Optimal Incentives without Expected Utility*

Víctor González-Jiménez†

August 10, 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, Risk Attitude, Incentives, Probability Weighting, Mechanism Design.

---

*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 nearly all fields of economics. The contracts it predicts, however, often do not match those observed in practice (Lazear and Oyer, 2007; Prendergast, 1999; Salanié, 2003). Notably, the bulk of the literature captures risk attitudes with expected utility. Expected utility, while theoretically appealing, is not an accurate description of choice under risk (Starmer, 2000).[2]

In our paper we investigate whether relaxing the assumption of expected utility maximization changes the type of contract predicted by the theory. In particular, we consider agents who *distort* probabilities as documented by abundant evidence from decision theory (Abdellaoui 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

[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 also 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).

general overconfidence. There, agents would, for example, violate first-order stochastic dominance.

Our main contribution is to show how the principal can take advantage of an agent who distorts probabilities. We consider different types of probability distortion. With these we find that the contracts offered mimic those observed in real-life, such as fixed salaries, high-performance commissions or long-shots, 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 if the highest performance levels realize—in other words, a long-shot.

We further show that, when optimism is moderate and 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, on its own, induces high effort; the agent's confidence that high performance levels realize is enough 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 realizations, which ultimately makes an incentive-compatible contract excessively costly. The principal would end up offering a contract with a constant payment—a fixed wage.

Second, we go beyond optimism and pessimism and consider also probability distortions stemming from the agents' cognitive limitations to perceive probabilities. These probability distortions are referred as likelihood insensitivity (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.

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., 2020; 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.

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 and also look at what happens when effort is *not* contractible, for different possible shapes of probability weighting, and do not restrict our analyses to one type of compensation scheme or to a parametric form of utility and probability weighting functions. Additionally, Gonzalez-Jimenez (2020) demonstrates that stochastic contracts are preferred to deterministic contracts when agents distort probabilities. That paper is silent about the specific shape of optimal stochastic contracts, as well as the power and location of incentives over the performance interval. We provide an answer to these unresolved questions.

We also contribute to the contract theory literature. We speak to a well-known paradox put forward by Salanié (2003) stating that the complex

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

theoretical 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 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 $c > 0$.

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* $u' > 0$*,* $u'' < 0$*, 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 observed by the principal. Additionally, 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 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 $f(q|e)$. 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. The agent can thus 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 is a function that 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.

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

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

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

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, i.e. a probability $\hat{p} \in (0,1)$ such that $w(\hat{p}) = \hat{p}$.

The seemingly drastic assumptions of extreme sensitivity to rare and almost-certain events, $\lim_{p\to 1} w'(p) = \infty$ 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) pg.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]

These assumptions have relevant implications that are formally presented next. We relegate all proofs to Appendix A.

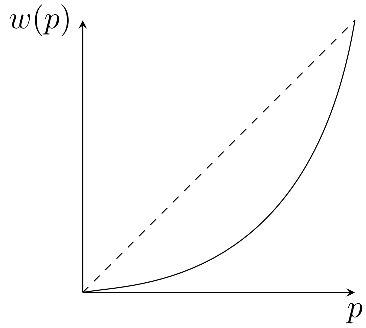
Lemma 1. *If* $\lim_{p\to 0} w'(p) = +\infty$, *then* $\lim_{p\to 0} w''(p) = -\infty$ *and* $\lim_{p\to 0} \frac{w''(p)}{w'(p)} = -\infty$.

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

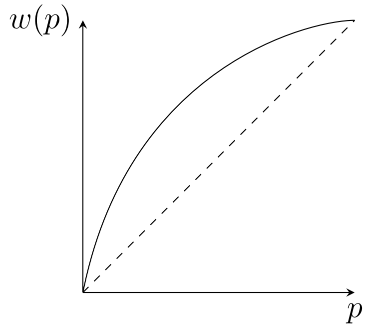
In words, the second derivative of the weighting function and the analog of the Arrow-Pratt measure in probabilities, $\frac{w''(p)}{w'(p)}$, are unbounded at small

[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.

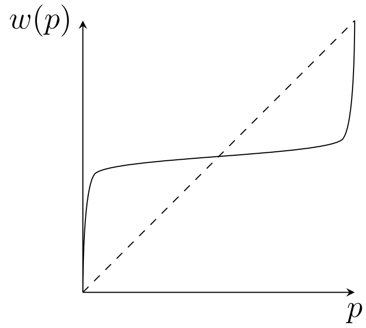
Figure 1: Examples of probability weighting functions



(a) Pessimism



(b) Optimism



(c) Likelihood insensitivity

Note: Dashed lines represent accurate perception of probability.

probabilities for a general probability weighting function. 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. These implications will not only be useful for proving our main results, but also formalize the findings of Dierkes and Sejdiu (2019).

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

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

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.

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 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* $\lim_{p\to 1} w'(p) = 0$*.*

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:

[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.

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 $\lim_{p \to 1} w'(p) = +\infty$.*

3.1 First best

If effort is contractible, the principal only needs to ensure participation. Interestingly, she can sometimes do so while at the same time taking advantage of the irrationalities exhibited by non-EUT agents. The following Proposition describes the resulting first-best optimal contracts for each agent.

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

1. *If the agent is EUT, the contract t^{fb}_{EU} is constant in q;*
2. *if the agent exhibits optimism, $t^{fb}_{O}(q)$ is everywhere increasing in q;*
3. *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 the fixed transfer ensures that the contract will be accepted by the agent.

When facing an optimist, the first-best contract increases in performance. While risky, this contract offers full insurance to the agent; it provides larger payments for realizations that are perceived to be more likely and lower payments for realizations that are perceived to be less likely. The principal is taking advantage of the agent's sensitivity to probabilities. From her point-of-view she provides larger payments for unlikely events and smaller payments for more likely 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 production sabotage and ensure participation, the principal simply offers a constant transfer that yields utility equal to his outside option. This solution, however, yields an impossibility: it eliminates risk, rendering impossible the exploitation of the agent's irrationality.[9]

To investigate how 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 is 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 however conditional on the agents' coefficient of absolute risk aversion not becoming larger with the contract's change.[10]

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 $\frac{dt^{fb}_{O,i}(q)}{dq} > \frac{dt^{fb}_{O,j}(q)}{dq}$.*

[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 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.

Furthermore, the opportunity cost faced by the principal from not being able to exploit pessimism while, at the same time, offering insurance (Proposition 1) becomes more severe as the agent becomes more pessimistic.

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. Therefore, the cost of offering t^{fb}_{EU} instead of t^{fb}_{P} 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 that 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. 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 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, who is fully informed about the agent's risk attitudes, anticipates how the agent's probability distortions affect his perception about the informativeness of output realizations 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. Assumption 4 entails that the results presented below take advantage of this naivete.

The W-MLRP is more stringent than the standard MLRP. 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-a-vis the MLRP.

Lemma 3. *The W-MLRP implies:*

1. $w\big(1-F(q|\bar{e})\big)\geq w\big(1-F(q|\underline{e})\big)$;
2. *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 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. 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.

Throughout, we will assume the W-MLRP holds; note, however, that thanks to Lemma 4 the results presented below can be obtained using the standard MLRP along with restrictions on the concavity or convexity of the weighting function.

The next Proposition describes the properties of the second-best contracts that solve the principal's program when she faces an optimist or a pessimist.

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:

1. *Identical to the first-best contract, $t_O^{fb}(q)$, in Proposition 1 if $c < \hat{c}_O$;*
2. *strictly increasing in q everywhere and introducing rewards and punishments with respect to $t_O^{fb}(q)$ if $c \geq \hat{c}_O$.*

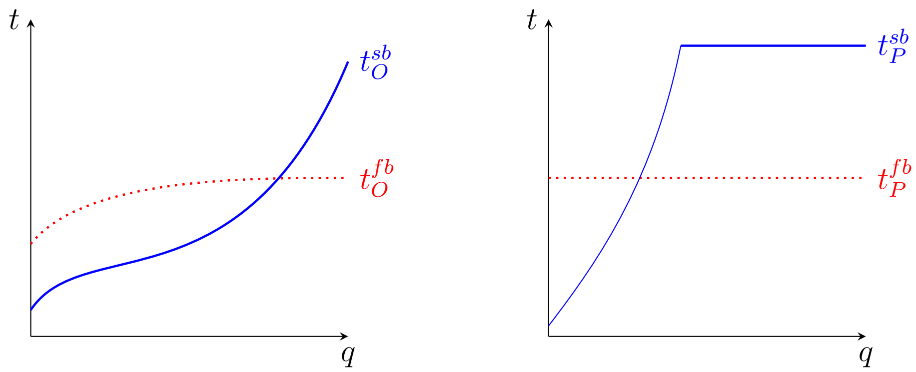
Under pessimism, the second-best contract, $t_P^{sb}(q)$, is:

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

For an optimistic agent, the first-best contract from Proposition 1 can suffice to incentivize high effort. Recall that such contract specifies transfers that increase in performance. Since the optimistic agent believes that high performance levels are more likely to realize, that contract can convince him that exerting high effort is profitable. Nevertheless, when the cost of exerting effort is sufficiently high, the optimistic agent needs to be incentivized with a contract that deviates from the first-best contract by providing lower transfers at low output levels and larger transfers at the high-end of the output space. The combination of large transfers at high performance levels and optimism inflates the agent's perceived benefit of exerting high effort. An example of these two contracts is presented in Figure 2a.

When the agent is pessimistic, the principal concentrates incentives at low performance levels. She offers a contract that provides larger transfers in exchange of higher performance levels at the low-end of the output interval

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.

while being performance-insensitive in the remainder of the output interval. These incentives, together with the agent's irrational perception that low performance levels realize more often, motivate him to exert high effort.[11] An example of these contracts is presented in Figure 2b.

We use Definition 3 to provide comparative statics that show how the optimal second-best contract changes as optimism (pessimism) becomes stronger.

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*

1. *Leads to a lower cost level $\hat{c}_O$ shortening the set $c \in (0, \hat{c}_O]$ in which $t_O^{sb}(q) = t_O^{fb}(q)$ if $q > q^*$;*
2. *Enlarges the set $q > q^*$.*

The comparative static from Corollary 3 shows that the principal implements one of the contracts in Proposition 3 according to the agent's degree of optimism. Specifically, stronger optimism makes it less likely that the

[11] Again, it would be inefficient to implement a schedule that is increasing everywhere in performance since the pesssimistic agent perceives high realizations to be unlikely. To motivate this agent with such a contract shape, the principal would need to incurr in excessively large transfers, yielding excessive expenditures.

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 can backfire and demotivate him. That is because this confidence might be excessive and make him believe that low effort also suffices to achieve high output levels. This makes necessary for the principal to implement powerful incentives that enhance the benefits from exerting high effort. Put differently, the agent needs to be incentivized not to "rest on his laurels."

We study next the consequences of stronger degrees of pessimism.

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

Stronger pessimism leads the principal to concentrate incentives at lower performance levels. Punishments and rewards are further shifted toward the lower-end of the output space, which implies that the incentive scheme becomes flat for a larger output segment.

Corollary 4 also shows that the principal implements one of the contracts in Proposition 3 according to the agent's degree of pessimism. When pessimism is strong enough, implementing incentives is no longer possible. The principal would need to concentrate rewards and punishments in a thin output set located at the neighborhood of $\underline{y}$. This yields an inefficient expenditure, since, to avoid sabotage, a flat amount equal to the largest reward needs to be paid in the remainder of the output space. This is unaffordable to the principal, leaving her no other remedy than to give up incentive compatibility and ensure participation with a salary.

To conclude this section, we compare the transfers of the second-best contracts presented in Proposition 2 and 3. We focus on the contracts that emerge when the agent has strong optimism or moderate pessimism, the cases in which incentive-compatibility can be implemented and requires a deviation from first-best contracts. The following corollary formalizes these comparisons.

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

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

2. $t_P^{sb}(q)$ *(Proposition 3) 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).*

When the incentive compatible constraint binds, the contract given to an optimist offers lower transfers at high performance levels as compared to the contract given to an EUT agent. This result shows, once again, how the principal exploits optimism to obtain incentives. She offers moderate transfers that will be complemented by the agent's confidence that high output levels realize more often. Furthermore, the contract of the optimist also specifies larger transfers at low output levels as compared to that given to an EUT agent. This incentivizes the optimistic agent to exert high effort because, even for realizations perceived to be unlikely, it is profitable to obtain larger output levels.

Compared to the EUT agent, a pessimist agent receives a contract that specifies larger transfers at low performance levels. Moreover, this agent is given lower transfers at high performance realizations. Incentives are thus concentrated where it matters to the pessimist: at the low-end of the output interval.

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 $\lim_{p \to 1} w'(p) = +\infty$.*

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 contract is an option-like incentive scheme.

Proposition 4. *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 performance.*

The increasing part of the first-best contract reflects the agent's risk seeking attitudes emerging from his tendency to overweight low probability events. Therefore, insurance is achieved by exposing the agent to risk at high output levels. Instead, the flat part of the contract reflects the agent's insensitivity to intermediate probabilities and the risk aversion from assigning large probability weights to large probabilities. The principal protects the agent from risk for those realizations.

To investigate how contracts change as likelihood insensitivity becomes stronger, we will talk about agents who are more likelihood insensitive than others. First, 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 exists constants ϵ, ϵ' such that*

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

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

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

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

We 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* $\phi : [0, 1] \rightarrow [0, 1]$ *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 middle-ranged probabilities. In other words, his weighting function exhibits a more pronounced inverse-S shape. The following corollary shows how the first-best contract of Proposition 4 changes as an agent is more likelihood insensitive.

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* $(q_I, \bar{q}]$ *in which pay is performance insensitive.*

When facing an agent with stronger likelihood insensitivity, the principal must offer a first-best contract specifying a larger performance-insensitive segment. Such contract matches the agent's increased insensitivity to probabilities, as well as the stronger risk aversion from assigning larger probability weights to worst events— smallest output levels realizing.

4.2 Second best

When effort is not contractible, likelihood insensitivity makes the principal's program more restrictive. Not only because the incentive compatibility constraint needs to be included, but also 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). Note that around $\tilde{p}$, probabilities are almost indistinguishable to the agent. Thus, 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, does not necessarily apply when those probabilities are in the neighborhood of $\tilde{p}$. Therefore, the W-MLRP must be explicitly assumed.

The next Proposition presents the second-best contract offered to the likelihood-insensitive agent.

Proposition 5. *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, $t_O^{sb}(q)$:*

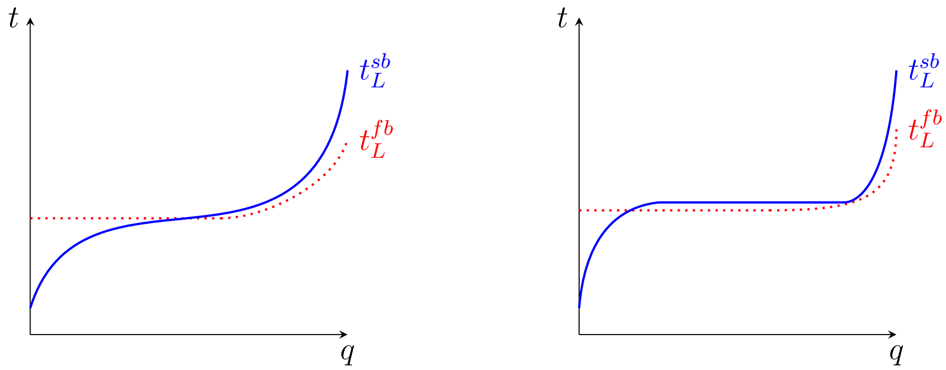
1. *is identical to the first-best contract, $t_L^{fb}(q)$ from Proposition 4 if $c < \hat{c}_L$;*
2. *increases everywhere in performance and exhibits steep payment increments at extreme performance levels if $c \geq \hat{c}_L$;*
3. *pays a fixed amount for some finite, fixed, compact interval, but above and/or below this interval pay steeply increases in performance if $c \geq \hat{c}_L$.*

The first part of the proposition shows that the solution to the principal's program does not require the incentive compatibility constraint to bind. The rationale for this solution is analogous to that presented in Proposition 3 for the optimist. Namely, the likelihood insensitive agent's perception that high output levels are more likely to realize along with the first-best contract having an increasing segment at those output levels, suffices to motivate the agent.

When high effort is sufficiently costly, i.e. $c \geq \hat{c}_L$, the optimal contract requires the implementation of rewards and punishments with respect to the first-best contract. These incentives are concentrated at extreme output levels, i.e. the probability events that are given the most weight by this agent. The resulting contract can be either everywhere increasing, as presented in Proposition 5 (2) and exemplified in Figure 4a, or insensitive in performance for intermediate performance levels, as presented in Proposition 5 (3) and exemplified Figure 4b. Whether one contract or the other applies depends on the degree of likelihood insensitivity as it will be shown next.

The following corollary shows how the second-best contract in Proposition 5 changes with stronger degrees of likelihood insensitivity.

Figure 3: Illustration of Proposition 4 and Proposition 5.



(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.

Corollary 7. *Assume that the incentive compatibility constraint binds. If agents i and j are likelihood insensitive and agent i is more likelihood insensitive than j, the second-best contract, t_L^{sb} (Proposition 5), offered to i exhibits a larger segment in which pay is performance insensitive.*

An agent with stronger likelihood insensitive is more likely to be offered the contract in Proposition 5 (3). That incentive scheme matches better his more severe insensitivity to probabilities. Instead, an agent with moderate likelihood insensitivity responds to incentives over the entire output space. He can be thus better motivated when obtaining the contract in Proposition 5 (2).

To conclude this section, we compare the transfers of the optimal contract presented in Proposition 5 (2) and (3) to those in the contract given to an EUT agent.

Corollary 8. *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 $w''\big(1 - F(q^{**}|\bar{e})\big) < 0$. Assume that the incentive compatibility constraint binds. Contract $t_L^{sb}(q)$ (Proposition 5) offers lower transfers in $q \in [\tilde{q}, q^{**})$ as compared to t_{SB}^{EU} (Proposition 2). However, that contract pays higher transfers in*

$q \in [\underline{q}, q^*)$, at the highest output levels in $q \in [q^{**}, \bar{q}]$, and at the lowest output levels in $q \in [q^*, \tilde{q})$.

Corollary 8 shows that the second-best contract given to a likelihood-insensitive agent specifies lower transfers at intermediate performance levels as compared to the contract given to the EUT agent. Moreover, the contracts from Proposition 5 (2) and (3) offer higher transfers at the smallest and largest performance levels. The principal thus concentrates incentives where it matters to the likelihood insensitive agent.

5 Extensions

5.1 Agents with 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 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*. Typically, 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). We follow this assumption to simplify the present analysis.

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,*

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

with the following properties:

- $\lambda > 1$;
- $v(0) = 0$;
- $v' \geq 0$ *for all* $q \in [\underline{q}, \bar{q}]$;
- $v'' < 0$ *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 an 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.[12]

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

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

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

[12] 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:

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

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

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

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

Proposition 6. *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}:*

1. *pays r everywhere if $\hat{q} = \bar{q}$;*
2. *pays r in $q < \hat{q}$ and depends on performance as in Proposition 3 or Proposition 5 in $q \geq \hat{q}$ if $\hat{q} \in (\underline{q}, \bar{q})$;*
3. *depends on performance as in Proposition 3 or Proposition 5 if $\hat{q} = \underline{q}$.*

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 5. 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 6, 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 $\lambda > 2$.

5.2 Additional extensions

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

Adverse selection

Thus far, we have considered a situation 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 setting. In the following, we consider a setting in which that assumption is relaxed. Consequently, the principal's goal is to screen first among the different types of agents to then incentivize high effort.

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 7. *The optimal menu of contracts, $\{t_{EU}^{sb}, t_L^{sb}\}$, exhibits the following properties:*

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

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 performance either as described by Proposition 2, or as described by Proposition 5. 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 7 (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 7 (1) becomes relevant in this case.

Non-EUT principal

Assume now that the principal also evaluates probabilities non-linearly. We however maintain the assumption that utility is linear. Thus, while she is able to pool large financial risks, say due to having sufficient liquidity, she might assign too large or small weights to the likelihood of events.

This problem can be solved with the tools provided in Sections 3 and 4. That is because, from the principal's point-of-view, who is fully informed about the agent's risk attitudes, the agent can be either more pessimistic or more optimistic than what she is— in the sense of Definition 3—, and/or more or less likelihood insensitive—in the sense of Definition 6.

For example, when the principal is pessimistic, an EUT agent is, from her point-of-view, optimistic. She will thus give him one of the contracts described in Proposition 3. The more pessimistic the principal becomes, the more important becomes the refinement given in Corollaries 5. Similarly, when the principal exhibits moderate likelihood insensitivity, an agent with severe likelihood insensitivity might exhibit, from her standpoint, a moderate degree of insensitivity. Thus the contract in Proposition 5 (2) applies. Corollary 7 can be used to provide adjustments to the contract according to the degree of likelihood insensitivity of both agent and principal.

All in all, the main results of Sections 3 and 4 and the refinements to these optimal contracts given in Corollaries 5, 4 and 7 solve this modified problem.

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 incentive schemes observed in practice. For example, performance-insensitive salaries under strong pessimism, long-shot contracts under moderate optimism, and bonuses under strong likelihood insensitivity. We thus provide a foundation for simple contracts based on preference.

This paper opens avenues for future research. A theoretical possibility is to abandon the one-shot framework considered in our model and investigate optimal contracting in a dynamic setting. In these settings, the standard solution requires contracts with the "Martingale Property." When agents aggregate risk using non-additive probability measures, it is unclear how this property emerges, and whether it leads to contracts that are often observed in practice. Another interestring possibility, would be performing an empirical test of the validity of our results. One could, for example, use the econometric tools developed in Barseghyan et al. (2013) and more recently in Barseghyan et al. (2021) to understand how risk preferences match labor contracts in different industries.

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

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

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, $t^{fb}_{EU}(q)$, is everywhere constant and satisfies

$$\frac{1}{u'\left(t^{fb}_{EU}(q)\right)} = \nu.$$

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, $t^{fb}_O(q)$, is increasing in q.

To better understand the shape of $t^{fb}_O(q)$ 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 $\lim_{q\to\bar{q}} t^{fb\prime}_O(q) = +\infty$. Moreover, from Definition 1 and Lemma 5 we have $\lim_{p\to 1}\frac{w''(p)}{w'(p)} = -\infty$, implying $\lim_{q\to\underline{q}} t^{fb\prime}_O(q) = +\infty$

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, $t^{fb}_P(q)$, is strictly decreasing in q.

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

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

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

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

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

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

Subtracting (9) from (8) gives

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

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

(10) is positive, so

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

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 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, the contract would be rejected.To ensure participation the principal must offer the pessimist $t_{EU}^{fb}(q)$. ■

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

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

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 1 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

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

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 assumption $w_i(1) = 1$. A similar rationale disregards $w_i'(p) > w_j'(p)$ for all $p \in (0,1)$. Hence, $w_i'(p) = w_j(p)$ for some $p \in (0,1)$.

Lemma 1 and Lemma 5 shows that $\lim_{p\to 0} w'(p) = +\infty$ and $\lim_{p\to 1} w'(p) = 0$. Let $w_J(p) := \eta\big(w_j(p)\big)$ where η is a concave, increasing, and continuous function. Accordingly, $-\frac{w_J''(p)}{w_J'(p)} > -\frac{w_j''(p)}{w_j'(p)}$ $\forall p \in (0,1)$ as shown in the first part of the Lemma. Therefore, $w'_J(p)$ tends to infinity faster than $w_j(p)$ as $p \to 0^+$. Assumption 3 states that $w'(p)$ 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 $w_i'(p)$ 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

$$\begin{aligned} &\frac{\mathrm{d}t^{fb}_{O,i}(q)}{\mathrm{d}q} > \frac{\mathrm{d}t^{fb}_{O,j}(q)}{\mathrm{d}q} \\ \Leftrightarrow\ &\frac{u'\left(t^{fb}_{O,i}(q)\right)}{u''\left(t^{fb}_{O,i}(q)\right)}\frac{w''_i(\bar{p})}{w'_i(\bar{p})}f(q|\bar{e}) > \frac{u'\left(t^{fb}_{O,j}(q)\right)}{u''\left(t^{fb}_{O,j}(q)\right)}\frac{w''_j(\bar{p})}{w'_j(\bar{p})}f(q|\bar{e}). \end{aligned}$$

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

$$\begin{aligned} &\frac{u'\left(t^{fb}_{O,i}(q)\right)}{u''\left(t^{fb}_{O,i}(q)\right)}\left\{\frac{\theta''\left(w_j(\bar{p})\right)}{\theta'\left(w_j(\bar{p})\right)}w'_j(\bar{p}) + \frac{w''_j(\bar{p})}{w'_j(\bar{p})}\right\} > \frac{u'\left(t^{fb}_{O,j}(q)\right)}{u''\left(t^{fb}_{O,j}(q)\right)}\frac{w''_j(\bar{p})}{w'_j(\bar{p})} \\ \Leftrightarrow\ &\frac{w''_j(\bar{p})}{w'_j(\bar{p})}\left\{\frac{u'\left(t^{fb}_{O,i}(q)\right)}{u''\left(t^{fb}_{O,i}(q)\right)} - \frac{u'\left(t^{fb}_{O,j}(q)\right)}{u''\left(t^{fb}_{O,j}(q)\right)}\right\} \\ &\qquad + \frac{u'\left(t^{fb}_{O,i}(q)\right)}{u''\left(t^{fb}_{O,i}(q)\right)}\frac{\theta''\left(w_j(\bar{p})\right)}{\theta'\left(w_j(\bar{p})\right)}w'_j(\bar{p}) > 0, \end{aligned} \tag{13}$$

which can be rewritten as

$$-\frac{\theta''\left(w_j(\bar{p})\right)}{\theta'\left(w_j(\bar{p})\right)}w'_j(\bar{p}) > \frac{w''_j(\bar{p})}{w'_j(\bar{p})}\left(1 - \frac{\frac{u''(t^{fb}_{O,i}(q))}{u'(t^{fb}_{O,i}(q))}}{\frac{u''(t^{fb}_{O,j}(q))}{u'(t^{fb}_{O,j}(q))}}\right) \tag{14}$$

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, (14) is true if

$$-\frac{u''(t^{fb}_{O,j}(q))}{u'(t^{fb}_{O,j}(q))} > -\frac{u''(t^{fb}_{O,i}(q))}{u'(t^{fb}_{O,i}(q))}. \tag{15}$$

Instead, (14) shows that under $-\frac{u''(t^{fb}_{O,j}(q))}{u'(t^{fb}_{O,j}(q))} < -\frac{u''(t^{fb}_{O,i}(q))}{u'(t^{fb}_{O,i}(q))}$ 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

$$\int_{\underline{q}}^{\bar{q}} u\Big(t^{fb}_{P,j}(q)\Big) w_j'\Big(1-F(q|\bar{e})\Big) f(q|\bar{e})\,\mathrm{d}q = u\Big(t^{fb}_{P,j}(\underline{q})\Big) + \int_{\underline{q}}^{\bar{q}} u'\Big(t^{fb}_{P,j}(q)\Big) {t^{fb}_{P,j}}'(q) w_j\Big(1-F(q|\bar{e})\Big)\,\mathrm{d}q. \quad (16)$$

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

$$\int_{\underline{q}}^{\bar{q}} u\Big(t^{fb}_{P,j}(q)\Big) w_i'\Big(1-F(q|\bar{e})\Big) f(q|\bar{e})\,\mathrm{d}q = u\Big(t^{fb}_{P,j}(\underline{q})\Big) + \int_{\underline{q}}^{\bar{q}} u'\Big(t^{fb}_{P,j}(q)\Big) {t^{fb}_{P,j}}'(q) w_i\Big(1-F(q|\bar{e})\Big)\,\mathrm{d}q. \quad (17)$$

Subtracting the right-hand side of (16) from the one of (17) gives

$$\int_{\underline{q}}^{\bar{q}} u'\Big(t^{fb}_{P,j}(q)\Big) {t^{fb}_{P,j}}'(q) \Big(w_i\Big(1-F(q|\bar{e})\Big) - w_j\Big(1-F(q|\bar{e})\Big)\Big) f(q|e)\,\mathrm{d}q. \quad (18)$$

Because ${t^{fb}_{P,j}}'(q) < 0$ (Proposition 1) and $w_i(p) < w_j(p)$ (Lemma 7) under pessimism, (18) is positive, so

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

(19)

For j, the candidate solution from the first-order condition ensures the participation constraint

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

so from (19)

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

Similarly, for i, the participation constraint holds:

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

Therefore,

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

which is implied by $t_{P,j}^{fb} > t_{P,i}^{fb}$ and, for the ironed solutions, $\tilde{t}_{P,j}^{fb} > \tilde{t}_{P,i}^{fb}$. Proposition 1 shows that to ensure participation t_{EU}^{fb} is given to both i and j. Since $t_{EU}^{fb} > \tilde{t}_{P,j}^{fb} > \tilde{t}_{P,i}^{fb}$ the cost borne by the principal of not being able to implement $\tilde{t}_{P,j}^{fb}$ and $\tilde{t}_{P,i}^{fb}$ is higher the more pessimistic the agent is. ∎

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

$$\begin{aligned} &\frac{w'\Big(1 - F(q_1|\underline{e})\Big) f(q_1|\underline{e})}{w'\Big(1 - F(q_1|\bar{e})\Big) f(q_1|\bar{e})} \leq \frac{w'\Big(1 - F(q_0|\underline{e})\Big) f(q_0|\underline{e})}{w'\Big(1 - F(q_0|\bar{e})\Big) f(q_0|\bar{e})} \\ \Leftrightarrow\ & w'\Big(1 - F(q_1|\underline{e})\Big) f(q_1|\underline{e}) w'\Big(1 - F(q_0|\bar{e})\Big) f(q_0|\bar{e}) \leq \\ & \qquad w'\Big(1 - F(q_0|\underline{e})\Big) f(q_0|\underline{e}) w'\Big(1 - F(q_1|\bar{e})\Big) f(q_1|\bar{e}). \end{aligned} \tag{20}$$

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

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

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

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

Integrating (20) again, but now with respect to q_1 from q_0 to $\bar{q}$, and following the same steps gives

$$\frac{w\Big(1 - F(q_0|\underline{e})\Big)}{w\Big(1 - F(q_0|\bar{e})\Big)} \leq \frac{w'\Big(1 - F(q_0|\underline{e})\Big) f(q_0|\underline{e})}{w'\Big(1 - F(q_0|\bar{e})\Big) f(q_0|\bar{e})}. \tag{22}$$

Letting $q_0 = q_1 = q$ and combining (21) and (22) gives

$$\begin{aligned}
&\frac{w\big(1 - F(q|\underline{e})\big)}{w\big(1 - F(q|\bar{e})\big)} \leq \frac{1 - w\big(1 - F(q|\underline{e})\big)}{1 - w\big(1 - F(q|\bar{e})\big)} \\
\Leftrightarrow\ & w\big(1 - F(q|\bar{e})\big) \geq w\big(1 - F(q|\underline{e})\big)
\end{aligned}$$

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

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{w'\big(1 - F(q|\underline{e})\big) f(q|\underline{e})}{w'\big(1 - F(q|\bar{e})\big) f(q|\bar{e})}\right) &= \frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{w'\big(1 - F(q|\underline{e})\big)}{w'\big(1 - F(q|\bar{e})\big)}\right) \frac{f(q|\underline{e})}{f(q|\bar{e})} \\
&\quad + \frac{w'\big(1 - F(q|\underline{e})\big)}{w'\big(1 - F(q|\bar{e})\big)} \frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{f(q|\underline{e})}{f(q|\bar{e})}\right) \\
&= \frac{w'\big(1 - F(q|\underline{e}\big)}{w'\big(1 - F(q|\bar{e}\big)}\left[\left(\frac{w''\big(1 - F(q|\bar{e})\big)}{w'\big(1 - F(q|\bar{e})\big)} f(q|\bar{e})\right.\right. \\
&\quad \left. - \frac{w''\big(1 - F(q|\underline{e})\big)}{w'\big(1 - F(q|\underline{e})\big)} f(q|\underline{e})\right) \frac{f(q|\underline{e})}{f(q|\bar{e})} \\
&\quad \left. + \frac{\mathrm{d}}{\mathrm{d}q} \frac{f(q|\underline{e})}{f(q|\bar{e})}\right].
\end{aligned} \tag{23}$$

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 (23). ■

Proposition 3

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

$$\begin{aligned}-f(q|\bar{e}) + \mu \big[u'(t^{sb}(q))(w'(1 - F(q|\bar{e}))f(q|\bar{e}) - w'(1 - F(q|\underline{e}))f(q|\underline{e})) \big] \\ + \nu u'(t^{sb}(q)) w'(1 - F(q|\bar{e})) f(q|\bar{e}) = 0,\end{aligned} \tag{24}$$

and, after re-arranging,

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

Incentive constraint is binding We first show that $\mu > 0$ might not always hold the optimum. Suppose instead 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 ν we get

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_O^{fb}(q)\Big) w'(1 - F(q|\bar{e})) f(q|\bar{e}) \, \mathrm{d}q - c$$
$$> \int_{\underline{q}}^{\bar{q}} u\Big(t_O^{fb}(q)\Big) w'(1 - F(q|\underline{e})) f(q|\underline{e}) \, \mathrm{d}q. \quad (26)$$

Integration by parts of (3) yields

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

which we use to rewrite (26) as

$$\int_{\underline{q}}^{\bar{q}} u'\Big(t_O^{fb}(q)\Big) t_O^{fb\prime}(q) \bigg(w\Big(1 - F(q|\bar{e})\Big) - w\Big(1 - F(q|\underline{e})\Big) \bigg) \, \mathrm{d}q > c. \quad (27)$$

According to Lemma 3, Assumption 4 implies $w\,(1 - F(q|\bar{e})) \geq w\,(1 - F(q|\underline{e}))$ which, together with $t_O^{fb\prime}(q) > 0$ (Proposition 1) and $u'(t) > 0$ (Assumption 1), imply that the left-hand side of (27) 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 $c \leq \hat{c}_O$, (27) holds: $\mu = 0$ and $t_O^{sb}(q) = t_O^{fb}(q)$; on the other hand,
- if $c > \hat{c}_O$, (27) does not hold: $\mu > 0$ and $t_O^{sb}(q)$ satisfies (25).

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

$$\int_{\underline{q}}^{\bar{q}} u\Big(\tilde{t}_P^{fb}(q)\Big) w'(1 - F(q|\bar{e})) f(q|\bar{e}) \,\mathrm{d}q - c > \int_{\underline{q}}^{\bar{q}} u\Big(\tilde{t}_P^{fb}(q)\Big) w'(1 - F(q|\underline{e})) f(q|\underline{e}) \,\mathrm{d}q$$
$$\Leftrightarrow u\Big(\tilde{t}_P^{fb}(q)\Big) - c > u\Big(\tilde{t}_P^{fb}(q)\Big)$$
$$\Leftrightarrow -c > 0.$$

The second inequality is due to $\tilde{t}_P^{fb}$ being constant in q. 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 $t^{sb}(q)$. Differentiate (25) with respect to q to obtain:

$$\begin{aligned} t^{sb\prime}(q) =& \frac{u'(t^{sb}(q)) w''(1 - F(q|\bar{e}))}{u''(t^{sb}(q)) w'\left(1 - F(q|\bar{e})\right)} f(q|\bar{e}) \\ &+ \mu \frac{w'\left(1 - F(q|\bar{e})\right) u'(t^{sb}(q))^2}{u''(t^{sb}(q))} \frac{\mathrm{d}}{\mathrm{d}q} \left(\frac{w'\Big(1 - F(q|\underline{e})\Big) f(q|\underline{e})}{w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})} \right). \end{aligned} \quad (28)$$

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), $u'(t^{sb}(q)) > 0$ and $u''(t^{sb}(q)) < 0$ (Assumption 1), and $w'(p) > 0$ (Assumption 3), so the second term on the right-hand side of (28) is always positive. The first term on the right-hand side of (28) is identical to the right-hand side of (7) in Proposition 1, which is 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 (28) are positive. Hence, $t_O^{sb\prime}(q) > 0$ everywhere.

We also study (28) 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 (28) is negative while the second one is positive. Hence, the sign of ${t_P^{sb}}'(q)$ depends on which of these terms dominates the other, which in turn depends on $w'(1 - F(q|\bar{e}))$. When q decreases, $w'(1 - F(q|\bar{e}))$ increases; the second term becomes larger and ${t_P^{sb}}'(q)$ is more likely to be positive. The opposite happens when q increases and $w'(1 - F(q|\bar{e}))$ decreases.

Again, we study (28) at the extremes, but this time 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 (28) goes to $-\infty$ while the second goes to 0. Therefore, $\lim_{q\to\bar{q}} {t_P^{sb}}'(q) = -\infty$.

We ask whether t_P^{sb} ever increases with output; that is, whether ${t_P^{sb}}'(q) > 0$ or equivalently, using (28),

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

We use

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

to rewrite the assumed inequality as

$$\frac{f(q|\underline{e})}{f(q|\bar{e})}\left[\frac{w''\Big(1-F(q|\bar{e})\Big)}{w'\Big(1-F(q|\bar{e})\Big)}f(q|\bar{e})\left(\frac{1}{\mu u'(t^{sb}(q))w'\Big(1-F(q|\underline{e})\Big)\frac{f(q|\underline{e})}{f(q|\bar{e})}}+1\right)\right.$$
$$\left.-\frac{w''\Big(1-F(q|\underline{e})\Big)}{w'\Big(1-F(q|\underline{e})\Big)}f(q|\underline{e})\right]\leq-\frac{\mathrm{d}}{\mathrm{d}q}\frac{f(q|\underline{e})}{f(q|\bar{e})}. \quad (29)$$

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

$$\frac{1}{\mu u'(t^{sb}(q))w'\Big(1-F(q|\underline{e})\Big)\frac{f(q|\underline{e})}{f(q|\bar{e})}} \quad (30)$$

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 ${t_P^{sb}}'(q)<0$ on some part of the output interval. To that end, find $q_{\mathcal{I}}$ that satisfies:

$$\int_{q_{\mathcal{I}}}^{q_h} t^{sb}(q)\,\mathrm{d}q-\int_{q_h}^{\bar{q}} t^{sb}(q)\,\mathrm{d}q=0. \quad (31)$$

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 $q_{\mathcal{I}}\in[\underline{q},q_h)$ ensures (31) 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. ■

Corollary 3

Proof. **Part 1.** Denote by i an agent who is more optimistic than agent j. According to Definition 3, $w_i(p) = \theta\Big(w_j(p)\Big)$ for all $p \in [0,1]$.

From (27) we know that the critical cost level for agent j is

$$\hat{c}_{O,j} := \int\limits_{\underline{q}}^{\bar{q}} u'\Big(t^{fb}_{O,j}(q)\Big) t^{fb}_{O,j}{}'(q) \Big(w_j\Big(1 - F(q|\bar{e})\Big) - w_j\Big(1 - F(q|\underline{e})\Big)\Big)\,\mathrm{d}q. \quad (32)$$

If agent i were given the same contract as that for j, then

$$\hat{c}_{O,i} := \int\limits_{\underline{q}}^{\bar{q}} u'\Big(t^{fb}_{O,j}(q)\Big) t^{fb}_{O,j}{}'(q) \Big(w_i\Big(1 - F(q|\bar{e})\Big) - w_i\Big(1 - F(q|\underline{e})\Big)\Big)\,\mathrm{d}q. \quad (33)$$

For $q < q^*$ we have $w_i'\Big(1 - F(q|e)\Big) < 1$. Let $w_0 := w_j\Big(1 - F(q|\underline{e})\Big)$ and $w_1 := w_j\Big(1 - F(q|\bar{e})\Big)$. Integrating $w_i'\Big(1 - F(q|e)\Big) = \theta'\Big(w_j\Big(1 - F(q|e)\Big)\Big)w'\Big(1 - F(q|e)\Big) < 1$ over $[w_0, w_1]$ gives

$$\int_{w_0}^{w_1} \theta'(s)\mathrm{d}s < \int_{w_0}^{w_1} \mathrm{d}s \Leftrightarrow \quad (34)$$

$$w_i\Big(1 - F(q|\bar{e})\Big) - w_i\Big(1 - F(q|\underline{e})\Big) < w_j\Big(1 - F(q|\bar{e})\Big) - w_j\Big(1 - F(q|\underline{e})\Big). \quad (35)$$

Equation (34) together with (32) and (33) show that $\hat{c}_{O,i} < \hat{c}_{O,j}$. If $q < q^*$, contract $t^{fb}_{O,i}(q)$ is different than $t^{fb}_{O,j}(q)$ as it specifies a lower threshold $\hat{c}_{O,i}$.

Part 2. The third part of Lemma 7 shows that the point p_l at which $w_i'(p_l) = w_j'(p_l)$ 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 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^ε 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. Hence, segment $q > q^\varepsilon$ for which $w_i'\Big(1 - F(q|e)\Big) < \varepsilon$ becomes larger.

Eq. (30) shows that this tendency of $w_i'(p)$ becoming smaller as i becomes more pessimistic, makes it more likely that $\frac{\mathrm{d}t^{sb}(q)}{\mathrm{d}q} < 0$ takes place in a larger segment of q. Consequently, ironing, if still possible, requires a smaller value $q_{\mathcal{I}}$.

■

Corollary 5

Proof. We start by rewriting (25) 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

$$\nu = \int_{\underline{q}}^{\bar{q}} \frac{1}{u'(t^{sb}(q))} f(q|\bar{e})\,\mathrm{d}q = \mathbb{E}_{\bar{e}}\left(\frac{1}{u'(t^{sb}(q))}\right) \tag{36}$$

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 (36) into (25) and multiplying by $u(t^{sb}(q))$, we obtain

$$\begin{aligned}&\mu u(t^{sb}(q))\Big[w'\Big(1-F(q|\bar{e})\Big)f(q|\bar{e})-w'\Big(1-F(q|\underline{e})\Big)f(q|\underline{e})\Big]\\&=f(q|\bar{e})u(t^{sb}(q))\left[\frac{1}{u'(t^{sb}(q))}-\mathbb{E}_{\bar{e}}\left(\frac{1}{u'(t^{sb}(q))}\right)w'\Big(1-F(q|\bar{e})\Big)\right].\end{aligned}\tag{37}$$

From the complementary slackness condition associated with μ we know

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

We can thus rewrite (37), after integrating with respect to q over $[\underline{q},\bar{q}]$, as

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

where $\tilde{\mathbb{E}}_{\bar{e}}$ is the expectation as perceived by an agent who distorts probabilities. Since $\tilde{\mathbb{E}}_{\bar{e}}\left(u(t^{sb}(q))\right)<\mathbb{E}_{\bar{e}}\left(u(t^{sb}(q))\right)$ under pessimism, and the opposite under optimism, (38) implies $\mu_P>\mu_{EU}>\mu_O\geq 0$.

We rewrite (25), the first-order condition, again, but this time as

$$\frac{1}{u'(t^{sb}_{nEU}(q))}=\nu w'\left(1-F(q|\bar{e})\right)+\\\mu_{nEU}w'\left(1-F(q|\bar{e})\right)\left(1-\frac{w'\left(1-F(q|\underline{e})\right)f(q|\underline{e})}{w'\left(1-F(q|\bar{e})\right)f(q|\bar{e})}\right),\tag{39}$$

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 $w' = 1$, (39) simplifies to

$$\frac{1}{u'(t^{sb}_{EU}(q))} = \nu + \mu_{EU}\left(1 - \frac{f(q|\underline{e})}{f(q|\bar{e})}\right). \tag{40}$$

Comparison of (39) and (40) 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, under certain conditions, hold 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, (39) and (40) 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 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 $t_L^{fb}(q)$. We have that $\nu > 0$ since $u' > 0$ and $w' > 0$. Moreover, (7) shows that, for $q \in [\underline{q}, \tilde{q}]$ where $w''(q) < 0$, $t_L^{fb'}(q) < 0$; and for $q \in [\tilde{q}, \bar{q}]$ where $w''(p) > 0$, $t_L^{fb'}(q) < 0$.

We look at how $t_L^{fb}(q)$ behaves at extremes. That is as q either approaches $\underline{q}$, $\bar{q}$ or $\tilde{q}$. 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 $\lim_{q \to \underline{q}} t_L^{fb'}(q) = +\infty$. Lemma 2 implies $\lim_{q \to \bar{q}} \frac{w''(1-F(q|\bar{e}))}{w'(1-F(q|\bar{e}))} = +\infty$, so $\lim_{q \to \bar{q}} t_L^{fb'}(q) = +\infty$. Finally, 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 $\lim_{q \to \bar{q}} t_L^{fb'}(q) = 0$.

$t_L^{fb'}(q) < 0$ 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:

$$\int_{\underline{q}}^{\tilde{q}} t_L^{fb}(q)\,\mathrm{d}q - \int_{\tilde{q}}^{q_I} t_L^{fb}(q)\,\mathrm{d}q = 0 \tag{41}$$

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

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

Letting $q_I := q_E$ would lead to non-monotonicities in the payment scheme. To fix that, note that continuity of $[\underline{q}, \bar{q}]$, u', w', and w'' imply that there exists $\varepsilon > 0$ such that $\tilde{q} + \varepsilon \leq \bar{q}$ and

$$t_L^{fb}(q_E - \delta) = \frac{\int_{\underline{q}}^{\tilde{q}+\varepsilon} t_L^{fb}(q)\,\mathrm{d}q}{\tilde{q} + \varepsilon - \underline{q}}, \tag{42}$$

for a $\delta > 0$ such that $q_E - \delta \geq \tilde{q}$. The tuple (δ, ε) can be adjusted to obtain $(\hat{\delta}, \hat{\varepsilon})$ such that (42) 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

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

■

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 $w_i(p) = \phi\big(w_j(p)\big)$,

$$\frac{w_i''(p)}{w_i'(p)} = \frac{\phi''\big(w_j(p)\big)}{\phi'\big(w_j(p)\big)} w_j'(p) + \frac{w_j''(p)}{w_j'(p)}. \tag{43}$$

Since $\phi'' < 0$ 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 the above equation 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, $[0, p_1]$, 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 $w_i'(p)$ and $w_j'(p)$ intersect at some point in $p < \tilde{p}$.

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 $\mathrm{w}_J(p) := \eta\big(w_j(p)\big)$ 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 the Lemma. Therefore, $w'_J(p)$ tends to infinity faster than $w_j'(p)$

as $p \to 0^+$. Moreover, due to the continuity of $w'(p)$, $w'(p)$ being decreasing in p in $p < \tilde{p}$, and the fact that $\lim_{p \to \tilde{p}} w'(p) = \min\{w'(p)\}$, 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 $-\frac{w''_i(p)}{w'_i(p)} > -\frac{w''_J(p)}{w'_J(p)}$ for all $p < \tilde{p}$ and $w'_i(p)$ tends to infinity faster than $w'_J(p)$ as $p \to 0^+$. 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 now that $w'_i(p) < w'_j(p)$ for all $p > \tilde{p}$. 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 $p_1 \in (\tilde{p}, p)$ corroborating the first part of the Lemma. We also have that $\int_{\tilde{p}}^{p_1} w'_i(p)\mathrm{d}p > \int_{\tilde{p}}^{p_1} w'_i(p)\mathrm{d}p \Leftrightarrow w_i(p_1) < w_j(p_1)$, contradicting 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 $w'_i(p)$ and $w'_j(p)$ intersect at some point in $p > \tilde{p}$.

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 $\mathrm{w}_J(p) := \eta\big(w_j(p)\big)$ 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 the Lemma. Therefore, $w'_j(p)$ tends to infinity faster than $w_J(p)$ as $p \to 1^-$. Moreover, due to the continuity of $w'(p)$, $w'(p)$ being increasing in p in $p > \tilde{p}$, and the fact that $\lim_{p \to \tilde{p}} w'(p) = \min\{w'(p)\}$, 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 $p > p_m > \tilde{p}$. Let $w_i := \phi\big(w_J(p)\big)$ where ϕ is a subadditive function. Then $\frac{w''_i(p)}{w'_i(p)} > \frac{w''_J(p)}{w'_J(p)}$ for all $p > \tilde{p}$ and $w'_J(p)$ tends to infinity faster than $w'_i(p)$ as $p \to 1^+$. 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 the interval $q \in [\underline{q}, \tilde{q}]$ as:

$$u\Big(t^{fb}_{L,j}(\underline{q})\Big)-u\Big(t^{fb}_{L,j}(\tilde{q})\Big)w_j\Big(1-F(\tilde{q}|\bar{e})\Big)+\int_{\underline{q}}^{\tilde{q}} u'(t^{fb}_{L,j}(q))\frac{\mathrm{d}t^{fb}_L(q)}{\mathrm{d}q}w_j\Big(1-F(q|\bar{e})\Big)\mathrm{d}q. \tag{44}$$

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

$$u\Big(t^{fb}_{L,j}(\underline{q})\Big)-u\Big(t^{fb}_{L,j}(\tilde{q})\Big)w_i\Big(1-F(\tilde{q}|\bar{e})\Big)+\int_{\underline{q}}^{\tilde{q}} u'(t^{fb}_{L,j}(q))\frac{\mathrm{d}t^{fb}_L(q)}{\mathrm{d}q}w_i\Big(1-F(q|\bar{e})\Big)\mathrm{d}q. \tag{45}$$

Subtracting (44) from (45) gives:

$$-u\Big(t^{fb}_{L,j}(\tilde{q})\Big)\Big(w_i\Big(1-F(q|\bar{e})\Big)-w_j\Big(1-F(q|\bar{e})\Big)\Big)+ \\ \int_{\underline{q}}^{\tilde{q}} u'(t^{fb}_{L,j}(q))\frac{\mathrm{d}t^{fb}_{L,j}}{\mathrm{d}q}\Big(w_i\Big(1-F(q|\bar{e})\Big)-w_j\Big(1-F(q|\bar{e})\Big)\Big)\mathrm{d}q. \tag{46}$$

Since $\frac{\mathrm{d}t^{fb}_{L,j}}{\mathrm{d}q} < 0$ if $q < \tilde{q}$ and because stronger likelihood insensitivity implies $w_j\Big(1-F(q|\bar{e})\Big) > w_i\Big(1-F(q|\bar{e})\Big)$ in $q < \tilde{q}$ (Lemma 10), Eq. (46) is positive. Therefore,

$$\int_{\underline{q}}^{\tilde{q}} u(t^{fb}_{L,j})w'_i\Big(1-F(q|\bar{e})\Big)f(q|e)\mathrm{d}q > \frac{\bar{U}}{2}, \tag{47}$$

where $\frac{\bar{U}}{2}$ is due to symmetry of t^{fb} around $\hat{p}$ which is in turn generated by the symmetry of $w(p)$ around that point, i.e. that $\tilde{p} = \hat{p} = \frac{1}{2}$.

Since the participation constraint binds at the optimum, then it must be that contract $t^{fb}_{L,i}(q)$ ensures $\int_{\underline{q}}^{\tilde{q}} u(t^{fb}_{L,i}(q))w'_i\Big(1-F(q|\bar{e})\Big)f(q|e)\mathrm{d}q = \frac{\bar{U}}{2}$.

Consequently,

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

The above relationship is implied by $t_{L,j}^{fb} > t_{L,i}^{fb}$ in $q \in [\underline{q}, \tilde{q}]$. So, ironing the solution 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. (46) for that interval is

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

Due to $\frac{\mathrm{d}t_{L,j}^{fb}}{\mathrm{d}q} > 0$ if $q > \tilde{q}$ and because stronger likelihood insensitivity implies $w_j\Big(1-F(q|\bar{e})\Big) < w_i\Big(1-F(q|\bar{e})\Big)$ (Lemma 10), Eq. (49) is positive. Therefore,

$$\int_{\tilde{q}}^{\bar{q}} u(t_{L,j}^{fb})w_i'\Big(1-F(q|\bar{e})\Big)f(q|e)\mathrm{d}q > \frac{\bar{U}}{2}, \tag{50}$$

Since the participation constraint binds at the optimum, then it must be that contract $t_{L,i}^{fb}(q)$ generates $\int_{\tilde{q}}^{\bar{q}} u(t_{L,i}^{fb}(q))w_i'\Big(1-F(q|\bar{e})\Big)f(q|e)\mathrm{d}q = \frac{\bar{U}}{2}$. Thus,

$$\int_{\tilde{q}}^{\bar{q}} u(t_{L,j}^{fb}(q))w_i'\Big(1-F(q|\bar{e})\Big)f(q|e)\mathrm{d}q > \int_{\tilde{q}}^{\bar{q}} u(t_{L,i}^{fb}(q))w_i'\Big(1-F(q|\bar{e})\Big)f(q|e)\mathrm{d}q. \tag{51}$$

Consequently, $t_{L,j}^{fb} > t_{L,i}^{fb}$ in $q \in [\tilde{q}, \bar{q}]$. Since ironing the first-order condition in $q \in [\underline{q}, \tilde{q}]$ generates $\tilde{t}_{L,j}^{fb} > \tilde{t}_{L,i}^{fb}$ and since $t_{L,j}^{fb} > t_{L,i}^{fb}$ in $q \in [\tilde{q}, \bar{q}]$ it must be that the point q_I, at which the ironed solution meets the solution from the first-order condition, is higher 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 (25) solves the maximization problem. Denote by $t_L^{sb}(q)$ the contract that satisfies that equation.

Incentive constraint is binding. We first show that $\mu > 0$ might not always hold at the optimum. Suppose indeed that $\mu = 0$. Then $t_L^{sb}(q) = t_L^{fb}(q)$, $t_L^{fb}(q)$ being the first-best contract presented in Proposition 4.

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

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_L^{fb}(q)\Big) w'(1 - F(q|\bar{e})) f(q|\bar{e}) \, \mathrm{d}q - c > \int_{\underline{q}}^{\bar{q}} u\Big(t_L^{fb}(q)\Big) w'(1 - F(q|\underline{e})) f(q|\underline{e}) \, \mathrm{d}q. \quad (52)$$

Using integration by parts we rewrite (52) as

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

Assumption 4 implies $w\left(1 - F(q|\bar{e})\right) - w\left(1 - F(q|\underline{e})\right) \geq 0$ which, together with $\frac{\mathrm{d}t_O^{fb}(q)}{\mathrm{d}q} > 0$ (Proposition 4) and $u'(t) > 0$ (Assumption 1), imply that the left-hand side of (27) is weakly positive. Since $w(p)$ and $u(t)$ are $\mathcal{C}^2$, and since c is a constant unbounded from above, there exists $\hat{c}_L > 0$ such that, for a given $t_L^{fb}(q)$,

- if $c \leq \hat{c}_L$, (53) holds: $\mu = 0$ and $t_L^{sb}(q) = t_L^{fb}(q)$; on the other hand,
- if $c > \hat{c}_L$, (53) does not hold: $\mu > 0$ and $t_L^{sb}(q)$ satisfies (25).

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 (28), which presents the derivative of $t^{sb}(q)$ with respect to q. Let $\tilde{q} \in (\underline{q}, \bar{q})$ such that $w\Big(1 - F(\tilde{q}|\bar{e})\Big) = 1 - F(\tilde{q}|\bar{e})$. When the agent exhibits likelihood insensitivity (Definition 4), $w'' < 0$ for all $q \in (\tilde{q}, \bar{q}]$. For those output levels the two terms on the right-hand side of (28) are positive. Hence, $\frac{\mathrm{d}t_L^{sb}(q)}{\mathrm{d}q} > 0$.

To further understand the shape of $t_O^{sb}(q)$ in $q \in (\tilde{q}, \bar{q}]$ we study (28) at 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 (28), $\lim_{q \to \bar{q}} \frac{\mathrm{d}t_L^{sb}(q)}{\mathrm{d}q} = +\infty$. 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.

Furthermore, 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 (28) is negative while the second one 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 (28) larger, and $\frac{\mathrm{d}t_P^{sb}(q)}{\mathrm{d}q}$ is more likely to be positive. The opposite happens when q increases.

Again, we study (28) 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 $\frac{\mathrm{d}t_L^{sb}(q)}{\mathrm{d}q} < 0$ in that interval.

The third part of Lemma 10 shows that for an agent i who is more likelihood insensitive than agent j, the point p_l at which $w_i'(p_l) = w_j'(p_l)$ in $p < \tilde{p}$ becomes smaller the more likelihood insensitive i is with respect to j. Also, the fourth part of that Lemma shows that the point p_n at which $w_i'(p_n) = w_j'(p_n)$ in $p > \tilde{p}$ becomes larger the more likelihood insensitive i is with respect to j. Hence, the output level q^ε such that $w_i'\Big(1 - F(q^\varepsilon|e)\Big) = \varepsilon$ for arbitrary small $\varepsilon > 0$ takes place at a lower and higher output levels the more likelihood insensitive i is. Hence, the segment $q > q^\varepsilon$ for which $w_i'\Big(1 - F(q|e)\Big) < \varepsilon$ becomes larger. Thus, stronger likelihood insensitivity implies that $w'(1 - F(q|\bar{e})$ becomes smaller for output levels away of $\underline{q}$. This property makes(30) larger, and consequently that $\frac{\mathrm{d}t_L^{sb}(q)}{\mathrm{d}q} < 0$ is more likely for intermediate values in $q \in (\tilde{q}, \bar{q}]$.

Ironing That $\frac{\mathrm{d}t_L^{sb}(q)}{\mathrm{d}q} < 0$ in $q \in [\underline{q}, \tilde{q}]$ takes place 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

$$\int_{q_{\mathcal{I}1}}^{q_{\mathcal{M}}} t_L^{sb}(q)\mathrm{d}q - \int_{q_{\mathcal{M}}}^{q_{\mathcal{S}}} t_L^{sb}(q)\mathrm{d}q + \int_{q_{\mathcal{S}}}^{q_{\mathcal{I}2}} t_L^{sb}(q)\mathrm{d}q = 0, \tag{54}$$

and

$$t_L^{sb}(q_{\mathcal{I}1}) = t_L^{sb}(q_{\mathcal{I}2}). \tag{55}$$

There are four cases worth discussion. 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:

$$\tilde{t}_L^{sb}(q) = \begin{cases} t_L^{sb}(q) & \text{if } q \in [\underline{q}, q_{\mathcal{I}1}) \cup (q_{\mathcal{I}2}, \bar{q}], \\ t_L^{sb}(q_{\mathcal{I}1}) & \text{if } q \in [q_{\mathcal{I}1}, q_{\mathcal{I}2}]. \end{cases} \tag{56}$$

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:

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

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:

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

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 sizeable 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. Accordingly, the third and fourth part of Lemma 10 show that the points $p_l \in (0, \tilde{p})$ and $p_m \in (\tilde{p}, 1)$ such that $w_i'(p_l) = w_j'(p_l)$ and $w_i'(p_n) = w_j'(p_n)$ become smaller and large, respectively, the more likelihood insensitive i becomes. Hence, the output level q^ε such that $w_i'\Big(1 - F(q^\varepsilon|e)\Big) = \varepsilon$ for arbitrary small $\varepsilon > 0$ takes place at a lower and higher output levels the more likelihood insensitive i is. Hence, the segment $q > q^\varepsilon$ for which $w_i'\Big(1 - F(q|e)\Big) < \varepsilon$ becomes larger.

Therefore, the expression in Eq. (30) becomes larger for agent i, and consequently $\frac{\mathrm{d}t_L^{sb}(q)}{\mathrm{d}q} < 0$ is more likely for him than for j. Ironing is applied in a larger interval.

■

Corollary 8

Proof. According to (36), the value of $\nu > 0$ is the same across agents with different probability weighting. Moreover, using (38) it can be established that the Lagrangian multiplier of the incentive compatibility multiplier, μ^L>0, exhibits $\mu^L > \mu^{EU}$ in $q > q^{**}$ and $q < q^*$ and $\mu^L < \mu^{EU}$ otherwise.

The first-order condition presented in equation (25) is rewritten to obtain:

$$\begin{aligned}\frac{1}{u'(t_L^{sb}(q))} =& \nu w'\Big(1 - F(q|\bar{e})\Big) + \\ & \mu^L w'\Big(1 - F(q|\bar{e})\Big)\left(1 - \frac{w'\left(1 - F(q|\underline{e})\right) f(q|\underline{e})}{w'\left(1 - F(q|\bar{e})\right) f(q|\bar{e})}\right),\end{aligned} \tag{59}$$

where t_L^{sb} the second-best contract (Proposition 5 (2) and (3)). For the EU agent, the optimal contract satisfies the first-condition presented in (40).

Comparison of (59) and (40) 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 $q \in [\tilde{q}, q^{**}]$.

The inequality $t_L^{sb}(q) < t_{EU}^{sb}(q)$ cannot hold everywhere in $q \in [q^*, \tilde{q}]$. To see how, let q approach q^*. In that case, $\lim_{q \to q^*} w'\Big(1 - F(q|\bar{e})\Big) = 1$ and (59) is similar to (40) with the exception that $\mu^L > \mu^{EU}$. Hence, $t_L^{sb}(q^*) > t_{EU}^{sb}(q^*)$. That inequality continues to hold for higher output levels in $q \in (q^*, \tilde{q}]$ 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 will make $w'\Big(1 - F(q|\bar{e})\Big)$ small in that segment and generate $w'\Big(1 - F(q|\bar{e})\Big) > w'\Big(1 - F(q|\underline{e})\Big)$, both of which might outweight the fact that $\mu^{EU} < \mu^L$.

Comparison of (59) and (40) 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 $t_L^{sb}(q) > t_{EU}^{sb}(q)$ 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 $\lim_{q \to \bar{q}} w'\Big(1 - F(q|\bar{e}\Big) = +\infty$, (39) and (40) demonstrate that $t_L^{sb}(\bar{q}) > t_{EU}^{sb}(\bar{q})$. To guarantee that inequality 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 $\mu^{EU} > \mu^L$. This happens with sufficiently large levels of likelihood insensitivity.

■

Proposition 6

Proof. Rewrite Eq. (4) using Assumption 6 as

$$CPT(t, e, r) = \int_{\underline{q}}^{\bar{q}} \theta v\left(t(q) - r\right) w'\Big(1 - F(q|e)\Big) - \lambda\Big(1 - \theta\Big) v(r - t(q)) w'\Big(1 - F(q|e)\Big) f(q|e)\, \mathrm{d}q - c(e), \quad (60)$$

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

$$\begin{aligned}\mathcal{L}(q,t) =&\Big(S(q)-t(q)\Big)f(q|\bar{e})\\&+\mu\Bigg(-\lambda v\Big(r-t(q)\Big)\Big(w'\Big(1-F(q|\bar{e})\Big)f(q|\bar{e})-w'\Big(1-F(q|\underline{e})\Big)f(q|\underline{e})\Big)-c\Bigg)\\&+\nu\Bigg(-\lambda v\Big(r-t(q)\Big)w'\Big(1-F(q|\bar{e})\Big)f(q|\bar{e})-c-\bar{U}\Bigg).\end{aligned} \tag{61}$$

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

$$\frac{1}{\lambda v'\Big(r-t\Big)\Big(w'\Big(1-F(q|\bar{e})\Big)\Big)}=\nu+\mu\left(1-\frac{w'\Big(1-F(q|\underline{e})\Big)f(q|\underline{e})}{w'\Big(1-F(q|\bar{e})\Big)f(q|\bar{e})}\right). \tag{62}$$

Denote by $t_C^{sb}(q)$ the transfer satisfying Eq. (62). 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

$$\lambda v\Big(r-t_C^{sb}(q)\Big)=\lambda\Big(1-w(\rho)\Big)v(r). \tag{63}$$

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

$$\lambda v\Big(r-t_C^{sb}(q)\Big)\le\lambda v\Big((1-w(\rho))r\Big). \tag{64}$$

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

$$CPT(L_\rho, \bar{e}, r) = -\left(1 - w\left(\frac{\bar{L}}{r}\right)\right)\lambda v(r) - c \tag{65}$$

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 (65) with respect to ρ, which gives us

$$w'(\rho)\lambda v(r) = 0. \tag{66}$$

Denote by ρ^{opt} the probability satisfying the condition in (66). The second-order condition evaluated at ρ^{opt} is

$$w''(\rho^{opt})\lambda v(r). \tag{67}$$

Hence, $\rho^{opt} \in (0, 1)$ whenever $w'' < 0$. That is onder 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 $p \in (0, 1). Since CPT(L_{\rho=1}, \bar{e}, r) = -c > -\lambda v(r) - c = CPT(L_{\rho=0}, \bar{e}, r),$

$$\tag{68}$$

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

$$\mathcal{L}(q,t) = \Big(S(q) - t(q)\Big) f(q|\bar{e}) \\ + \mu\Bigg(v\Big(t(q) - r\Big)\Big(w'\Big(1 - F(q|\bar{e})\Big)f(q|\bar{e}) - w'\Big(1 - F(q|\underline{e})\Big)f(q|\underline{e})\Big) - c\Bigg) \\ + \nu\Bigg(v\Big(t(q) - r\Big)w'\Big(1 - F(q|\bar{e})\Big)f(q|\bar{e}) - c - \bar{U}\Bigg). \tag{69}$$

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

$$\frac{1}{v'\Big(t-r\Big)w'\Big(1-F(q|\bar{e})\Big)} = \nu + \mu\left(1 - \frac{w'\Big(1-F(q|\underline{e})\Big)f(q|\underline{e})}{w'\Big(1-F(q|\bar{e})\Big)f(q|\bar{e})}\right). \tag{70}$$

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 5, 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 5), denote by $\hat{q} \in [\underline{q}, \bar{q}]$ the performance level satisfying:

$$\frac{1}{\frac{\lambda v(r)}{r}} = \nu + \mu\left(1 - \frac{w'\Big(1-F(\hat{q}|\underline{e})\Big)f(\hat{q}|\underline{e})}{w'\Big(1-F(\hat{q}|\bar{e})\Big)f(\hat{q}|\bar{e})}\right). \tag{71}$$

Where the left-hand side of (71) 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. (71) 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 (71) 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 5, depending on the shape of w. Finally, if $\hat{q} \in [\underline{q}, \bar{q}]$ then

$$t_C^{sb}(q) = \begin{cases} r & \text{if } q < \hat{q}, \\ t_P^{sb}(q), t_O^{sb}(q) \text{ (Proposition 3), or } t_L^{sb}(q) \text{ (Proposition 5)} & \text{if } q \geq \hat{q}. \end{cases} \tag{72}$$

■

Proposition 7

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

$$\int_{\underline{q}}^{\bar{q}} u\big(t_{EU}(q)\big) f(q|\bar{e})\,\mathrm{d}q - c \geq \int_{\underline{q}}^{\bar{q}} u\big(t_{EU}(q)\big) f(q|\underline{e})\,\mathrm{d}q, \tag{73}$$

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

$$\int_{\underline{q}}^{\bar{q}} u\big(t_L(q)\big) w'\big(1 - F(q|\bar{e})\big) f(q|\bar{e})\,\mathrm{d}q - c \geq \int_{\underline{q}}^{\bar{q}} u\big(t_L(q)\big) w'\big(1 - F(q|\underline{e})\big) f(q|\underline{e})\,\mathrm{d}q. \tag{74}$$

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:

$$\int_{\underline{q}}^{\bar{q}} u\big(t_{EU}(q)\big) f(q|\bar{e})\,\mathrm{d}q - c \geq \max_{e \in \{\underline{e}, \bar{e}\}} \left\{ \int_{\underline{q}}^{\bar{q}} u\big(t_L(q)\big) f(q|\bar{e})\,\mathrm{d}q - c(e) \right\}, \tag{75}$$

and for the non-EUT agent:

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big)w'\Big(1 - F(q|\bar{e})\Big)f(q|\bar{e})\,\mathrm{d}q - c$$
$$\geq \max_{e\in\{\underline{e},\bar{e}\}} \left\{ \int_{\underline{q}}^{\bar{q}} u\Big(t_{EU}(q)\Big)w'\Big(1 - F(q|\bar{e})\Big)f(q|\bar{e})\,\mathrm{d}q - c(e) \right\}. \tag{76}$$

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

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_{EU}(q)\Big)f(q|\bar{e})\,\mathrm{d}q - c \geq 0, \tag{77}$$

and

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big)w'\Big(1 - F(q|\bar{e})\Big)f(q|\bar{e})\,\mathrm{d}q - c \geq 0. \tag{78}$$

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:

$$\int_{\underline{q}}^{\bar{q}} w'\left(1 - F(q|\bar{e})\right) f(q|\bar{e})\mathrm{d}q - \int_{\underline{q}}^{\bar{q}} w'\left(1 - F(q|\underline{e})\right) f(q|\underline{e})\mathrm{d}q =$$
$$w\left(1 - F(q|\bar{e})\right) - w\left(1 - F(q|\underline{e})\right). \tag{79}$$

Instead, for the EU agent, efficiency amounts to:

$$\int_{\underline{q}}^{\bar{q}} f(q|\bar{e})\mathrm{d}q - \int_{\underline{q}}^{\bar{q}} f(q|\underline{e})\mathrm{d}q = \left(1 - F(q|\bar{e})\right) - \left(1 - F(q|\underline{e})\right). \tag{80}$$

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 (79) to be larger than (80) is $w'(1-F(q|e)) > 1$ for any e. That is because

$$\int_{1-F(q|\underline{e})}^{1-F(q|\bar{e})} w'(s)\mathrm{d}s > \int_{1-F(q|\underline{e})}^{1-F(q|\bar{e})} \mathrm{d}s \Leftrightarrow$$
$$w\left(1-F(q|\bar{e})\right) - w\left(1-F(q|\underline{e})\right) > F(q|\underline{e}) - F(q|\bar{e}) \tag{81}$$

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 (77) and (76) immediately imply (78). Hence, at the optimum the participation constraint in (77) binds, while the participation constraint in (78) slacks.

From equation (75) and the constraint in (77), which binds at the optimum, we obtain:

$$0 \geq \max_{e\in\{\underline{e},\bar{e}\}} \left\{ \int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) f(q|\bar{e})\,\mathrm{d}q - c(e) \right\}, \tag{82}$$

which implies that EUT agents cannot afford to mimic non-EUT agents. Hence, the relevant adverse selection constrain is that in (76), which states that the non-EUT agent derives rents from mimicking the EUT agent. In contrast, equation (75) slacks at the optimum.

A direct implication that (76) binds is $t_L(q) \geq t_{EU}(q)$, which in turn gives

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_{EU}(q)\Big) f(q|\bar{e})\,\mathrm{d}q - c > \int_{\underline{q}}^{\bar{q}} u\Big(t_{EU}(q)\Big) f(q|\underline{e})\,\mathrm{d}q. \tag{83}$$

Hence, the moral hazard constraint in (73) slacks at the optimum.

Next, from the inequality in (78), which slacks at the optimum, along with equation (82), which holds with strict inequality, we obtain:

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})\,\mathrm{d}q - c > 0 \geq \max_{e \in \{\underline{e}, \bar{e}\}} \left\{ \int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) f(q|\bar{e})\,\mathrm{d}q - c(e) \right\}. \tag{84}$$

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

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) f(q|\underline{e})\,\mathrm{d}q > \int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) w'\left(1 - F(q|\underline{e})\right) f(q|\underline{e})\,\mathrm{d}q, \tag{85}$$

Equations (84) and (85) imply

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})\,\mathrm{d}q - c > \int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) w'\left(1 - F(q|\underline{e})\right) f(q|\underline{e})\,\mathrm{d}q. \tag{86}$$

and equation (74) is implied by other constraints in the principal's program.

Hence, at the solution only equations (76) and (77) 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 (77). 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 (76).

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.

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

Barseghyan, Levon, Francesca Molinari, and Matthew Thirkettle, "Discrete choice under risk with limited consideration," *American Economic Review*, 2021, *111* (6), 1972–2006.

_ , _ , 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.

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

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

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

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

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

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.

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.

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.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

■