb}(q)\,\mathrm{d}q-\int_{q_h}^{\bar{q}} t^{sb}(q)\,\mathrm{d}q=0. \quad (33)$$

There are two cases. If $\int_{\underline{q}}^{q_h} t_P^{sb}(q)\,\mathrm{d}q>\int_{q_h}^{\bar{q}} t_P^{sb}(q)\,\mathrm{d}q$, there exists $q_{\mathcal{I}}\in[\underline{q},q_h)$ that ensures (33) In that case ironing can be implemented and the modified solution is:

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

Instead, if $\int_{\underline{q}}^{q_h} t_P^{sb}(q)\,\mathrm{d}q\leq\int_{q_h}^{\bar{q}} t_P^{sb}(q)\,\mathrm{d}q$ ironing cannot be implemented and the contract is constant everywhere: the Principal cannot implement an incentive-compatible scheme. We assume that in this case the principal is still interested in contracting and offers the first-best. ■

Corollary 3

Proof. **Part 1.** Let agent i be more optimistic than agent j. Accordingly, $w_i(p) = \theta\big(w_j(p)\big)$ for all $p \in [0,1]$ (Definition 3).

From Proposition 3 we know that there exists a critical cost level $\hat{c}_{O,j}$, such that

$$\hat{c}_{O,j} := \int\limits_{\underline{q}}^{\bar{q}} u'\Big(t_{O,j}^{fb}(q)\Big) {t_{O,j}^{fb}}'(q) \Big(w_j\Big(1 - F(q|\bar{e})\Big) - w_j\Big(1 - F(q|\underline{e})\Big)\Big)\,\mathrm{d}q. \tag{34}$$

If agent i were given the same contract as that given to j, his critical cost level satisfies:

$$\hat{c}_{O,i} := \int\limits_{\underline{q}}^{\bar{q}} u'\Big(t_{O,j}^{fb}(q)\Big) {t_{O,j}^{fb}}'(q) \Big(w_i\Big(1 - F(q|\bar{e})\Big) - w_i\Big(1 - F(q|\underline{e})\Big)\Big)\,\mathrm{d}q. \tag{35}$$

The existence of $\hat{c}_{O,i}$ is guaranteed by the facts that $w_i\Big(1 - F(q|\bar{e})\Big) > w_i\Big(1 - F(q|\underline{e})\Big)$ (Lemma 3) and ${t_{O,j}^{fb}}'(q) > 0$ (Proposition 1) which imply that the right-hand side of (35) is positive. So there exists a set $c \in (0, \hat{c}_{O,i})$ in which $t_{O,j}^{fb}$ incentivizes agent i.

Consider $q < q^*$. For those output levels, optimism implies $w_i'\Big(1 - F(q|e)\Big) < 1$. Moreover, let $w_0 := w_j\Big(1 - F(q|\underline{e})\Big)$ and $w_1 := w_j\Big(1 - F(q|\bar{e})\Big)$. Note that $w_0, w_1 \in (0,1)$ and that $w_1 > w_0$ (Lemma 3).

Since $w_i'\Big(1 - F(q|e)\Big) = \theta'\Big(w_j\Big(1 - F(q|e)\Big)\Big) w_j'\Big(1 - F(q|e)\Big)$, we integrate the inequality $w_i'\Big(1 - F(q|e)\Big) < 1$ over $[w_0, w_1]$ to obtain

$$\begin{aligned} \int_{w_0}^{w_1} \theta'(s)\mathrm{d}s < \int_{w_0}^{w_1} \mathrm{d}s \Leftrightarrow \\ w_i\Big(1 - F(q|\bar{e})\Big) - w_i\Big(1 - F(q|\underline{e})\Big) < w_j\Big(1 - F(q|\bar{e})\Big) - w_j\Big(1 - F(q|\underline{e})\Big). \end{aligned} \tag{36}$$

Equation (36) together with (34) and (35) imply that $\hat{c}_{O,i} < \hat{c}_{O,j}$. Under stronger optimism the set $c < \hat{c}_O$ becomes smaller.

Part 2. The third part of Lemma 7 shows that the point $p_l \in (0,1)$ such that $w_i'(p_l) = w_j'(p_l)$ becomes smaller the more optimistic i is with respect to j. Let $w_j' = 1$. Accordingly, the output level q^* such that $w_i'\Big(1 - F(q^*|e)\Big) = 1$ takes place at a higher output level the more optimistic i is. Thus, increasing the length of the interval satisfying $q < q^*$.

■

Corollary 4

Proof. Let agent i be more pessimistic than agent j. Accordingly, the third part of Lemma 7 shows that the point $p_l \in (0,1)$ at which $w_i'(p_l) = w_j'(p_l)$ becomes larger the more pessimistic i is. Thus, the segment $p \in (p_l, 1]$ for which $w_i'(p) > w_j'(p)$ becomes smaller.

As a consequence, the output level $q^\varepsilon \in [\underline{q}, \bar{q}]$ such that $w_i'\Big(1 - F(q^\varepsilon|e)\Big) = \varepsilon$ for arbitrary small $\varepsilon > 0$ takes place at a lower output level the more pessimistic i is. Making the segment $q > q^\varepsilon$ for which $w_i'\Big(1 - F(q|e)\Big) < \varepsilon$ larger.

Eq. (32) shows that this tendency of $w_i'(p)$ becoming smaller as i becomes more pessimistic, makes more likely that $\frac{\mathrm{d}t_P^{sb}(q)}{\mathrm{d}q} < 0$ for a larger segment of q. Consequently, ironing, if still possible, requires a smaller value $q_{\mathcal{I}}$; the ironed segment becomes larger.

■

Corollary 5

Proof. We start by rewriting (27) as

$$\frac{1}{u'(t^{sb}(q))} f(q|\bar{e}) = \nu w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e}) + \mu w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e}) \\ - \mu w'\Big(1 - F(q|\underline{e})\Big) f(q|\underline{e}).$$

Integrating both sides with respect to q over $[\underline{q}, \bar{q}]$, and noting that

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

gives us

$$\nu = \int_{\underline{q}}^{\bar{q}} \frac{1}{u'(t^{sb}(q))} f(q|\bar{e})\,\mathrm{d}q = \mathbb{E}_{\bar{e}}\left(\frac{1}{u'(t^{sb}(q))}\right) \tag{37}$$

where $\mathbb{E}_{\bar{e}}$ is the expectation with respect to the probability distribution of q induced by $\bar{e}$. Hence, $\nu > 0$ and its value is the same for different agents with different probability weighting functions w.

After plugging (37) into (27) and multiplying by $u(t^{sb}(q))$, we obtain

$$\begin{aligned}&\mu u(t^{sb}(q))\Big[w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e}) - w'\Big(1 - F(q|\underline{e})\Big) f(q|\underline{e})\Big]\\ &= f(q|\bar{e})u(t^{sb}(q))\left[\frac{1}{u'(t^{sb}(q))} - \mathbb{E}_{\bar{e}}\left(\frac{1}{u'(t^{sb}(q))}\right) w'\Big(1 - F(q|\bar{e})\Big)\right].\end{aligned} \tag{38}$$

From the complementary slackness condition associated with μ we know

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

We can thus rewrite (38), after integrating with respect to q over $[\underline{q}, \bar{q}]$, as

$$\begin{aligned}
\mu c &= \int_{\underline{q}}^{\bar{q}} u(t^{sb}(q)) \left[\frac{1}{u'(t^{sb}(q))} - \mathbb{E}_{\bar{e}} \left(\frac{1}{u'(t^{sb}(q))} \right) w'\Big(1 - F(q|\bar{e})\Big) \right] f(q|\bar{e})\, \mathrm{d}q \\
&= \mathbb{E}_{\bar{e}} \left(\frac{u(t^{sb}(q))}{u'(t^{sb}(q))} \right) - \mathbb{E}_{\bar{e}} \left(\frac{1}{u'(t^{sb}(q))} \right) \int_{\underline{q}}^{\bar{q}} u(t^{sb}(q)) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})\, \mathrm{d}q \\
&= \mathbb{E}_{\bar{e}} \left(\frac{u(t^{sb}(q))}{u'(t^{sb}(q))} \right) - \mathbb{E}_{\bar{e}} \left(\frac{1}{u'(t^{sb}(q))} \right) \tilde{\mathbb{E}}_{\bar{e}} \Big(u(t^{sb}(q)) \Big).
\end{aligned} \tag{39}$$

where $\tilde{\mathbb{E}}_{\bar{e}}$ is the expectation as perceived by an agent who distorts probabilities. Since $\tilde{\mathbb{E}}_{\bar{e}}\Big(u(t^{sb}(q))\Big) < \mathbb{E}_{\bar{e}}\Big(u(t^{sb}(q))\Big)$ under pessimism, and the opposite under optimism, (39) implies $\mu_P > \mu_{EU} > \mu_O \geq 0$.

We rewrite (27), the first-order condition, again, but this time as

$$\begin{aligned}
\frac{1}{u'(t^{sb}_{nEU}(q))} &= \nu w'\left(1 - F(q|\bar{e})\right) + \\
&\qquad \mu_{nEU} w'\left(1 - F(q|\bar{e})\right) \left(1 - \frac{w'\left(1 - F(q|\underline{e})\right) f(q|\underline{e})}{w'\left(1 - F(q|\bar{e})\right) f(q|\bar{e})}\right),
\end{aligned} \tag{40}$$

where $t^{sb}_{nEU} \in \{t^{sb}_O, t^{sb}_P\}$ and $\mu_{nB} \in \{\mu_O, \mu_P\}$. For the EU agent, for whom $w' = 1$, (40) simplifies to

$$\frac{1}{u'(t^{sb}_{EU}(q))} = \nu + \mu_{EU} \left(1 - \frac{f(q|\underline{e})}{f(q|\bar{e})}\right). \tag{41}$$

Comparison of (40) and (41) shows that $t^{sb}_{nEU}(q) < t^{sb}_{EU}(q)$ only if $w'\Big(1 - F(q|\bar{e})\Big) < 1$, $w'\Big(1 - F(q|\bar{e})\Big) < w'\Big(1 - F(q|\underline{e})\Big)$, and $\mu_{EU} > \mu_{nEU}$.

For the optimist, these conditions hold for $q \in [\underline{q}, q^*)$, where q^* is the output level such that $w'\Big(1 - F(q^*|\bar{e})\Big) = 1$. For $q \in [q^*, \bar{q}]$, however, $w'(1 - F(q|\bar{e})) \geq 1$. Since $\lim_{q \to \bar{q}} \frac{f(q|\underline{e})}{f(q|\bar{e})} = 0$ and $\lim_{q \to \bar{q}} w'(1 - F(q|\bar{e})) = +\infty$, we have $t^{sb}_O(\bar{q}) > t^{sb}_{EU}(\bar{q})$ The inequality $t^{sb}_O(q) > t^{sb}_{EU}(q)$ can also hold, under certain conditions, for output levels lower than $\bar{q}$. Let $\mu_O \to \mu_{EU}$, which happens in case of moderate optimism. In that case, $t^{sb}_O(q) > t^{sb}_{EU}(q)$ holds for sufficiently small

$\frac{f(q|\underline{e})}{f(q|\bar{e})}$, or equivalently, high output levels. Note that moderate optimism is consistent with the assumption $\mu^O > 0$, that is, the incentive compatibility constraint holds (Proposition 3).

Similarly, (40) and (41) show that $t^{sb}_{nEU}(q) > t^{sb}_{EU}(q)$ only if $w'\big(1 - F(q|\bar{e})\big) > 1$, $w'\big(1 - F(q|\bar{e})\big) > w'\big(1 - F(q|\underline{e})\big)$, and $\mu_{nEU} \geq \mu_{EU}$.

For the pessimist, these conditions hold for $q \in [\underline{q}, q^*)$. For $q \in [q^*, \bar{q}]$, however, $w'\left(1 - F(q|\bar{e})\right) \leq 1$. Since $\lim_{q \to \bar{q}} \frac{f(q|\underline{e})}{f(q|\bar{e})} = 0$ and $\lim_{q \to \bar{q}} w'\big(1 - F(q|\bar{e})\big) = 0$, we have $t^{sb}_P(\bar{q}) < t^{sb}_{EU}(\bar{q})$.

The inequality $t^{sb}_P(q) < t^{sb}_{EU}(q)$ can also hold for lower output levels than $\bar{q}$ under some conditions: for sufficiently large $\frac{f(q|\underline{e})}{f(q|\bar{e})}$, that is, low enough output levels so $q \to q^*$; or for sufficiently small $w'\big(1 - F(q|\bar{e})\big)$, that is, strong pessimism. ■

Proposition 4

Proof. The considered problem is similar to the one solved in Proposition 1, the difference being that $w(p)$ now exhibits likelihood insensitivity and is thus inverse-S shaped. The maximization problem is otherwise unchanged, so the contract satisfying (6) is the candidate solution from the first-order condition. Denote that contract by $t^{fb}_L(q)$. We have that $\nu > 0$ since $u' > 0$ and $w' > 0$. Moreover, (7) shows that, for $q \in [\underline{q}, \tilde{q}]$ where $w''(q) > 0$, $t^{fb'}_L(q) < 0$; and for $q \in [\tilde{q}, \bar{q}]$ where $w''(p) < 0$, $t^{fb'}_L(q) > 0$.

We look at how $t^{fb}_L(q)$ behaves at extremes. That is as q either approaches $\underline{q}$ or $\bar{q}$. Lemma 1 implies $\lim_{q \to \underline{q}} \frac{w''(1-F(q|\bar{e}))}{w'(1-F(q|\bar{e}))} = +\infty$, so from (7) we get $\lim_{q \to \underline{q}} t^{fb'}_L(q) = +\infty$. Lemma 2 implies $\lim_{q \to \bar{q}} \frac{w''(1-F(q|\bar{e}))}{w'(1-F(q|\bar{e}))} = +\infty$, so $\lim_{q \to \bar{q}} t^{fb'}_L(q) = +\infty$. Moreover, note that $\lim_{q \to \tilde{q}} w''\big(1 - F(q|\bar{e})\big) = 0$, this is the inflection point from concavity to convexity, so it must be that $\lim_{q \to \tilde{q}} t^{fb'}_L(q) = 0$.

That $t^{fb'}_L(q) < 0$ for $q \in [\underline{q}, \tilde{q}]$ is undesirable, so we apply ironing. In this case, it consists on finding $q_I \in (\tilde{q}, \bar{q})$ such that:

$$\int_{\underline{q}}^{\tilde{q}} t^{fb}_L(q)\, \mathrm{d}q - \int_{\tilde{q}}^{q_I} t^{fb}_L(q)\, \mathrm{d}q = 0 \tag{42}$$

Symmetry of w around $\hat{p} = \tilde{p} = 0.5$ implies that $t_L^{fb}(q)$ is symmetric around $\tilde{q}$, the output level that generates probability $\tilde{p}$. Hence, there exists a $q_E \in (\tilde{q}, \bar{q})$ such that

$$t_L^{fb}(q_E) = \frac{\int\limits_{\underline{q}}^{\tilde{q}} t_L^{fb}(q)\,\mathrm{d}q}{\tilde{q} - \underline{q}}.$$

If we were to set $q_I = q_E$, we would be generating to non-monotonicities in the payment scheme. As, by construction, $\lim_{q \to \tilde{q}^+} t_L^{fb}(q) \neq t_L^{fb}(q_E)$. To fix that, note that continuity of $[\underline{q}, \bar{q}]$ and $t_L^{fb}(q)$, the latter in turn due to continuity of u', w', and w'', imply that for $\varepsilon > 0$ such that $\tilde{q} + \varepsilon \leq \bar{q}$, there exists a $\delta > 0$ such that $q_E - \delta \geq \tilde{q}$ and

$$t_L^{fb}(q_E - \delta) = \frac{\int\limits_{\underline{q}}^{\tilde{q}+\varepsilon} t_L^{fb}(q)\,\mathrm{d}q}{\tilde{q} + \varepsilon - \underline{q}}. \tag{43}$$

The tuple (δ, ε) can be adjusted to obtain $(\hat{\delta}, \hat{\varepsilon})$ such that both (43) holds and $q_E - \hat{\delta} = \tilde{q} + \hat{\varepsilon}$. Letting $q_I := \tilde{q} + \hat{\varepsilon}$, existence of $q_I \in (\tilde{q}, \bar{q})$ follows immediately.

The resulting ironed solution $\tilde{t}_L^{fb}(q)$ is thus

$$\tilde{t}_L^{fb}(q) = \begin{cases} t_L^{fb}(q_I) & \text{if } q < q_I, \\ t_L^{fb}(q) & \text{if } q \geq q_I. \end{cases}$$

For $q_I \in (\underline{q}, \bar{q})$.

■

Lemma 9. *A function ϕ that satisfies Assumption 4 and $\tilde{p} = 0.5$ exhibits subadditivity.*

Proof. Concavity of ϕ in $p < \tilde{p} = 0.5$ implies that

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

for $q < r$, $q < 0.5$, and $r + q < 0.5$, which satisfies the first condition in Definition 5.

Convexity of ϕ in $p > \tilde{p} = 0.5$ implies that

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

for $q < r$ and $q > 0.5$, which satisfies the second condition in Definition 5. ■

Lemma 10. *If agent i is more likelihood insensitive than agent j then:*

1. $-\frac{w_i''(p)}{w_i'(p)} > -\frac{w_j''(p)}{w_j'(p)}$ *if* $p < \tilde{p}$ *and* $\frac{w_i''(p)}{w_i'(p)} > \frac{w_j''(p)}{w_j'(p)}$ *if* $p > \tilde{p}$*;*
2. $w_i(p) > w_j(p)$ *if* $p < \tilde{p}$ *and* $w_i(p) < w_j(p)$ *if* $p > \tilde{p}$*;*
3. *There exists a unique* $p_k \in (0, \tilde{p})$ *such that* $w_i'(p_k) = w_j'(p_k)$*, this point becomes smaller the more likelihood insensitive i is with respect to j.*
4. *There exists a unique* $p_m \in (\tilde{p}, 1)$ *such that* $w_i'(p_m) = w_j'(p_m)$*, this point becomes larger the more likelihood insensitive i is with respect to j.*

Proof. **Part 1.** Consider first $p < \tilde{p}$. Since $w_i(p) = \phi\big(w_j(p)\big)$,

$$\frac{w_i''(p)}{w_i'(p)} = \frac{\phi''\big(w_j(p)\big)}{\phi'\big(w_j(p)\big)} w_j'(p) + \frac{w_j''(p)}{w_j'(p)}. \tag{44}$$

Due to $\phi'' < 0$ in $p < \tilde{p}$ (Lemma 9), it must be that

$$-\frac{w_i''(p)}{w_i'(p)} > -\frac{w_j''(p)}{w_j'(p)}.$$

A similar procedure gives that stronger likelihood insensitivity implies $\frac{w_i''(p)}{w_i'(p)} > \frac{w_j''(p)}{w_j'(p)}$ in $p > \tilde{p}$.

Part 2. Let $p_0, p_1 \in [0, 0.5]$ such that $p_1 > p_0$. Integrate $-\frac{w_i''(p)}{w_i'(p)} > -\frac{w_j''(p)}{w_j'(p)}$ over $[p_0, p_1]$ to obtain

$$\int_{p_0}^{p_1} \frac{-w_i''(s)}{w_i'(s)} \mathrm{d}s > \int_{p_0}^{p_1} \frac{-w_j''(s)}{w_j(s)} \mathrm{d}s \Leftrightarrow \ln\left(\frac{w_j'(p_1)}{w_j'(p_0)}\right) > \ln\left(\frac{w_i'(p_1)}{w_i'(p_0)}\right).$$

Integrating over the range of p_0, namely $[0, p_1]$, gives

$$\int_0^{p_1} w_j'(p_1) w_i'(s) \mathrm{d}s > \int_0^{p_1} w_i'(p_1) w_j'(s) \mathrm{d}s \Leftrightarrow w_j'(p_1) w_i(p_1) > w_i'(p_1) w_j(p_1).$$

Integrating again, but this time over the range of p_1 gives

$$\int_p^{0.5} \frac{w_j'(s)}{w_j(s)} \mathrm{d}s > \int_p^{0.5} \frac{w_i'(s)}{w_i(s)} \mathrm{d}s \Leftrightarrow w_i(p) > w_j(p).$$

A similar procedure gives that when i is more likelihood insensitive than j in the sense of Definition 6, then $w_i(p) < w_j(p)$ in $p > \hat{p}$.

Part 3.

Suppose that $w_i'(p) > w_j'(p)$ for all $p < \tilde{p}$. While $\int_0^{p_1} w_i'(p)\mathrm{d}p > \int_0^{p_1} w_i'(p)\mathrm{d}p \Leftrightarrow w_i(p_1) > w_j(p_1)$ for arbitrary $p_1 \in (0, \tilde{p})$, corroborating the first part of the Lemma. We also have that $\int_{p_1}^{0.5} w_i'(p)\mathrm{d}p > \int_{p_1}^{0.5} w_i'(p)\mathrm{d}p \Leftrightarrow w_i(p_1) < w_j(p_1)$, contradicting the first part of the Lemma. A similar rationale leads to a contradiction when $w_i'(p) < w_j'(p)$ for all $p < \tilde{p}$ is assumed. Hence, it must be that $w_i'(p) > w_j'(p)$ for some for some $p < \tilde{p}$.

Assumption 3 states that $w'(p)$ is decreasing in $p < \tilde{p}$. Moreover, Lemma 1 shows that $\lim_{p \to 0} w'(p) = +\infty$. Let $w_J(p) := \eta\big(w_j(p)\big)$ where η is a subadditive function. Accordingly, $-\frac{w_J''(p)}{w_J'(p)} > -\frac{w_j''(p)}{w_j'(p)}$ in $p < \tilde{p}$ as shown in the first part of this Lemma. The function $w_J'(p)$ tends to infinity faster than $w_j'(p)$ as $p \to 0^+$ as it exhibits a larger slope in $p < \tilde{p}$.

Due to the continuity of $w'(p)$, $w'(p)$ being decreasing in p in $p < \tilde{p}$, $-\frac{w_J''(p)}{w_J'(p)} > -\frac{w_j''(p)}{w_j'(p)}$, $w_J'(p)$ tending to infinity faster than $w_j'(p)$ as $p \to 0^+$ and the fact that $\lim_{p \to \tilde{p}} w'(p) = \min\{w'(p)\}$ for generic $w(p)$, there exists a unique point $p_k \in (0, \tilde{p})$ such that $w_J'(p_k) = w_j'(p_k)$. For $p < p_k$ then $w_J'(p) > w_j'(p)$ but instead $w_J'(p) < w_j'(p)$ if $\tilde{p} > p > p_k$.

Next, let $w_i := \phi\big(w_J(p)\big)$ where ϕ is a subadditive function. Then, $w_i'(p)$ is a decreasing function in $p < \tilde{p}$, exhibits $-\frac{w_i''(p)}{w_i'(p)} > -\frac{w_J''(p)}{w_J'(p)}$ for all $p < \tilde{p}$, tends to infinity faster than $w_J'(p)$ as $p \to 0^+$, and converges to $\min\{w'(p)\}$ as $p \to \tilde{p}$. Therefore, the point p_l such that $w_i'(p_l) = w_J'(p_l)$ is such that $p_l < p_k < \tilde{p}$.

Part 4. Suppose that $w_i'(p) < w_j'(p)$ for $p > \tilde{p}$. While $\int_{p_1}^{1} w_i'(p)\mathrm{d}p < \int_{\tilde{p}}^{1} w_i'(p)\mathrm{d}p \Leftrightarrow w_i(p_1) < w_j(p_1)$ for arbitrary $p_1 \in (\tilde{p}, \bar{p})$, corroborating the first part of the Lemma. We also have $\int_{\tilde{p}}^{p_1} w_i'(p)\mathrm{d}p > \int_{\tilde{p}}^{p_1} w_i'(p)\mathrm{d}p \Leftrightarrow w_i(p_1) < w_j(p_1)$, contradicting it. A similar rationale leads to a contradiction when

$w_i'(p) > w_j'(p)$ is assumed. Hence, it must be that if $w_i'(p) < w_j'(p)$ holds it must do so for some $p > \tilde{p}$.

Assumption 3 states that $w'(p)$ is increasing in $p > \tilde{p}$. Moreover, Lemma 2 shows that $\lim_{p\to 1} w'(p) = +\infty$. Let $w_J(p) := \eta\big(w_j(p)\big)$ where η is a subadditive function. Accordingly, $\frac{w_J''(p)}{w_J'(p)} > \frac{w_j''(p)}{w_j'(p)}$ in $p > \tilde{p}$ as shown in the first part of this Lemma. Therefore, $w_j'(p)$ tends to infinity faster than $w_J(p)$ as $p \to 1^-$.

Due to the continuity of $w'(p)$, $w'(p)$ being increasing in p in $p > \tilde{p}$, $\frac{w_J''(p)}{w_J'(p)} > \frac{w_j''(p)}{w_j'(p)}$ in $p > \tilde{p}$, $\lim_{p\to\tilde{p}} w'(p) = \min\{w'(p)\}$, and the fact that $w_j'(p)$ tends to infinity faster than $w_J(p)$ as $p \to 1^-$, there exists a unique point $p_m \in (0,1)$ such that $w_J'(p_m) = w_j'(p_m)$. For $p < p_m$ then $w_J'(p) < w_j'(p)$ but instead $w_J'(p) > w_j'(p)$ if $p > p_m > \tilde{p}$.

Next, Let $w_i := \phi\big(w_J(p)\big)$ where ϕ is a subadditive function. Then, w_i' is an increasing function in $p > \tilde{p}$, exhibits $\frac{w_i''(p)}{w_i'(p)} > \frac{w_J''(p)}{w_J'(p)}$ for all $p > \tilde{p}$, $w_J'(p)$ tends to infinity faster than $w_i'(p)$ as $p \to 1^+$, and converges to $\min\{w'(p)\}$ as $p \to \tilde{p}$. Hence, the point p_n such that $w_i'(p_n) = w_J'(p_n)$ is such that $p_n > p_m > \tilde{p}$.

■

Corollary 6

Proof. Denote by w_i and w_j the probability weighting functions of agent i and j, respectively. Using integration by parts, we express the utility of agent j over $q \in [\underline{q}, \tilde{q}]$ as:

$$u\Big(t_{L,j}^{fb}(\underline{q})\Big) - u\Big(t_{L,j}^{fb}(\tilde{q})\Big) w_j\Big(1 - F(\tilde{q}|\bar{e})\Big) + \int_{\underline{q}}^{\tilde{q}} u'(t_{L,j}^{fb}(q)) \frac{\mathrm{d}t_L^{fb}(q)}{\mathrm{d}q} w_j\Big(1 - F(q|\bar{e})\Big) \mathrm{d}q. \tag{45}$$

If agent i obtained $t_{L,j}^{fb}(q)$, he would derive utility:

$$u\Big(t_{L,j}^{fb}(\underline{q})\Big) - u\Big(t_{L,j}^{fb}(\tilde{q})\Big) w_i\Big(1 - F(\tilde{q}|\bar{e})\Big) + \int_{\underline{q}}^{\tilde{q}} u'(t_{L,j}^{fb}(q)) \frac{\mathrm{d}t_L^{fb}(q)}{\mathrm{d}q} w_i\Big(1 - F(q|\bar{e})\Big) \mathrm{d}q. \tag{46}$$

Subtracting (45) from (46) gives:

$$-u\Big(t_{L,j}^{fb}(\tilde{q})\Big)\Big(w_i\Big(1-F(q|\bar{e})\Big)-w_j\Big(1-F(q|\bar{e})\Big)\Big)+ \int_{\underline{q}}^{\tilde{q}} u'(t_{L,j}^{fb}(q))\frac{\mathrm{d}t_{L,j}^{fb}}{\mathrm{d}q}\Big(w_i\Big(1-F(q|\bar{e})\Big)-w_j\Big(1-F(q|\bar{e})\Big)\Big)\mathrm{d}q. \tag{47}$$

Since $\frac{\mathrm{d}t_{L,j}^{fb}}{\mathrm{d}q} < 0$ if $q < \tilde{q}$ and because stronger likelihood insensitivity implies $w_j\Big(1-F(q|\bar{e})\Big) > w_i\Big(1-F(q|\bar{e})\Big)$ in $q < \tilde{q}$ (Lemma 10), the expression in (47) is positive.

Symmetry of $t_{L,j}^{fb}$ around $\tilde{q}$, in turn generated by symmetry of $w'(p)$ around $\hat{p} = 0.5$, together with $t_{L,j}^{fb}$ making the participation constraint bind at the optimum for j entail $\int_{\underline{q}}^{\tilde{q}} u(t_{L,j}^{fb}(q))w_j'\Big(1-F(q|\bar{e})\Big)f(q|e)\mathrm{d}q = \frac{\bar{U}}{2}$.

Since (47) is positive, then

$$\int_{\underline{q}}^{\tilde{q}} u(t_{L,j}^{fb})w_i'\Big(1-F(q|\bar{e})\Big)f(q|e)\mathrm{d}q > \frac{\bar{U}}{2}. \tag{48}$$

Since $t_{L,i}^{fb}$ makes the participation constraint bind at the optimum for i, then $\int_{\underline{q}}^{\tilde{q}} u(t_{L,i}^{fb}(q))w_i'\Big(1-F(q|\bar{e})\Big)f(q|e)\mathrm{d}q = \frac{\bar{U}}{2}$. Therefore,

$$\int_{\underline{q}}^{\tilde{q}} u(t_{L,j}^{fb}(q))w_i'\Big(1-F(q|\bar{e})\Big)f(q|e)\mathrm{d}q > \int_{\underline{q}}^{\tilde{q}} u(t_{L,i}^{fb}(q))w_i'\Big(1-F(q|\bar{e})\Big)f(q|e)\mathrm{d}q. \tag{49}$$

Consequently, $t_{L,j}^{fb} > t_{L,i}^{fb}$ in $q \in [\underline{q}, \tilde{q}]$. Ironing $t_{L,j}^{fb}$ and $t_{L,i}^{fb}$ in $q \in [\underline{q}, \tilde{q}]$ gives $\tilde{t}_{P,j}^{fb} > \tilde{t}_{P,i}^{fb}$.

Consider now output levels $q > \tilde{q}$. The analog of Eq. (47) for that interval is

$$u(t^{fb}_{L,j}(\tilde{q}))\Big(w_i\Big(1-F(q|\bar{e})\Big)-w_j\Big(1-F(q|\bar{e})\Big)\Big)+$$
$$\int_{\tilde{q}}^{\bar{q}} u'(t^{fb}_{L,j}(q))\frac{\mathrm{d}t^{fb}_{L,j}}{\mathrm{d}q}\Big(w_i\Big(1-F(q|\bar{e})\Big)-w_j\Big(1-F(q|\bar{e})\Big)\Big)\mathrm{d}q. \tag{50}$$

Due to $\frac{\mathrm{d}t^{fb}_{L,j}}{\mathrm{d}q}>0$ if $q>\tilde{q}$ and because stronger likelihood insensitivity implies $w_j\Big(1-F(q|\bar{e})\Big)<w_i\Big(1-F(q|\bar{e})\Big)$ (Lemma 10) in $q>\tilde{q}$, The expression in (50) is positive.

Because $t^{fb}_{L,j}$ makes the participation constraint bind at the optimum for j, then $\int_{\underline{q}}^{\tilde{q}} u(t^{fb}_{L,j}(q))w'_j\Big(1-F(q|\bar{e})\Big)f(q|e)\mathrm{d}q=\frac{\bar{U}}{2}$. Moreover, (50) implies

$$\int_{\tilde{q}}^{\bar{q}} u(t^{fb}_{L,j})w'_i\Big(1-F(q|\bar{e})\Big)f(q|e)\mathrm{d}q>\frac{\bar{U}}{2}, \tag{51}$$

Given that $t^{fb}_{L,i}(q)$ makes the participation constraint bind at the optimum, then $\int_{\tilde{q}}^{\bar{q}} u(t^{fb}_{L,i}(q))w'_i\Big(1-F(q|\bar{e})\Big)f(q|e)\mathrm{d}q=\frac{\bar{U}}{2}$. Thus,

$$\int_{\tilde{q}}^{\bar{q}} u(t^{fb}_{L,j}(q))w'_i\Big(1-F(q|\bar{e})\Big)f(q|e)\mathrm{d}q>\int_{\tilde{q}}^{\bar{q}} u(t^{fb}_{L,i}(q))w'_i\Big(1-F(q|\bar{e})\Big)f(q|e)\mathrm{d}q. \tag{52}$$

Thus it must be that, $t^{fb}_{L,j}>t^{fb}_{L,i}$ in $q\in[\tilde{q},\bar{q}]$.

Since ironing the first-order condition in $q\in[\underline{q},\tilde{q}]$ generates $\tilde{t}^{fb}_{L,j}>\tilde{t}^{fb}_{L,i}$ and because $t^{fb}_{L,j}>t^{fb}_{L,i}$ in $q\in[\tilde{q},\bar{q}]$, the point q_I, at which the ironed solution meets the solution from the first-order condition, takes place at a larger output value for i than for j. ■

Proposition 5

Proof. The problem is similar to the one solved in Proposition 3 with the difference that $w(p)$ is now inverse-S shaped. Therefore, the first-order condition (27) solves the maximization problem. Denote by $t^{sb}_L(q)$ the contract that satisfies that equation.

Incentive constraint is binding. We first show that $\mu > 0$ might not be true at the optimum. Suppose that $\mu = 0$. Then $t_L^{sb}(q) = t_L^{fb}(q)$, $t_L^{fb}(q)$ being the first-best contract presented in Proposition 5.

From the complementary slackness condition of $\mu = 0$ we get

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_L^{fb}(q)\Big) w'(1 - F(q|\bar{e})) f(q|\bar{e}) \, \mathrm{d}q - c$$

$$> \int_{\underline{q}}^{\bar{q}} u\Big(t_L^{fb}(q)\Big) w'(1 - F(q|\underline{e})) f(q|\underline{e}) \, \mathrm{d}q. \quad (53)$$

Using integration by parts we rewrite (53) as

$$\int_{\underline{q}}^{\bar{q}} u'\Big(t_L^{fb}(q)\Big) \frac{\mathrm{d}t_L^{fb}(q)}{\mathrm{d}q} \Big[w(1 - F(q|\bar{e})) - w(1 - F(q|\underline{e}))\Big] \, \mathrm{d}q > c. \quad (54)$$

Assumption 4 implies $w\left(1 - F(q|\bar{e})\right) - w\left(1 - F(q|\underline{e})\right) \geq 0$ which, together with $\frac{\mathrm{d}t_O^{fb}(q)}{\mathrm{d}q} > 0$ in $q > \tilde{q}$ (Proposition 5) and $u'(t) > 0$ (Assumption 1), imply that the left-hand side of (29) can be weakly positive. Because $w(p)$ and $u(t)$ are $\mathcal{C}^2$, and since c is a constant unbounded from above, there exists $\hat{c}_L > 0$ such that, for a given $t_L^{fb}(q)$,

- if $c \leq \hat{c}_L$, (54) holds: $\mu = 0$ and $t_L^{sb}(q) = t_L^{fb}(q)$; on the other hand,
- if $c > \hat{c}_L$, (54) does not hold: $\mu > 0$ and $t_L^{sb}(q)$ satisfies (27).

In the remainder of the proof we assume that $c > \hat{c}_L$ so $\mu > 0$.

Shape of $t^{sb}(q)$ The second part of the proof analyzes the shape of $t^{sb}(q)$. To that end use (30), which presents the derivative of $t^{sb}(q)$ with respect to q. Denote by $\tilde{q} \in (\underline{q}, \bar{q})$ the output level satisfying $w\Big(1 - F(\tilde{q}|\bar{e})\Big) = 1 - F(\tilde{q}|\bar{e})$. Under likelihood insensitivity (Definition 4), $w'' < 0$ for all $q \in (\tilde{q}, \bar{q}]$. Hence, the two terms on the right-hand side of (30) are positive and $\frac{\mathrm{d}t_L^{sb}(q)}{\mathrm{d}q} > 0$.

To further understand the shape of $t_O^{sb}(q)$ in $q \in (\tilde{q}, \bar{q}]$ we study its behavior as q approaches the extremes of that set. From Definition 4 and Lemma 1 we know that $\lim_{q \to \bar{q}} \frac{w''\left(1-F(q|\bar{e})\right)}{w'\left(1-F(q|\bar{e})\right)} = -\infty$, so, according to (30), $\lim_{q \to \bar{q}} \frac{\mathrm{d}t_L^{sb}(q)}{\mathrm{d}q} =$

$+\infty$. Moreover, since $w''\big(1-F(\tilde{q}|\bar{e})\big)=0$ then $\lim_{q\to\tilde{q}^+}\frac{\mathrm{d}t_L^{sb}(q)}{\mathrm{d}q}>0$ due to Assumption 4.

Under likelihood insensitivity (Definition 4), $w''>0$ for all $q\in[\underline{q},\tilde{q}]$. Hence, the first term on the right-hand side of (30) is negative while the second term there is positive. The sign of $\frac{\mathrm{d}t_L^{sb}(q)}{\mathrm{d}q}$ depends on which of these terms dominates the other, which in turn depends on the magnitude of $w'(1-F(q|\bar{e}))$. When q decreases in the considered set, $w'(1-F(q|\bar{e}))$ increases, which in turn makes the second term in (30) larger, and $\frac{\mathrm{d}t_P^{sb}(q)}{\mathrm{d}q}$ is more likely to be positive. The opposite happens when q increases in $[\underline{q},\tilde{q}]$.

Again, we study (30) at the extremes of the considered set $q\in(\tilde{q},\bar{q}]$. Given that $w''\big(1-F(\tilde{q}|\bar{e})\big)=0$ then $\lim_{q\to\tilde{q}^-}\frac{\mathrm{d}t_L^{sb}(q)}{\mathrm{d}q}>0$ due to Assumption 4. It remains to be shown whether t^{sb} ever decreases with output; that is whether $\frac{\mathrm{d}t_L^{sb}(q)}{\mathrm{d}q}<0$ in $[\underline{q},\tilde{q}]$.

Lemma 10 shows that for agent i, who is more likelihood insensitive than agent j, p_l inducing $w_i'(p_l)=w_j'(p_l)$ in $p<\tilde{p}$ becomes smaller as the degree of insensitivity of i becomes more severe with respect to that of j. That Lemma also shows that p_n such that $w_i'(p_n)=w_j'(p_n)$ in $p>\tilde{p}$ becomes larger as i becomes more insensitive with respect to j. Hence,

$q_1^{\varepsilon}\in[\underline{q},\tilde{q}]$ and $q_2^{\varepsilon}\in[\tilde{q},\bar{q}]$ guaranteeing $w_i'\big(1-F(q_1^{\varepsilon}|e)\big)=w_i'\big(1-F(q_2^{\varepsilon}|e)\big)=\varepsilon$ for arbitrary small $\varepsilon>0$, take place at lower and higher output levels, respectively, as i becomes more insensitive with respect to j. Stronger likelihood insensitivity expands the segment $q\in[q_1^{\varepsilon},q_2^{\varepsilon}]$ in which $w_i'\big(1-F(q|e)\big)<\varepsilon$ becomes larger. This property makes (32) larger, which in turn implies that $\frac{\mathrm{d}t_L^{sb}(q)}{\mathrm{d}q}<0$ becomes more likely for intermediate values in $q\in(\underline{q},\tilde{q}]$.

Ironing That $\frac{\mathrm{d}t_L^{sb}(q)}{\mathrm{d}q}<0$ is possible in $q\in[\underline{q},\tilde{q}]$ for a sufficiently strong level of likelihood insensitivity is undesirable. We iron the solution to avoid this problem. Denote by $q_{\mathcal{M}}:=max\left(t_L^{sb}(q)\right)$ in $q\in[\underline{q},\tilde{q})$ and $q_{\mathcal{S}}:=min\left(t_L^{sb}(q)\right)$ in $q\in(q_{\mathcal{M}},\tilde{q})$. Ironing requires finding $q_{\mathcal{I}1}$ in $\in[\underline{q},q_{\mathcal{M}})$ and $q_{\mathcal{I}2}\in(q_{\mathcal{S}},\bar{q}]$ such that

$$\int_{q_{\mathcal{I}1}}^{q_{\mathcal{M}}} t_L^{sb}(q)\mathrm{d}q - \int_{q_{\mathcal{M}}}^{q_{\mathcal{S}}} t_L^{sb}(q)\mathrm{d}q + \int_{q_{\mathcal{S}}}^{q_{\mathcal{I}2}} t_L^{sb}(q)\mathrm{d}q = 0, \tag{55}$$

and

$$t_L^{sb}(q_{\mathcal{I}1}) = t_L^{sb}(q_{\mathcal{I}2}). \tag{56}$$

There are four possible cases. First, if $min\{t_L^{sb}(q)\} = t_L^{sb}(\underline{q})$ and $max\{t_L^{sb}(q)\} = t_L^{sb}(\bar{q})$, then $q_{\mathcal{I}1}$ in $\in (\underline{q}, q_{\mathcal{M}})$ and $q_{\mathcal{I}2} \in (q_{\mathcal{S}}, \bar{q})$ exist. The ironed incentives scheme becomes:

$$\tilde{t}_L^{sb}(q) = \begin{cases} t_L^{sb}(q) & \text{if } q \in [\underline{q}, q_{\mathcal{I}1}) \cup (q_{\mathcal{I}2}, \bar{q}], \\ t_L^{sb}(q_{\mathcal{I}1}) & \text{if } q \in [q_{\mathcal{I}1}, q_{\mathcal{I}2}]. \end{cases} \tag{57}$$

Second, if $min\{t_L^{sb}(q)\} = t_L^{sb}(q_{\mathcal{S}})$ and $max\{t_L^{sb}(q)\} = t_L^{sb}(\bar{q})$, then $q_{\mathcal{I}1} = \underline{q}$ and $q_{\mathcal{I}2} \in (q_{\mathcal{S}}, \bar{q})$. The resulting incentives scheme becomes:

$$\tilde{t}_L^{sb}(q) = \begin{cases} t_L^{sb}(q) & \text{if } q \in [q_{\mathcal{I}2}, \bar{q}], \\ t_L^{sb}(q_{\mathcal{I}1}) & \text{if } q \in [\underline{q}, q_{\mathcal{I}2}]. \end{cases} \tag{58}$$

Third, if $min\{t_L^{sb}(q)\} = t_L^{sb}(\underline{q})$ and $max\{t_L^{sb}(q)\} = t_L^{sb}(q_{\mathcal{M}})$, then $q_{\mathcal{I}1} \in (\underline{q}, q_{\mathcal{M}})$ and $q_{\mathcal{I}2} = \bar{q}$. The resulting incentives scheme in such case is:

$$\tilde{t}_{SB}^{LI}(q) = \begin{cases} t_L^{sb}(q_{\mathcal{I}2}) & \text{if } q \in [q_{\mathcal{I}2}, \bar{q}], \\ t_L^{sb}(q) & \text{if } q \in [\underline{q}, q_{\mathcal{I}2}]. \end{cases} \tag{59}$$

Finally, if $min\{t_L^{sb}(q)\} = t_L^{sb}(q_{\mathcal{S}})$ and $max\{t_L^{sb}(q)\} = t_L^{sb}(q_{\mathcal{M}})$, $q_{\mathcal{I}l}$ and $q_{\mathcal{I}2}$ do not exist, because $t(q)_{SB}^{LI}$ exhibits a sizable interval in which $\frac{\mathrm{d}t_L^{sb}(q)}{\mathrm{d}q} < 0$, the first-order solution cannot be ironed and incentive compatibility cannot be implemented without including perverse incentives. ■

Corollary 7

Proof. Let agent i be more likelihood insensitive than agent j. Lemma 10 shows that $p_l \in (0, \tilde{p})$ guaranteeing $w_i'(p_l) = w_j'(p_l)$ becomes smaller and that $p_m \in (\tilde{p}, 1)$ guaranteeing $w'_i(p_n) = w'_j(p_n)$ becomes larger as i becomes more insensitive with respect to j. Consequently, $q_1^\varepsilon \in [\underline{q}, \tilde{q}]$ and $q_2^\varepsilon \in [\tilde{q}, \bar{q}]$ guaranteeing $w'_i\Big(1 - F(q_1^\varepsilon|e)\Big) = w'_i\Big(1 - F(q_2^\varepsilon|e)\Big) = \varepsilon$ for arbitrary small $\varepsilon > 0$, take place at lower and higher output levels, respectively, as i becomes more insensitive. Stronger likelihood insensitivity expands the segment $q \in [q_1^\varepsilon, q_2^\varepsilon]$ in which $w'_i\Big(1 - F(q|e)\Big) < \varepsilon$ becomes larger. Accordingly, the expression in Eq. (32) becomes larger as agent i becomes more insensitive, which in turn yields that $\frac{\mathrm{d}t_L^{sb}(q)}{\mathrm{d}q} < 0$ is more likely for him than for j. Ironing is applied for a larger interval.

■

Corollary 8

Proof. According to (37), the value of $\nu > 0$ is the same across agents with different probability weighting functions. Moreover, (39) shows that the Lagrangian multiplier of the incentive compatibility constraint, $\mu^L{>}0$, exhibits $\mu^L > \mu^{EU}$ in the segments $q > q^{**}$ and $q < q^*$ and $\mu^L < \mu^{EU}$ otherwise.

The first-order condition presented in equation (27) is rewritten to obtain:

$$\begin{aligned}\frac{1}{u'(t_L^{sb}(q))} =& \nu w'\Big(1 - F(q|\bar{e})\Big) + \\ & \mu^L w'\Big(1 - F(q|\bar{e})\Big)\left(1 - \frac{w'\left(1 - F(q|\underline{e})\right) f(q|\underline{e})}{w'\left(1 - F(q|\bar{e})\right) f(q|\bar{e})}\right),\end{aligned} \tag{60}$$

where t_L^{sb} the second-best contract when the incentive compatibility constraint binds at the optimum (Proposition 6 (ii) and (iii)). For the EU agent, the optimal contract satisfies the first-condition presented in (41).

Comparison of (60) and (41) gives that $t_L^{sb}(q) < t_{EU}^{sb}(q)$ only if $w'\Big(1 - F(q|\bar{e})\Big) < 1$, $w'\Big(1 - F(q|\bar{e})\Big) < w'\Big(1 - F(q|\underline{e})\Big)$, and $\mu^{EU} > \mu^L$. These conditions hold in the interval $q \in [\tilde{q}, q^{**}]$.

Instead, in $q \in [q^*, \tilde{q}]$ the inequality $t_L^{sb}(q) < t_{EU}^{sb}(q)$ cannot hold. To see why let q approach q^*. Since $\lim_{q \to q^*} w'\Big(1 - F(q|\bar{e})\Big) = 1$, the expression in (60) becomes similar to that in (41) with the important exception that $\mu^L > \mu^{EU}$, which leads to $t_L^{sb}(q^*) > t_{EU}^{sb}(q^*)$. For higher output levels in $q \in (q^*, \tilde{q}]$, the relation $t_L^{sb}(q^*) > t_{EU}^{sb}(q^*)$ can hold whenever $w'\Big(1-F(q|\bar{e})\Big)$ is sufficiently large. This happens with sufficiently moderate levels of likelihood insensitivity. In contrast, sufficiently high likelihood insensitivity yields $w'\Big(1-F(q|\bar{e})\Big)$ small at the considered region(Lemma 10) and lead to $w'\Big(1-F(q|\bar{e})\Big) > w'\Big(1-F(q|\underline{e})\Big)$, both of which might outweigh $\mu^{EU} < \mu^L$.

Comparison of (60) and (41) gives that $t_L^{sb}(q) > t_{EU}^{sb}(q)$ only if $w'\Big(1 - F(q|\bar{e})\Big) > 1$, $w'\Big(1 - F(q|\underline{e})\Big) < w'\Big(1 - F(q|\bar{e})\Big)$, and $\mu^L \geq \mu^{EU}$. These conditions hold in the interval $q \in [\underline{q}, q^*)$.

The validity of the relation $t_L^{sb}(q) > t_{EU}^{sb}(q)$ in $q \in [q^{**}, \bar{q}]$ is first evaluated as q approaches $\bar{q}$. Since $\lim_{q \to \bar{q}} \frac{f(q|\underline{e})}{f(q|\bar{e})} = 0$ and $\lim_{q \to \bar{q}} w'\Big(1 - F(q|\bar{e}\Big) = +\infty$, equations (60) and (41) imply that $t_L^{sb}(\bar{q}) > t_{EU}^{sb}(\bar{q})$. To guarantee $t_L^{sb}(q) > t_{EU}^{sb}(q)$ for lower output levels, $w'\Big(1 - F(q|\bar{e})\Big)$ has to be sufficiently large to outweigh $w'\Big(1-F(q|\bar{e})\Big) < w'\Big(1-F(q|\underline{e})\Big)$ and $\mu^{EU} > \mu^L$. A possibility under sufficiently large levels of likelihood insensitivity.

■

Appendix B: Prospect Theory Preferences (Online)

In this Appendix, we extend the model to account for reference dependence. To that end, we enrich the agent's risk attitudes by characterizing them according to Cumulative Prospect Theory (CPT from here onward, Tversky and Kahneman, 1992). Accordingly, the agent does not evaluate the transfers in $t(q)$ as final carriers of wealth, but does so relative to a reference point $r > 0$.

For simplicity, we assume that the reference point r is assumed to be exogenous to the alternatives faced by the decision-maker. For instance, it can be the agent's current wealth at the moment of making decisions (Kahneman and Tversky, 1979; Tversky and Kahneman, 1981). This reference point rule has been recently validated empirically by Baillon et al. (2020a) as it explains most of subjects' behavior.

As stated in the main text, the main departure of CPT with respect to RDU and EUT is that the agent can exhibit different risk preferences for gains and losses. This is captured with two ingredients. First, transfers enter the agent's utility differently depending on whether they are classified as gains or losses. A property that is captured by the following assumption on the agent's utility.

Assumption 5. *The value function, $V(t,r)$, is a piece-wise function,*

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

with the following properties:

- $\lambda > 1$*;*
- $v(0) = 0$*;*
- $v' > 0$ *for all* $q \in [\underline{q}, \bar{q}]$*;*
- $v'' < 0$ *for all* $q \in [\underline{q}, \bar{q}]$.

The agent's utility is convex for losses, generating risk seeking attitudes, and concave for gains, generating risk aversion. Furthermore, Assumption 5 introduces loss aversion. That is, transfers counting as losses loom larger than equally-sized transfers counting as gains. This latter property is captured by the parameter $\lambda > 1$ and expresses a special dislike for losses.

The second ingredient is that the probability weighting function is defined separately over gains and losses. Probabilities associated with gains are transformed by the probability weighting function w, introduced in Assumption 3. 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.[19]

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

Assumption 6. *A probability weighting function for losses is a function* $z : [0,1] \to [0,1]$ *satisfying the duality condition* $z(F(q|e)) = 1 - w(1 - F(q|e))$ *for any* e.

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

$$CPT(t,e,r) = \int_{\underline{q}}^{\bar{q}} \Big[\theta v\left(t(q) - r\right) w'\Big(1 - F(q|e)\Big) \\ - \lambda(1-\theta) v(r - t(q)) z'\Big(F(q|e)\Big)\Big] f(q|e)\,\mathrm{d}q - c(e), \tag{61}$$

where θ is an indicator function taking the value $\theta = 1$ if $t(q) \geq r$ and $\theta = 0$ otherwise.

[19] In other words, the CPT agent orders possible transfers counting as losses from the least-desirable, $t(\underline{q})$, to the closest to the reference point from below, and uses a separate weighting function z to transform the probabilities that emerge from these—as the literature describes them—loss ranks.

The principal's program when facing a CPT agent is:

$$\begin{aligned}
\max_{t(q)} \quad & \int_{\underline{q}}^{\bar{q}} \Big(S(q) - t(q)\Big) f(q|\bar{e}) \,\mathrm{d}q \\
\text{s.t.} \quad & CPT(t, \bar{e}, r) \geq \bar{U}, \\
& CPT(t, \bar{e}, r) \geq CPT(t, \underline{e}, r)
\end{aligned}$$

The optimal incentive scheme offered to agents with CPT preferences is characterized next.

Proposition 7. *Let Assumptions 3 to 6 hold. There exists a threshold $\hat{q} \in [\underline{q}, \bar{q}]$ such that the second best-contract, t_C^{sb}:*

(i) pays r everywhere if $\hat{q} = \bar{q}$;

(ii) pays r in $q < \hat{q}$ and depends on performance as in Proposition 3, Proposition 4, or Proposition 6 in $q \geq \hat{q}$ if $\hat{q} \in (\underline{q}, \bar{q})$;

(iii) depends on performance as in Proposition 3, Proposition 4, or Proposition 6 if $\hat{q} = \underline{q}$.

Proof. Rewrite Eq. (61) using Assumption 6 as

$$\begin{aligned}
CPT(t, e, r) = \int_{\underline{q}}^{\bar{q}} \Big[& \theta v\left(t(q) - r\right) w'\Big(1 - F(q|e)\Big) \\
& - \lambda(1-\theta) v(r - t(q)) w'\Big(1 - F(q|e)\Big)\Big] f(q|e) \,\mathrm{d}q - c(e), \quad (62)
\end{aligned}$$

where θ is an indicator function taking a value one if $t \geq r$. Let first $\theta = 0$. Denoting by ν and μ the multipliers associated to the participation and the incentive compatibility constraints, respectively, the Lagrangian of the principal's problem can be written as

$$\begin{aligned}\mathcal{L}(q,t) =& \Big(S(q)-t(q)\Big)f(q|\bar{e}) \\ &+\mu\Bigg(-\lambda v\Big(r-t(q)\Big)\Big(w'\Big(1-F(q|\bar{e})\Big)f(q|\bar{e})-w'\Big(1-F(q|\underline{e})\Big)f(q|\underline{e})\Big)-c\Bigg) \\ &+\nu\Bigg(-\lambda v\Big(r-t(q)\Big)w'\Big(1-F(q|\bar{e})\Big)f(q|\bar{e})-c-\bar{U}\Bigg).\end{aligned} \tag{63}$$

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

$$\frac{1}{\lambda v'\Big(r-t\Big)\Big(w'\Big(1-F(q|\bar{e})\Big)\Big)} = \nu + \mu\left(1 - \frac{w'\Big(1-F(q|\underline{e})\Big)f(q|\underline{e})}{w'\Big(1-F(q|\bar{e})\Big)f(q|\bar{e})}\right). \tag{64}$$

Denote by $t_C^{sb}(q)$ the transfer satisfying Eq. (64). We show next that a lottery $L = (p, r; 1-p, 0)$ improves upon the solution $t_C^{sb}(q)$ whenever $0 < t_C^{sb}(q) < r$. Since $-\lambda v\Big(r - t_C^{sb}(q)\Big)$ is increasing in $t_C^{sb}(q)$, there exists a number $\rho \in [0,1]$ for each realization q such that

$$\lambda v\Big(r - t_C^{sb}(q)\Big) = \lambda\Big(1-w(\rho)\Big)v(r). \tag{65}$$

Hence $L_\rho := (\rho, r; 1-\rho, 0)$ leaves the agent's participation and incentive compatibility constraints unchanged. Using the fact that $v'' < 0$ gives

$$\lambda v\Big(r - t_C^{sb}(q)\Big) \leq \lambda v\Big((1-w(\rho))r\Big). \tag{66}$$

Since $v' > 0$ is increasing then $t_C^{sb}(q) > w(\rho)r$. The lottery contract L_ρ can be cost-efficient for the principal, it provides the same incentives at a lower perceived expected cost. Note that when $w(\rho) < \rho$ the lottery contract has a lower expected cost.

The incentives of offering L_ρ are studied next. Let $\bar{L} := \rho r$. The utility of an agent is

$$CPT(L_\rho, \bar{e}, r) = -\left(1 - w\left(\frac{\bar{L}}{r}\right)\right)\lambda v(r) - c \tag{67}$$

The above equation is not linear in $\bar{L}$ due to w having curvature (Assumption 3). Hence, changes in $\bar{L}$ affect marginal utility. To understand how changes in $\bar{L}$ affect the marginal incentives of offering the lottery, we compute the first-order condition of (67) with respect to ρ, which gives us

$$w'(\rho)\lambda v(r) = 0. \tag{68}$$

Denote by ρ^{opt} the probability satisfying the condition in (68). The second-order condition evaluated at ρ^{opt} is

$$w''(\rho^{opt})\lambda v(r). \tag{69}$$

Hence, $\rho^{opt} \in (0,1)$ whenever $w'' < 0$. This holds under optimism or likelihood insensitivity.

Due to Assumption 3, $\lim_{\rho\to 1} w'(\rho) = 0$ under optimism so in that case $\rho^{opt} \to 1$. Instead, $\rho^{opt} \in \{0,1\}$ if $w'' > 0$ for any interval in $p \in (0,1)$. Since

$$CPT(L_{\rho=1}, \bar{e}, r) = -c > -\lambda v(r) - c = CPT(L_{\rho=0}, \bar{e}, r), \tag{70}$$

then $\rho^{opt} = 1$ in that case. Therefore, either for optimism or whenever $w(p)$ is convex in any interval, the principal avoids exposing the agent to losses by paying $t = r$.

Let now $\theta = 1$. The Lagrangian of the principal's problem in that case can be written as

$$\begin{aligned}\mathcal{L}(q,t) =& \Big(S(q) - t(q)\Big)f(q|\bar{e}) \\ &+ \mu\Bigg(v\Big(t(q) - r\Big)\Big(w'\Big(1 - F(q|\bar{e})\Big)f(q|\bar{e}) - w'\Big(1 - F(q|\underline{e})\Big)f(q|\underline{e})\Big) - c\Bigg) \\ &+ \nu\Bigg(v\Big(t(q) - r\Big)w'\Big(1 - F(q|\bar{e})\Big)f(q|\bar{e}) - c - \bar{U}\Bigg).\end{aligned} \tag{71}$$

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

$$\frac{1}{v'\big(t-r\big)w'\big(1-F(q|\bar{e})\big)} = \nu + \mu\left(1 - \frac{w'\big(1-F(q|\underline{e})\big)f(q|\underline{e})}{w'\big(1-F(q|\bar{e})\big)f(q|\bar{e})}\right). \tag{72}$$

Since $v' > 0$ and $v'' < 0$ and $w(p)$ is as described by Assumption 3, the solution is similar to that presented in Proposition 3 and Proposition 6, except that it can be that $r > 0$. Hence, r is now taken as the initial value for those solutions.

To establish the location shift from paying the amount $t = r$, given to protect the agent from losses, to a solution that increases in performance, as given by Proposition 3 or Proposition 6), denote by $\hat{q} \in [\underline{q}, \bar{q}]$ the performance level satisfying:

$$\frac{1}{\frac{\lambda v(r)}{r}} = \nu + \mu\left(1 - \frac{w'\big(1-F(\hat{q}|\underline{e})\big)f(\hat{q}|\underline{e})}{w'\big(1-F(\hat{q}|\bar{e})\big)f(\hat{q}|\bar{e})}\right). \tag{73}$$

Where the left-hand side of (73) denote the marginal incentives of offering $L_{\rho=1}$. The existence and uniqueness of $\hat{q}$ is guaranteed by the fact that the left-hand side of Eq. (73) of is positive and constant in q while the right-hand side of that equation increases with q (Assumption 4) over $[0, +\infty)$.

There are three cases. When $\frac{\lambda v(r)}{r}$ is small and the right-hand side of (73) is large enough, then $\hat{q} \geq \bar{q}$. In that case $t_C^{sb} = r$. Alternatively, $\frac{\lambda v(r)}{r}$ can be large so that $\hat{q} \leq \bar{q}$ and the solution is fully given by Proposition 3 and Proposition 6, depending on the shape of w. Finally, if $\hat{q} \in [\underline{q}, \bar{q}]$ then

$$t_C^{sb}(q) = \begin{cases} r & \text{if } q < \hat{q}, \\ t_P^{sb}(q), t_O^{sb}(q) \text{ (Proposition 3), or } t_L^{sb}(q) \text{ (Proposition 6)} & \text{if } q \geq \hat{q}. \end{cases} \tag{74}$$

■

Under CPT preferences, the optimal contract often includes a performance-insensitive segment paying the amount r. The reason behind these segments is loss aversion. Exposing the agent to losses by paying amounts lower than r would generate large disutility, leading eventually to rejection. To prevent this, the principal can either introduce large rewards that compensate the agent for facing such risk of losses, or she can eradicate the possibility of losses. The former solution is expensive since losses loom larger than equally sized gains by a factor of λ. Consequently, the principal offers, wherever necessary, the minimum amount required to locate the agent in the domain of gains: $t(q) = r$. This payment is given unless the realization of output crosses a critical threshold $\hat{q}$.

Moreover, the optimal contract might as well include transfers that depend on performance in the same way as the contracts described in Propositions 3 or 6. Depending on the agent's probability perception in gains, the shape of one of these contracts applies for all $q > \hat{q}$. That is because in the domain of gains, the CPT agent exhibits risk attitudes equivalent to those of the RDU agent. So, the second-best contract that motivates an RDU agent, also suffices to incentivize a CPT agent with the same probability weighting function.

The contract characterized in Proposition 7, leads to incentive schemes that are often observed in practice. For instance, when the CPT agent is sufficiently pessimistic the resulting optimal contract can be binary. It pays a fixed salary, $t(q) = r$ in $q < \hat{q}$, and a lump-sum bonus, paid in in $q > \hat{q}$. This shape reflects different sources of risk aversion. The first fixed-pay level ensures that the agent does not face losses, while the second fixed-pay level reflects the impossibility faced by the principal to implement incentives due to the agent's severe pessimism. 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 the agent's risk attitudes are not characterized by CPT. Also, our result holds for any level of loss aversion, i.e. even if $\lambda > 2$.

Appendix C: Continuous Effort (Online)

Let $e \in [\underline{e}, \bar{e}]$ with $\underline{e} \geq 0$. The following assumptions are made on $c(e)$ the function capturing the cost of effort.

Assumption 7 (cost of effort). *$c(e) : [\underline{e}, \bar{e}] \to [0, +\infty)$ is $\mathcal{C}^2$ with $c'(e) > 0$ and $c''(e) > 0$.*

Furthermore, we impose the following assumptions on the cumulative distribution function.

Assumption 8 (output distribution). *$F(y|e) : [\underline{y}, \bar{y}] \to [0, 1]$ is $\mathcal{C}^2$ with respect to e and y, and exhibits $F_{ee}(y|e) > 0$.*

As in the main body of the paper, the probability density function is defined as $f(y|e) := F_q(q|e)$. Note that the convexity of the CDF, $F_{ee}(y|e) > 0$, has been shown to ensure the validity of the first-order approach.

Furthermore, we extend the continuous MLRP, $\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{f_e(q|e)}{f(q|e)}\right) > 0$, to account for probability distortions.

Assumption 9 (continuous WMLRP). $\frac{d}{dq}\left(\frac{\frac{d}{de}\Big(w'(1-F(q|e))f(q|e)\Big)}{w'(1-F(q|e)f(q|e))}\right) > 0$

A central implication of Assumption 9 is that it implies first-order stochastic dominance, $F_e(q|e) \leq 0$.

We are in a position to show that stronger conditions are required to guarantee the validity of the first-order approach under probability distortion.

Lemma 11. *For the first-order approach to be valid it suffices that $\frac{w_{ee}\Big(1-F(q|e)\Big)F_e(q|e)}{w_e\Big(1-F(q|e)\Big)} < \frac{F_{ee}(q|e)}{F_e(q|e)}$, or it is necessary and sufficient that $c_{ee}(e) > \mathcal{B}$, where*

$$\mathcal{B} := \int_{\bar{q}}^{\underline{q}} u'\Big(t(q)\Big)\frac{dt(q)}{dq}\left(w_e\Big(1-F(q|e)\Big)F_{ee}(q|e) - w_{ee}\Big(1-F(q|e)\Big)\Big(F_e(q|e)\Big)^2\right)dq.$$

Proof. Using integration by parts, rewrite the agent's utility in Eq. (3) as

$$RDU(t,e) = u\Big(t(\underline{q})\Big) - \int_{\bar{q}}^{\underline{q}} u'\Big(t(q)\Big)\frac{\mathrm{d}t(q)}{\mathrm{d}q}w\Big(1 - F(q|e)\Big)\mathrm{d}y - c(e). \qquad (75)$$

Denote by t^{fo} the solution to the following principal's program:

$$\begin{aligned}
\max_{\{t(y)\}} \quad & \int_{\underline{y}}^{\bar{y}} \Big(S(y) - t(y)\Big) f(y|e)\mathrm{d}y \\
\text{s.t.} \quad & u\Big(t(\underline{q})\Big) - \int_{\bar{q}}^{\underline{q}} u'\Big(t(q)\Big)\frac{\mathrm{d}t(q)}{\mathrm{d}y} w\Big(1 - F(q|e)\Big)\mathrm{d}q - c(e) \geq \bar{U}, \qquad (76)\\
& \int_{\bar{q}}^{\underline{q}} u'\Big(t(q)\Big)\frac{\mathrm{d}t(y)}{\mathrm{d}q} w_e\Big(1 - F(q|e)\Big)F_e(y|e)\mathrm{d}q - c'(e)
\end{aligned}$$

In the above program, the incentive compatibility constraint is replaced by the first-order condition of Eq. (75) with respect to e. This approach is necessary and sufficient if the following condition holds:

$$\int_{\bar{q}}^{\underline{q}} u'\Big(t(q)\Big)\frac{\mathrm{d}t(q)}{\mathrm{d}q}\bigg(w_e\Big(1-F(q|e)\Big)F_{ee}(q|e)-w_{ee}\Big(1-F(q|e)\Big)\Big(F_e(q|e)\Big)^2\bigg)\mathrm{d}q-c''(e) < 0. \tag{77}$$

Since $c''(e) > 0$ (Assumption 7), $u' > 0$ (Assumption 1), $\frac{\mathrm{d}t(q)}{\mathrm{d}q} \geq 0$ (Assumption 2), the following condition suffices for the concavity of $RDU(t, e)$:

$$w_e\Big(1 - F(q|e)\Big)F_{ee}(q|e) - w_{ee}\Big(1 - F(q|e)\Big)\Big(F_e(q|e)\Big)^2 < 0 \tag{78}$$

Due to $F_{ee}(q|e) > 0$ (Assumption 8) and $w_e\Big(1 - F(q|e)\Big) > 0$ (Assumption 3), a probability weighting function that exhibits $w_{ee}\Big(1 - F(y|e)\Big) < 0$ cannot fulfill the condition in Eq. (78). Hence, for the optimality of t^{fo} it suffices that $w_{ee}\Big(1 - F(q|e)\Big) > 0$. Letting $p = 1 - F(q|e)$, that condition can be written as $w''(p) > 0$.

■

The Lemma shows that the first-order condition suffices to characterize the incentive constraints when the weighting function is sufficiently convex, so as to guarantee $\frac{w_{ee}\left(1-F(q|e)\right)F_e(q|e)}{w_e\left(1-F(q|e)\right)} < \frac{F_{ee}(q|e)}{F_e(q|e)}$, or when the cost function is sufficiently convex. For simplicity we assume that when $w''(p)$ is not sufficiently convex, $c(e)$ attains the bound presented in the above Lemma. If that were not the case, the principal might require other means to incentivize the agent. Gonzalez-Jimenez (2020) stochastic contracts are optimal when this condition does not hold.

We are in a position to characterize the optimal contracts when effort is observable. It turns out that they are identical to those presented under the binary case.

Proposition 8. *The optimal first-best under optimism, pessimism, or likelihood insensitivity exhibit the shapes of the contracts presented in Proposition 1 and Proposition 5.*

Proof. Denoting the Lagrange multiplier of the agent's participation constraint by ν, the Lagrangian of the principal's problem writes as:

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

Pointwise optimization with respect to $t(q)$ and algebraic manipulations yield

$$\frac{1}{u'\Big(t^{fb}(q)\Big)w'\Big(1-F(q|e)\Big)}=\nu. \tag{79}$$

By assumption, $u'(t)>0$ and $w'(p)>0$, so $\nu>0$. The participation constraint binds at the optimum.

The optimal effort level, e^* satisfies

$$\begin{aligned}&\int_{\bar{q}}^{\bar{q}}\Big(S(q)-t^{fb}(q)\Big)f_e(q|e^*)\mathrm{d}q+\\ &\qquad\nu\left(-\int_{\underline{q}}^{\bar{q}}u'\Big(t^{fb}(q)\Big)\Big(w_e\Big(1-F(q|e^*)\Big)\Big)F_e(q|e^*)\mathrm{d}q-c'(e^*)\right)=0.\end{aligned} \tag{80}$$

Since $-\int_{\underline{q}}^{\bar{q}}u'\Big(t^{fb}(q)\Big)\Big(w_e\Big(1-F(q|e^*)\Big)\Big)F_e(q|e^*)\mathrm{d}q-c'(e^*)=0$, the above equation becomes:

$$\int_{\bar{q}}^{\bar{q}}\Big(S(q)-t^{fb}(q)\Big)f_e(q|e^*)\mathrm{d}q=0. \tag{81}$$

The solution of the principal's program is thus given by $\{(t^{fb}(q), e^*)\}$, where $t^{fb}(q)$ is the transfer satisfying Eq. (79) and e^* satisfies Eq. (81).

To investigate the shape of $t^{fb}(q)$ we differentiate (79) with respect to q, giving us

$$t^{fb\prime}(q) = \frac{u'\Big(t^{fb}(q)\Big)}{u''\Big(t^{fb}(q)\Big)} \frac{w''\Big(1 - F(q|e^*)\Big)}{w'\Big(1 - F(q|e^*)\Big)} f(q|e). \tag{82}$$

This is exactly the equality in (7) when letting $\bar{e} = e^*$. The analyses of the shape of $t^{fb}(q)$ under optimism, pessimism, and likelihood insensitivity in Propositions 5 and 1 immediately follow. ■

Consider now a setting of moral hazard. First, we show that when optimism or likelihood insensitivity are moderate, the first-best may suffice to elicit high effort levels. This solution is the analog of Proposition 3 if $c < \hat{c}_O$ and Proposition 6 if $c < \hat{c}_L$, but, as a direct consequence of considering a continuous output space, we condition on the values of e^{**}, the optimal effort level implemented by the principal, rather than on c.

Proposition 9. *Assume Optimism or Likelihood Insensitivity. There exists a unique effort level $\hat{e} \in [\underline{e}, \bar{e}]$ such that if e^{**}, the effort level implemented by the principal, is such that $e^{**} < \hat{e}$, the optimal contract is the first-best contract from Proposition 8.*

Proof. Denote by ν the Lagrange multiplier of the agent's participation constraint, and μ, of the incentive compatibility constraint. The Lagrangian of the principal's maximization problem writes as

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

Pointwise optimization with respect to $t(q)$ and algebraic manipulations yield

$$\frac{1}{u'\Big(t^{sb}(q)\Big)w'\Big(1-F(q|e)\Big)} = \nu + \mu\left(\frac{\frac{\mathrm{d}}{\mathrm{d}e}\Big(w'(1-F(q|e))f(q|e)\Big)}{w'(1-F(q|e)f(q|e))}\right) \tag{83}$$

The optimal transfer under moral hazard, $t^{sb}(q)$ results from the condition above.

The optimal effort level under moral hazard, e^{**}, must satisfy

$$\int_{\bar{q}}^{\bar{q}}\Big(S(q)-t^{fb}(q)\Big)f_e(q|e^{**})\mathrm{d}q+$$
$$\mu\left(-\int_{\underline{q}}^{\bar{q}}u'\Big(t^{fb}(q)\Big)\Big(w_e\Big(1-F(q|e^{**})\Big)F_{ee}(q|e^{**})-w_{ee}\Big(1-F(q|e^{**})\Big)F_e(q|e^{**})\Big)\mathrm{d}q-c''(e^{**})\right)=0. \tag{84}$$

The solution of the principal's program is thus given by $\{(t^{sb}(q), e^{**})\}$, where $t^{sb}(q)$ is the transfer satisfying Eq. (83) and e^{**} satisfies Eq. (84).

We next show that $\mu > 0$ might not hold the optimum under optimism or likelihood insensitivity and the solution to the principal's problem becomes $\{(t^{fb}(q), e^{**})\}$. Suppose instead that $\mu = 0$. Accordingly, $t^{sb}(q) = t^{fb}(q)$, where $t^{fb}(q)$ is the first-best contract presented in Proposition 8.

Optimism Consider the case of an agent with optimism in the sense of Definition 1. From the complementary slackness condition from μ we get

$$u'\Big(t(q)\Big)\frac{\mathrm{d}t(y)}{\mathrm{d}q}w_e\Big(1-F(q|e)\Big)F_e(y|e)\mathrm{d}q > c'(e) \tag{85}$$

Assumption 9 implies $F_e(q|e) < 0$ which, together with $\frac{\mathrm{d}t(q)}{\mathrm{d}q} > 0$ (Proposition 8), $w' > 0$ (Assumption 3) and $u'(t) > 0$ (Assumption 1), imply that the left-hand side of (85) is weakly positive, rendering the inequality in (Assumption 3) feasible.

The right-hand side of (85) is increasing because $c'(0) = 0$ and $c''(e) > 0$. Also, because $w_e\Big(1-F(q|e)\Big)F_e(q|e)$ is decreasing (Lemma 11) the left-hand side of (85) is decreasing. Hence, there exists an effort level $\hat{e} \in [\underline{e}, \bar{e}]$ such that

$$u'\Big(t(q)\Big)\frac{\mathrm{d}t(y)}{\mathrm{d}q}w_e\Big(1 - F(q|\hat{e})\Big)F_e(y|\hat{e})\mathrm{d}q = c'(\hat{e}).$$

Hence, for $e \in [\underline{e}, \hat{e})$, the inequality in (85) holds .

Likelihood insensitivity For likelihood insensitivity $\frac{\mathrm{d}t(q)}{\mathrm{d}q} > 0$ (Proposition 8) so the inequality in (Assumption 3) is feasible. Since, $w''(p) > 0$ in $(0, \tilde{p})$, then $w_e\Big(1-F(q|e)\Big)F_e(q|e)$ is decreasing (Lemma 11) in that probability interval, the existence of $\hat{e}$ is guaranteed.

■

Second, it is shown that the contract shapes presented in Proposition 3 and Proposition 6 continue to hold when effort is continuous.

Proposition 10. *The optimal second-best exhibits the shapes of the contracts presented in Proposition 3 and Proposition 6 under pessimism, or if $e^{**} > \hat{e}$ and under either likelihood insensitivity or optimism.*

Proof. Assume $\mu > 0$. Differentiate (83) with respect to q to obtain:

$$\begin{aligned} t^{sb\prime}(q) =& \frac{u'(t^{sb}(q))w''(1 - F(q|e^{**}))}{u''(t^{sb}(q))w'\left(1 - F(q|e^{**})\right)}f(q|e^{**}) \\ &+ \mu\frac{w'\left(1 - F(q|e^{**})\right)u'(t^{sb}(q))^2}{u''(t^{sb}(q))}\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{\frac{\mathrm{d}}{\mathrm{d}e}\Big(w'(1 - F(q|e))f(q|e)\Big)}{w'(1 - F(q|e)f(q|e))}\right). \end{aligned} \tag{86}$$

The above equation and Eq. (30) differ only in that $\bar{e}$ is now e^{**} and the discrete MLRP is replaced by its continuous analog. Therefore, the analysis of $t^{sb\prime}(q)$ is similar to that presented in Proposition 3.

Under optimism, $w''(p) < 0$ for all $p \in (0, 1)$ implies that both terms in Eq. (86) are positive, implying that t^{sb} is everywhere increasing. Moreover, since $\lim_{q\to\bar{q}} w'(p) = +\infty$ and $\lim_{q\to\underline{q}} w'(p) = 0$, then $t^{sb\prime}(q) \to +\infty$ at both extremes.

Under pessimism, $w''(p) > 0$ for all $p \in (0, 1)$. Hence, the first term in the right-hand side of Eq. (86) is negative, while the second one is positive.

Due to $\lim_{q\to\bar{q}} w'(p) = 0$, then $\lim_{q\to\bar{q}} \frac{w''(p)}{w'(p)} = +\infty$; the first term in Eq. (86) dominates and $\lim_{q\to\bar{q}} {t^{sb}}'(q) = -\infty$.

Eq. (86) implies that ${t^{sb}}'(q) > 0$ under pessimism requires:

$$-\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{\frac{\mathrm{d}}{\mathrm{d}e}\Big(w'(1-F(q|e))f(q|e)\Big)}{w'(1-F(q|e)f(q|e))}\right) > \frac{w''(1-F(q|e))f(q|e)}{w'(1-F(q|e))}\left(\frac{1}{\mu w'\Big(1-F(q|e)\Big)u'(t^{sb}(q))}\right). \tag{87}$$

The W-MLRP gives

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{\frac{\mathrm{d}}{\mathrm{d}e}\Big(w'(1-F(q|e))f(q|e)\Big)}{w'(1-F(q|e)f(q|e))}\right) =& \frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{f_e(q|e)}{f(q|e)}\right)\\ &+\left(-\frac{\Big(w''(1-F(q|e))\Big)^2 F_e(q|e)f(q|e)}{\Big(w'(1-F(q|e))\Big)^2}\right.\\ &\left.-\frac{\Big(w''(1-F(q|e))\Big)f(q|e)}{w'(1-F(q|e))}\right),\end{aligned} \tag{88}$$

we use the above expression to rewrite Eq. (86) as

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{f_e(q|e)}{f(q|e)}\right) >& \frac{\Big(w''(1-F(q|e))\Big)^2 f(q|e)}{\Big(w'(1-F(q|e))\Big)^2}\left(-F_e(q|e)+\frac{1}{\mu u'(t^{sb}(q))w''(1-F(q|e))}\right)\\ &\frac{-\frac{\mathrm{d}}{\mathrm{d}e}\Big(w''(1-F(q|e))f(q|e)\Big)}{w'(1-F(q|e))}.\end{aligned} \tag{89}$$

Since $\lim_{q\to\bar{q}} w'(p) = +\infty$, then $\lim_{q\to\bar{q}} w''(p) = +\infty$. Therefore, the quantity

$$\frac{1}{\mu u'(t^{sb}(q))w''(1-F(q|e))}$$

goes to 0 as q approaches $\underline{q}$. All is left is

$$\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{f_e(q|e)}{f(q|e)}\right) > \frac{\Big(w''(1-F(q|e))\Big)^2 f(q|e)}{\Big(w'(1-F(q|e))\Big)^2}\left(-F_e(q|e) + \frac{-\frac{\mathrm{d}}{\mathrm{d}e}\Big(w''(1-F(q|e))f(q|e)\Big)}{w'(1-F(q|e))}\right), \tag{90}$$

which holds from the WMLRP (See Eq. (88)). Therefore, there exists an output level $q_h \in (\underline{q}, \bar{q})$ such that ${t^{sb}}'(q) > 0$ if $q \in [\underline{q}, q_h)$ and ${t^{sb}}'(q) < 0$ otherwise. The method for ironing is the same as in Proposition 3.

■

Appendix D: Adverse selection (Online)

Assume for simplicity that there are two types of agents: EUT and non-EUT. Also, suppose that non-EUT agents have RDU preferences with likelihood insensitivity and pessimism. Their weighting function exhibits an inverse-S shape and it yields $\mathbb{E}(t) > \tilde{\mathbb{E}}(t)$, where $\tilde{\mathbb{E}}(t|e) := \int_{\underline{q}}^{\bar{q}} u(t)\mathrm{d}w\Big(1 - F(q|e)\Big)$— a non-additive expectation. Various studies support this assumption (Bruhin et al., 2010; Harrison and Rutström, 2009).

The principal knows that she contracts with a EUT agent with probability π_E and with a non-EUT agent with the complement $1 - \pi_E$. The timing of her problem is as follows:

1. The agent learns his type: EU or L.
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.

The solution to this problem of moral hazard followed by adverse selection is provided next.

Proposition 11. *The optimal menu of contracts, $\{t^{sb}_{EU}, t^{sb}_{L}\}$, exhibits the following properties:*

1. *t^{sb}_{EU} satisfies $\mathbb{E}\left(u(t^{sb}_{EU})|\bar{e}\right) = c$ while t^{sb}_{L} satisfies $\tilde{\mathbb{E}}\left(u(t^{sb}_{L})|\bar{e}\right) = \tilde{\mathbb{E}}\left(u(t^{sb}_{L})|\bar{e}\right)$ if $w'\left(1 - F(q|\bar{e})\right) > 1$.*
2. *t^{sb}_{L} satisfies $\tilde{\mathbb{E}}(t^{sb}_{L}|\bar{e}) = c$ while t^{sb}_{EU} satisfies $\tilde{\mathbb{E}}(t^{sb}_{EU}|\bar{e}) = \tilde{\mathbb{E}}(t^{sb}_{EU}|\bar{e})$ if $w'\left(1 - F(q|\bar{e})\right) \leq 1$.*

Proof. the moral hazard incentive constraint of the EUT agent when given a contract t_{EU} is

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_{EU}(q)\Big) f(q|\bar{e})\,\mathrm{d}q - c \geq \int_{\underline{q}}^{\bar{q}} u\Big(t_{EU}(q)\Big) f(q|\underline{e})\,\mathrm{d}q, \tag{91}$$

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

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e}) \,\mathrm{d}q - c \geq \int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) w'\Big(1 - F(q|\underline{e})\Big) f(q|\underline{e}) \,\mathrm{d}q. \tag{92}$$

To distinguish between the two agents, t_L and t_{EU} must satisfy the adverse selection incentive-compatible constraints. That is for the EUT agent:

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_{EU}(q)\Big) f(q|\bar{e}) \,\mathrm{d}q - c \geq \max_{e \in \{\underline{e}, \bar{e}\}} \left\{ \int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) f(q|\bar{e}) \,\mathrm{d}q - c(e) \right\}, \tag{93}$$

and for the non-EUT agent:

$$\begin{aligned} &\int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e}) \,\mathrm{d}q - c \\ &\geq \max_{e \in \{\underline{e}, \bar{e}\}} \left\{ \int_{\underline{q}}^{\bar{q}} u\Big(t_{EU}(q)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e}) \,\mathrm{d}q - c(e) \right\}. \end{aligned} \tag{94}$$

Finally, the participation constraint of both agents, when the contracts targeted to them are selected, are

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_{EU}(q)\Big) f(q|\bar{e}) \,\mathrm{d}q - c \geq 0, \tag{95}$$

and

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e}) \,\mathrm{d}q - c \geq 0. \tag{96}$$

The standard approach to solve the adverse selection problem is to provide rents to the more efficient agent, which in turn depends on the impact of exerting high effort. Formally, efficiency for the non-EUT agent amounts to:

$$\int_{\underline{q}}^{\bar{q}} w'\left(1-F(q|\bar{e})\right) f(q|\bar{e})\mathrm{d}q - \int_{\underline{q}}^{\bar{q}} w'\left(1-F(q|\underline{e})\right) f(q|\underline{e})\mathrm{d}q = \\ w\left(1-F(q|\bar{e})\right) - w\left(1-F(q|\underline{e})\right). \tag{97}$$

Instead, for the EU agent, efficiency amounts to:

$$\int_{\underline{q}}^{\bar{q}} f(q|\bar{e})\mathrm{d}q - \int_{\underline{q}}^{\bar{q}} f(q|\underline{e})\mathrm{d}q = \left(1-F(q|\bar{e})\right) - \left(1-F(q|\underline{e})\right). \tag{98}$$

The W-MLRP (Assumption 4) 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 (97) to be larger than (98) is $w'(1-F(q|e)) > 1$ for any e. That is because

$$\int_{1-F(q|\underline{e})}^{1-F(q|\bar{e})} w'(s)\mathrm{d}s > \int_{1-F(q|\underline{e})}^{1-F(q|\bar{e})} \mathrm{d}s \Leftrightarrow \\ w\left(1-F(q|\bar{e})\right) - w\left(1-F(q|\underline{e})\right) > F(q|\underline{e}) - F(q|\bar{e}) \tag{99}$$

Under likelihood insensitivity $w'(1-F(q|e)) > 1$ holds in $q \in [\underline{q}, q_l^{**})$, where q_l^{**} satisfies $w'\left(1-F(q_l^{**}|e)\right) = 1$ and $w''\left(1-F(q_l^{**}|e)\right) > 0$, and also in $q \in (q_h^{**}, \bar{q}]$, where q_h^{**} is such that $w'\left(1-F(q_h^{**}|e)\right) = 1$ and $w''\left(1-F(q_h{**}|e)\right) < 0$.

Suppose 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. We first reduce the number of constraints to solve the principal's problem. Equations (95) and (94) immediately imply (96). Hence, at the optimum the participation constraint in (95) binds, while the participation constraint in (96) slacks.

From equation (93) and the constraint in (95), which binds at the optimum, we obtain:

$$0 \geq \max_{e\in\{\underline{e},\bar{e}\}} \left\{ \int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) f(q|\bar{e})\,\mathrm{d}q - c(e) \right\}, \tag{100}$$

which implies that EUT agents cannot afford to mimic non-EUT agents. Hence, the relevant adverse selection constrain is that in (94), which states that the non-EUT agent derives rents from mimicking the EUT agent. In contrast, equation (93) slacks at the optimum.

A direct implication that (94) binds is $t_L(q) \geq t_{EU}(q)$, which in turn gives

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_{EU}(q)\Big) f(q|\bar{e})\,\mathrm{d}q - c > \int_{\underline{q}}^{\bar{q}} u\Big(t_{EU}(q)\Big) f(q|\underline{e})\,\mathrm{d}q. \tag{101}$$

Hence, the moral hazard constraint in (91) slacks at the optimum.

Next, from the inequality in (96), which slacks at the optimum, along with equation (100), which holds with strict inequality, we obtain:

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})\,\mathrm{d}q - c > 0 \geq \max_{e \in \{\underline{e}, \bar{e}\}} \left\{ \int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) f(q|\bar{e})\,\mathrm{d}q - c(e) \right\}. \tag{102}$$

The above equation, together with the assumption of likelihood insensitivity with pessimism, implies that the non-EUT agent's perception of probabilities generate:

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) f(q|\underline{e})\,\mathrm{d}q > \int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) w'\left(1 - F(q|\underline{e})\right) f(q|\underline{e})\,\mathrm{d}q, \tag{103}$$

Equations (102) and (103) imply

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})\,\mathrm{d}q - c > \int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) w'\left(1 - F(q|\underline{e})\right) f(q|\underline{e})\,\mathrm{d}q. \tag{104}$$

and equation (92) is implied by other constraints in the principal's program.

Hence, at the solution only equations (94) and (95) bind. Thus, the optimal transfer given to the EUT agent, t_{EU}, must guarantee $\mathbb{E}\Big(u(t_{EU})|\bar{e}\Big) :=$

$\int_{\underline{q}}^{\bar{q}} u(t_{EU}) f(q|\bar{e})\mathrm{d}q = c$, satisfying the binding constraint in (95). Moreover, the transfer offered to the non-EUT, t_L, should satisfy

$$\tilde{\mathbb{E}}\Big(u(t_L)|\bar{e}\Big) := \int_{\underline{q}}^{\bar{q}} u(t_L) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})\mathrm{d}q = \tilde{\mathbb{E}}\Big(u(t_{EU})|\bar{e}\Big),$$

as implied by (94).

At implied probabilities that make the EUT is more efficient, the proof follows a similar logic. 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_L guarantees $\tilde{\mathbb{E}}\Big(u(t_L)|\bar{e}\Big) = c$ and t_{EU} guarantees $\mathbb{E}\Big(u(t_{EU})|\bar{e}\Big) = \mathbb{E}\Big(u(t_L)|\bar{e}\Big)$, at those output intervals. f

■

The principal offers a menu of contracts with a contract targeting each existing type. Thus, in our case the optimal menu consists of two contracts. Moreover, the principal implements high effort by making each of these contracts contingent on perfffformance either as described by Proposition 2, or as described by Proposition 6. This guarantees that incentives are given according to they way in which each type perceives output realizations. Importantly, to guarantee self-selection into the right contract, informational rents are included in one of the contracts. Specifically, the contract that targets the most efficient type is embellished with rents to discourage mimicking.

So far this solution seems standard. However, whether one agent is more efficient than the other crucially depends on probability weighting. When the agent's actions yield high and/or low probability, the agent suffering from likelihood insensitivity inflates the impact of his action on the probability of obtaining higher output levels. In that case, this irrational agent is more efficient; he is more likely to exert high effort with lower pay. In this situation, the menu in Proposition 11 (2) becomes relevant as it disincentivizes the non-EUT agent to mimic the EUT agent. Alternatively, when the agent's actions yield intermediate probability events, exerting effort seems pointless to the likelihood insensitive agent. The EUT agent is more efficient as he would require lower incentives to be motivated. The menu of contracts in Proposition 11 (1) becomes relevant in this case.

References

Abdellaoui, Mohammed, 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.

_ , **Han Bleichrodt, and Corina Paraschiv**, "Loss Aversion Under Prospect Theory: A Parameter-Free Measurement," *Management Science*, 2007, *53* (10), 1659–1674.

Baillon, Aurélien and Aysil Emirmahmutoglu, "Zooming in on ambiguity attitudes," *International Economic Review*, 2018, *59* (4), 2107–2131.

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

_ , _ , **Aysil Emirmahmutoglu, Johannes Jaspersen, and Richard Peter**, "When risk perception gets in the way: Probability weighting and underprevention," *Operations Research*, 2020, *70* (3), 1371–1392.

Baillon, Aurélien, Zhenxing Huang, Asli Selim, and Peter P. Wakker, "Measuring Ambiguity Attitudes for All (Natural) Events," *Econometrica*, 2018, *86* (5), 1839–1858.

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.

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, "Event exchangeability: Probabilistic sophistication without continuity or monotonicity," *Econometrica*, 2006, *74* (3), 771–786.

_ **and** _ , "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.

de Meza, David and David C. Webb, "Incentive design under loss aversion," *Journal of the European Economic Association*, 2007, *5* (1), 66–92.

Dierkes, Maik and Vulnet Sejdiu, "Indistinguishability of small probabilities, subproportionality, and the common ratio effect," *Journal of Mathematical Psychology*, 2019, *93*, 102283.

Ellsberg, Daniel, "Risk, ambiguity, and the Savage axioms," *The quarterly journal of economics*, 1961, pp. 643–669.

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.

French, Kenneth R and James M Poterba, "Were Japanese stock prices too high?," *Journal of Financial Economics*, 1991, *29* (2), 337–363.

Ghirardato, Paolo, "Agency theory with non-additive uncertainty," Technical Report, mimeo, http://web. econ. unito. it/gma/paolo/age. pdf 1994.

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.

Halevy, Yoram, "Ellsberg revisited: An experimental study," *Econometrica*, 2007, *75* (2), 503–536.

Harrison, Glenn W and E Elisabet Rutström, "Expected utility theory and prospect theory: One wedding and a decent funeral," *Experimental economics*, 2009, *12* (2), 133–158.

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.

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.

Kellner, Christian, "The principal-agent problem with smooth ambiguity," *Review of Economic Design*, 2017, *21* (2), 83–119.

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

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

Lang, Matthias, "First-order and second-order ambiguity aversion," *Management Science*, 2017, *63* (4), 1254–1269.

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

l'Haridon, Olivier and Ferdinand M Vieider, "All over the map: A worldwide comparison of risk preferences," *Quantitative Economics*, 2019, *10* (1), 185–215.

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.

_ , "The theory of moral hazard and unobservable behaviour: Part I," *The Review of Economic Studies*, 1999, *66* (1), 3–21.

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.

Rogerson, William P, "The first-order approach to principal-agent problems," *Econometrica*, 1985, pp. 1357–1367.

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.

Savage, Leonard J, *The foundations of statistics*, John Wiley and Sons, 1954.

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.

Trautmann, Stefan T and Gijs Van De Kuilen, "Ambiguity attitudes," *The Wiley Blackwell handbook of judgment and decision making*, 2015, *2*, 89–116.

Tversky, Amos 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.

Vierø, Marie-Louise, "Bait contracts," *Oxford Economic Papers*, 2014, *66* (3), 702–720.

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.

Yaari, Menahem E, "The Dual Theory of Choice under Risk," *Econometrica*, 1987, *55* (1), 95–115.