# When Not Paying for Performance is Optimal: Probability Weighting and Moral Hazard *

Víctor González-Jiménez†

August 14, 2025

**Abstract**

I investigate the optimal design of contracts when agents exhibit probability weighting. Integrating probability weighting into a standard moral hazard framework reveals that optimal contracts should include substantial performance-insensitive segments. This occurs because these agents overweight the probabilities of extreme outcomes— the highest or lowest performance levels— and tend to be overly sensitive to incentives linked to those outcomes, while displaying limited sensitivity to incentives implemented elsewhere. As a result, contracts that do not vary with performance over a broad range of performance levels become optimal. Empirical evidence from a representative sample of American households supports the model's prediction: individuals with stronger probability weighting are less likely to be compensated based on performance. This paper provides a novel explanation for the widespread use of simple and performance-insensitive contracts by organizations.

**JEL Classification** : D86, D82, J41, M52, M12.
**Keywords:** Contract Theory, Probability Weighting, Incentives, Rank-Dependent Utility.

*This paper supersedes an earlier version, entitled "Optimal Incentives without Expected Utility." I would like to thank Aurelien Baillon, Geoffrey Castillo, Paul Heidhues, Mats Köster, Daniel Garcia, Gyula Seres, and Cole Williams for their valuable comments. I also would like to thank participants in seminars at Erasmus University, Düsseldorf Institute for Competition Economics, University of Amsterdam, and the University of Vienna.

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

# 1 Introduction

The moral hazard framework in contract theory studies how a principal can design incentives to motivate an agent whose actions are unobservable. The contracts it predicts, however, often diverge from those observed in practice (Lazear and Oyer, 2007; Prendergast, 1999; Salanié, 2003). Notably, the bulk of the literature examines risk attitudes based on expected utility, which, while theoretically appealing, is not an accurate description of choice under uncertainty (Starmer, 2000). In what follows, I investigate whether relaxing the assumption of expected utility maximization improves the type of contract predicted by the theory.

In particular, I consider agents who exhibit *probability weighting*, a phenomenon supported by abundant evidence from behavioral economics (Abdellaoui et al., 2011; Bruhin et al., 2010; Fehr-Duda and Epper, 2011; Kahneman and Tversky, 1979; l'Haridon and Vieider, 2019; Tversky and Kahneman, 1992).[1] These agents deviate from expected utility because their preferences between risky alternatives are not linear in probabilities (Abdellaoui, 2000). An assumption that underlies the most prominent alternative models of decision making under risk and uncertainty, such as rank-dependent utility (Quiggin, 1982) and cumulative prospect theory (Tversky and Kahneman, 1992). I incorporate these models within the moral hazard framework in contract theory, thus bridging the gap between the two literatures.

The main result of this paper is that optimal contracts for agents who exhibit probability weighting feature substantial performance-insensitive segments. Therefore, it proposes an explanation for the prevalence of simple and low-powered contracts observed in practice, such as salaries, lump-sum bonuses, and commissions awarded only for reaching exceptionally high performance levels. According to the framework of this paper, such simple contracts emerge because the principal anticipates that probability weighting will attenuate the agent's responsiveness to incentives, thereby reducing the effectiveness of performance-based pay.

I first examine the design of contracts for agents who display probability weighting due to pessimism or optimism. These agents pay disproportionate attention to unfavorable outcomes (low performance levels), in the case of pessimism, and favorable ones (high performance levels), in the case of optimism. As a result, they overweight the probabilities of these outcomes and underweight all others. The principal reacts to this pattern of probability weighting by offering a contract that concentrates incentives at the performance levels whose probabilities the agent overweights.

In the specific case of a pessimistic agent, the optimal contract concentrates incentives at low performance levels. Moreover, since this agent underweights the probabilities of all other outcomes, that contract must provide a fixed and high pay-

[1] See Wakker (2010, p. 204) for an extensive list of references.

ment for those realizations. This results in a performance-insensitive segment at the upper end of the performance set. According to this incentive scheme, commissions have a cap because offering incentives beyond a certain threshold is ineffective; the agent does not respond to them.

Importantly, I also find that the stronger the agent's pessimism is, the larger will be the performance-insensitive segment included by the optimal contract. This is because the agent increasingly overweights the probability of obtaining the lowest performance level, which leads him to more heavily underweight all other probabilities. Thus, under severe pessimism, the performance segment in which the fixed and high payment is given expands, making incentive-compatible contracts increasingly costly to implement. Moreover, I demonstrate that when the cost of inducing high effort exceeds the corresponding benefits, the principal optimally abandons incentive compatibility. In such cases, if contracting remains desirable, the principal ends up offering a contract with a constant payment for all performance levels; in other words, a salary.

Under optimism, the principal offers a contract that provides large transfers only if the highest performance levels are realized. For all other outcomes, transfers are substantially lower. This type of incentive scheme, where incentives are concentrated at high performance levels, resembles a high-performance commission (Joseph and Kalwani, 1998; Oyer, 2000). Remarkably, when optimism is moderate, incentive-compatible contracts are unnecessary. This occurs because the agent's tendency to overweight the probability of high performance inflates the perceived benefits from exerting high effort, making the incentive compatibility constraint to be slack at the optimum. Thus, as optimism becomes less severe, the optimal contract features lower-powered incentives.

I also consider probability weighting due to likelihood insensitivity (Tversky and Wakker, 1995; Wakker, 2001). According to this type of probability weighting, individuals misperceive probabilities due to cognitive and perceptual limitations. These limitations make individuals insufficiently sensitive to changes in intermediate probabilities, leading them to overweight the probabilities of extreme outcomes, both best and worst (Wakker, 2010).

In the case of such agents, the principal tends to concentrate incentives at extreme performance levels while offering minimal incentives in between. Hence, the optimal contract includes a performance-insensitive segment at intermediate outcomes. That solution resembles compensation schemes that include both a low-performance lump-sum bonus, which rewards (or avoids penalizing) the agent for preventing the worst performance levels, and a high-performance lump-sum bonus. Moreover, as likelihood insensitivity increases, the optimal contract includes a larger performance-insensitive segment. Hence, similar to the case of pessimism, stronger probability weighting due to likelihood insensitivity reduces the effectiveness of incentives, and results in contracts without incentives over a broader range of performance levels.

I empirically corroborate the theoretical prediction that stronger probability weighting, driven by either more severe likelihood insensitivity or pessimism, decreases the extent to which contracts depend on performance. To that end, I use data from the American Life Panel, a representative sample of American households that are regularly surveyed. Specifically, I combine data from the incentivized experiment conducted by Dimmock et al. (2021), which elicited respondents' probability weighting functions, with data from the survey of Maestas et al. (2023), which records the type of compensation respondents receive in their current job. These data allow me to examine the relationship between the strength of probability weighting and the extent to which compensation is performance-based. Consistent with the model's predictions, I find that individuals exhibiting more severe probability weighting, due to pessimism and insensitivity, are on average less likely to be paid with contracts that depend on performance.

The paper concludes by discussing several extensions to the theoretical model. Some of these extensions examine the impact of other deviations from expected utility, such as loss aversion and ambiguity attitude, on optimal contracting. With these, I provide a comprehensive account of optimal incentives without expected utility.[2] Importantly, these extensions demonstrate the flexibility of the model, as it can easily incorporate these phenomena, and the generality of its predictions, since qualitatively similar results are obtained. The remaining extensions are technical, and show that the main theoretical results hold when some assumptions of the model are relaxed.

## 2 Related literature

This paper contributes to the contract theory literature by proposing a solution to the paradox put forward by Salanié (2003), which suggests that the complex solutions predicted by theory do not align with the simplicity of contracts observed in practice. I demonstrate that incorporating probability weighting into a canonical principal-agent model yields contracts that do not depend everywhere on performance optimal. This feature of the solution generates simpler and more realistic contracts. For instance, when individuals are overly pessimistic the resulting contract can involve a salary, and when they are likelihood insensitive, the optimal contract includes lump-sum bonuses.[3]

[2] I do not discuss deviations from expected utility due to violations of transitivity such as preference reversals (Lichtenstein and Slovic, 1971). The most prominent theories of risk developed to capture those behaviors, such as regret theory (Loomes and Sugden, 1982), assume that agents evaluate probabilities accurately, in stark contrast to this setting, but that they suffer from a "distortion of utilities", which calls for an entirely different framework of analysis.

[3] In Online Appendix B, I show that when overly pessimistic agents are also loss averse, then the optimal contract becomes a salary and a lump-sum bonus for high performance. This is consistent with

These two contracts are among the most popular compensation practices.

I also contribute to the behavioral contract theory literature, which incorporates behavioral biases such as loss aversion, dynamic inconsistency, and biased beliefs into principal-agent settings (see Kőszegi, 2014, for a review). I focus on incorporating probability weighting in a setting of moral hazard. To the best of my knowledge, I am the first to do so. This feature is reminiscent of Spalt (2013) who shows that when agents exhibit cumulative prospect theory preferences it is *first-best optimal* to use stock options. I find a similar result in which the optimal contract offered to likelihood-insensitive agents can exhibit an option-like shape. However, the analysis of Spalt (2013) is extended in several ways: by considering the case in which the incentive compatibility constraint *is binding*; by examining the impact of other types of probability weighting, such as those caused by pessimism and optimism; by not restricting the analyses to a parametric form of utility and probability weighting; and by extending his results to a setting of ambiguous probabilities.

The model also demonstrates that the main findings of De La Rosa (2011) and Santos-Pinto (2008) extend to a setting of probability weighting. Specifically, I show that the principal can exploit optimistic agents by overpaying high performance levels and underpaying all other outcomes. This intuition is similar to what De La Rosa (2011) and Santos-Pinto (2008) find for overconfidence and positive self-image, respectively. However, I also demonstrate that this intuition does not necessarily apply to pessimism and likelihood insensitivity, which are the empirically prevalent forms of probability weighting (Fehr-Duda and Epper, 2011; Wakker, 2010). In the case of pessimism, the technical requirement that the optimal contract be monotonic prevents the principal from offering the highest transfers for the lowest performance levels; the outcomes this agent overweights the most. As a result, the principal must seek alternative, often more costly, methods to incentivize the agent, which may sometimes compromise incentive compatibility. Thus, an important insight of the present paper is that pessimism cannot be as easily exploited as optimism. For likelihood insensitivity, a similar difficulty in providing incentives arises at low output levels, which are also overweighted by the agent.[4]

The closest related works are González-Jiménez (2024a) and González-Jiménez (2024b). The former paper demonstrates that agents exhibiting probability weighting can be optimally incentivized through stochastic contracts, which introduce additional risk beyond that inherent in performance. This study complements his findings in two

the result of Herweg et al. (2010). However, I show that this result does not hold when pessimism is not severe, or when individuals exhibit other types of probability weighting, such as optimism or likelihood insensitivity.

[4]This model also differs from De La Rosa (2011) and Santos-Pinto (2008) in that it relies on weaker assumptions as it does not require heterogeneous priors between the agent and principal to generate optimistic attitudes. In the model, probabilities are objectively known to both principal and agent, which disregards, by construction, the formation of heterogeneous priors.

key ways: (i) by providing empirical evidence from a representative sample that probability weighting is associated with contracts that are performance-insensitive rather than riskier; and (ii) by identifying the theoretical conditions under which deterministic contracts may be preferable for the principal. The latter result is presented in Online Appendix D. Furthermore, González-Jiménez (2024b) examines the optimal design of incentives when the agent exhibits loss aversion. I extend his analysis by characterizing optimal contracts in the presence of both loss aversion and probability weighting. Notably, incorporating loss aversion into the framework of this paper expands the range over which the optimal contract is performance-insensitive, making the predictions of the model more realistic. This result is presented in Online Appendix E .

# 3 Theoretical Framework

## 3.1 The General Setup

Consider a principal (referred to as *she/her*) who can hire an agent (referred to as *he/him*) to perform a task. The agent's action consists of exerting effort on the task $e$, which can take two values $\{\underline{e}, \bar{e}\}$. Exerting effort generates a disutility equal to $c(e)$. Specifically, it is assumed that a high level of effort, $\bar{e}$, is costly to the agent $c(\bar{e}) = c$, with $c > 0$, whereas choosing the low effort entails no cost, $c(\underline{e}) = 0$. In Online Appendix D, I relax the assumption that effort is binary and show that the main results of the model can hold in a more general setting in which $e$ is continuous and $c(e)$ is a strictly convex function.

Output from the task is denoted as $q$ and is also referred to as performance. Throughout, I assume that output is stochastic. Hence, $q$ is a random variable that can take any value in the compact interval $[\underline{q}, \bar{q}]$. Moreover, both principal and agent know that $q$ is distributed according to the conditional distribution function $F(q|e)$, which admits the probability density $f(q|e)$.

The relationship between output and effort is assumed to be governed by the monotone likelihood ratio property (MLRP, henceforth):

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

The MLRP implies that high-output realizations are more likely to be drawn from the distribution of output when the agent's level of effort is high than when it is low. Consequently, observing output realizations allow the principal to infer the agent's likely action: higher performance level constitutes clearer evidence that the agent has exerted high effort.

To convince the agent to work on the task, the principal offers a take-it-or-leave-it contract that specifies a transfer schedule $t(q)$. Throughout, it is assumed that the principal has the following utility:

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

where $S(q)$ is a captures the principal's benefit from output. That function is assumed to exhibit $S' > 0$, $S'' \leq 0$ for all $q$, and $S(\underline{q}) = 0$. Importantly, the model focuses on a setting in which the principal is interested in eliciting a high level of effort. Accordingly, it is assumed, wherever possible, that the expected benefit from obtaining that level of effort outweighs its expected cost, namely the expected transfer paid to the agent when effort is high.

The timing of the interaction between agent and principal is as follows. First, the principal offers the contract to the agent. If the contract is accepted, then the agent proceeds to choose $e$. Subsequently, $q$ is realized and the agent's transfer is determined by the schedule $t(q)$. Alternatively, if the contract is rejected, the interaction between the parties immediately ends, and each is left with their reservation utility. I denote the agent's reservation utility as $\bar{U} \geq 0$.

Furthermore, I assume that the transfers specified in $t(q)$ are assumed to be monotonically increasing in output:

**Assumption 2.** *The contract $t(q)$ exhibits $t(q_2) \geq t(q_1)$ for any $q_2, q_1 \in [\underline{q}, \bar{q}]$ such that $q_2 > q_1$.*

Besides the well-known motivations for assuming monotonocity, this property has two novel implications that are relevant to the incentive design problem considered in this paper.[5] First, it guarantees that principal and agent share the same ranking of performance outcomes; that is, both prefer higher performance levels to lower ones. This alignment allows us to focus on the relevant setting in which the principal leverages the agent's probability weighting to extract higher output. Notice that otherwise, we would be in a setting in which probability weighting gets in the way of the principal's objectives. Second, monotonicity prevents the agent from engaging in hedging behavior, which would disregard probability weighting as a valid representation of the agent's preferences. Specifically, hedging would violate the condition of rank-tradeoff consistency (also known as comononotonicity) (Abdellaoui, 2002; Wakker, 2010), which is a key axiom for the models that I use to characterize preferences with probability weighting.

The transfers specified in $t(q)$ enter the agent's utility through the function $u$, about which I make the following assumption:

[5]One of the standard motivations for monotonicity is that it ensures that the agent does not engage in sabotage. Notice that otherwise, he would be incentivized to destroy output with the aim of claiming the highest transfer included in the non-monotonic contract. This justification is also applicable in the current setting.

**Assumption 3.** *The consumption utility function* $u : \mathbb{R}_+ \to \mathbb{R}_+$ *is twice continuously differentiable and exhibits* $u(0) = 0$, $u' > 0$, $u'' < 0$, *and* $-\frac{u''}{u'} < B$ *for* $B < +\infty$.

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

Under the aforementioned assumptions, the agent's preferences can be written as:

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

To remain consistent with standard notation in the literature, I work throughout with *decumulative probabilities*. Thus, a probability in this model refers to the likelihood of obtaining a realization better than a given performance level $Q \in [\underline{q}, \bar{q}]$. Formally, let a probability be $p = 1 - F(Q|e)$ for any $e$. This alternative representation of probabilities has no impact on the solution to the incentive design problem. To see this, note that the agent's preference in equation (1) is equivalent to the following representation in terms of decumulative probabilities:[6]

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

## 3.2 Probability Weighting Functions and Rank-Dependent Utility

When the agent perceives probabilities accurately, expected utility theory (EUT, henceforth) presented in equation (2), captures his preferences. In contrast, I relax this assumption by allowing the agent to exhibit probability weighting. I model this feature by means of a *probability weighting function*, $w$, that transforms probabilities $p$. The following assumptions are imposed on $w$:

**Assumption 4.** *The probability weighting function* $w : [0,1] \to [0,1]$ *is twice continuously differentiable, and fulfills the following conditions:*

*i)* $w(0) = 0$ *and* $w(1) = 1$;

*ii)* $w'(p) > 0$ *for all* $p \in (0,1)$;

*iii) For some* $\tilde{p} \in [0,1]$, $w''(p) < 0$ *if* $p \in [0,\tilde{p})$ *and* $w''(p) > 0$ *if* $p \in (\tilde{p},1]$;

*iv) If* $\tilde{p} = 1$, *then* $\lim_{p\to 0^+} w'(p) = +\infty$ *and* $\lim_{p\to 1^-} w'(p) = 0$;

*v) If* $\tilde{p} = 0$, *then* $\lim_{p\to 0^+} w'(p) = 0$ *and* $\lim_{p\to 1^-} w'(p) = +\infty$; *and*

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

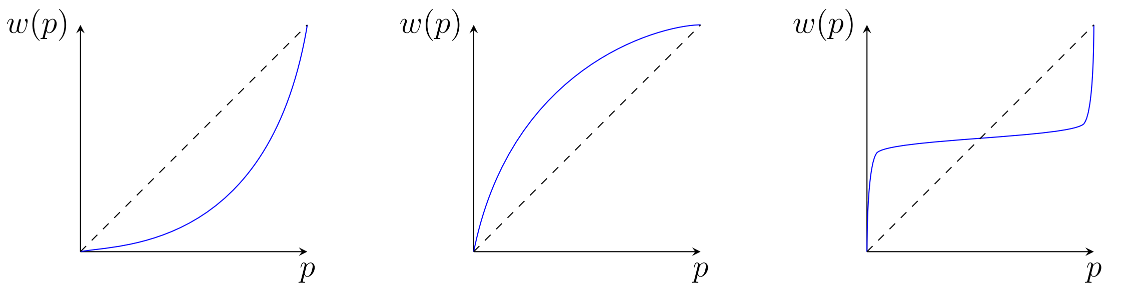
$$\int_{q_1}^{q_2} \mathrm{d}F(q|e) = F(q_2|e) - F(q_1|e) = 1 - F(q_1|e) - (1 - F(q_2|e)) = \int_{q_2}^{q_1} \mathrm{d}\left(1 - F(q|e)\right).$$

*vi) If* $\tilde{p} \in (0,1)$, *then* $\lim_{p\to 0^+} w'(p) = +\infty$ *and* $\lim_{p\to 1^-} w'(p) = +\infty$.

The probability weighting function maps the unitary interval onto itself and is continuous and increasing. It has at least two fixed points: one at impossibility ($w(0) = 0$), which corresponds to the probability of obtaining a realization *better than the highest output level*, and one at certainty ($w(1) = 1$), which corresponds to the probability of obtaining a *better or equal realization than the lowest output level*.

Moreover, the probability weighting function has an inflection point denoted by $\tilde{p}$ that crucially determines its shape. When $\tilde{p} = 0$, the function is convex everywhere. That shape, along with the fact that decumulative probabilities are being transformed, implies that large probabilities—the probabilities of low performance levels— receive larger weight than small probabilities—the probabilities of high performance levels. Figure 1a presents an example of a convex weighting function. In contrast, when $\tilde{p} = 1$ the probability weighting function is concave everywhere. That shape implies that small probabilities—those associated with high performance levels— receive larger weight than large probabilities—those associated with low performance levels (Figure 1b). Finally, when $\tilde{p} \in (0,1)$, the function exhibits an inverse-S shape (Figure 1c). In this latter case, the agent assigns large weights to small and large probabilities, the probabilities of extreme performance levels, while assigning similar weights to the probabilities of intermediate output levels. Notably, the latter shape implies the existence of an interior point $\hat{p} \in (0,1)$ where $w(\hat{p}) = \hat{p}$.[7]

Figure 1: Examples of probability weighting functions




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

Note: Dashed lines represent accurate perception of probability.

The preferences of an agent who exhibits probability weighting are characterized by

[7]The seemingly drastic assumptions of extreme sensitivity to almost-certain and almost-impossible realizations, i.e. $\lim_{p\to 1^-} w'(p) = \infty$ and $\lim_{p\to 0^+} w'(p) = \infty$, are incorporated in the most prominent proposals of parametric forms of probability weighting functions, such as those of Prelec (1998), Tversky and Kahneman (1992), and Goldstein and Einhorn (1987). Furthermore, these assumptions have relevant implications for the results that are formalized in Online Appendix A .

rank-dependent utility (RDU, henceforth):

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

RDU generalizes EUT by applying probability weighting to the decumulative function $1 - F(q|e)$. Thus, for a given outcome, $Q \in [\underline{q}, \bar{q}]$, and a given effort level, $e' \in \{\underline{e}, \bar{e}\}$, the agent considers the probability of obtaining a higher outcome, which is $1 - F(Q|e')$, but this probability is perceived as $w\Big(1 - F(Q|e')\Big)$. Accordingly, when an outcome that is infinitesimally worse than $Q$ is taken as benchmark, obtaining $Q$ implies a marginal difference in perceived probabilities captured by the expression $\mathrm{d}\Big(w\Big(1 - F(Q|e')\Big)\Big)$. Notice that this expression is the differential of the integral in (3) when $q$ is replaced by $Q$. Therefore, the rank-dependent functional given in (3) implies that the utility derived from an outcome $Q$, $u(t(Q))$, is weighted by its contribution to the perceived probability $\mathrm{d}\Big(w\Big(1 - F(Q|e')\Big)\Big)$, and all such weighted utilities are summed over the set of all possible performance levels $q \in [\underline{q}, \bar{q}]$.

Notably, under RDU, the agent's risk attitude is jointly determined by the curvature of the functions $u$ and $w$. The influence of the curvature of $u$ on risk attitude is common to EUT and RDU. However, the influence of the curvature of $w$ on risk attitude is exclusive to RDU and is referred to as *probabilistic risk attitude* (Wakker, 1994). It captures the influence of deviations from expected utility in decision making under risk. The aim of this model is to establish how this novel source of risk attitude affects optimal contracting in a moral hazard setting.

## 3.3 The Principal's Problem

When facing an agent with preferences described by equation (3), the principal's problem consists of implementing a contract that is accepted by the agent, a condition I refer to as participation constraint (PC, henceforth), that incentivizes him to exert high effort, a condition I refer to as incentive compatibility constraint (IC, henceforth), and that satisfies the monotonicity constraint from Assumption 2 (M, henceforth). Formally, her program is:

$$
\begin{aligned}
\max_{t(q)} \quad & \int_{\underline{q}}^{\bar{q}} \Big( S(q) - t(q) \Big) \, \mathrm{d}F(q|\bar{e}) \\
\text{s.t.} \quad & \int_{\bar{q}}^{\underline{q}} u(t) \, \mathrm{d}w(1 - F(q|\bar{e})) - c \geq \bar{U} && \text{(PC)}, \\
& \int_{\bar{q}}^{\underline{q}} u(t) \, \mathrm{d}w(1 - F(q|\bar{e})) - c \geq \int_{\bar{q}}^{\underline{q}} u(t) \, \mathrm{d}w(1 - F(q|\underline{e})) && \text{(IC)}, \\
& t(q_2) \geq t(q_1) \text{ for all } q_2, q_1 \in [\underline{q}, \bar{q}] \text{ such that } q_2 > q_1 && \text{(M)}
\end{aligned}
\tag{4}
$$

The program described in (4) indicates that the principal, as it is standard in the moral hazard literature, is assumed to be fully informed about the agent's preferences. Consequently, she designs a contract that fully incorporates that information in order to elicit a high level of effort. In Online Appendix D, I examine how the optimal contract would change if this assumption were relaxed. This alternative analysis aligns better with empirical research showing significant variability in RDU preferences over time (Kilka and Weber, 2001; Rottenstreich and Hsee, 2001; Zeisberger et al., 2012), which suggests that the assumption of the principal knowing the agent's preferences may be overly restrictive. Notably, I find that similar contracts emerge in that setting.

Moreover, the program in (4) also assumes that the RDU agent cannot "debias" himself by interpreting the contract offered by the principal as a signal of his probability weighting. One possible interpretation of this assumption is that the agent is naive about the *dynamic inconsistency* caused by probability weighting (Barberis, 2012; Ebert and Strack, 2015).[8] Specifically, while the agent may initially plan to reject a contract that signals that he suffers from probability weighting, he deviates from this plan because of his misperception of probabilities and accepts such a contract.[9] Another interpretation of this assumption is based on the well-known property that probability weighting prevents the agent from performing backward induction (Karni and Schmeidler, 1991; Machina, 1989). Thus, he is unable to understand the reasoning that the principal used to design the contract, which relies on backward induction: first account for the influence of the agent's probability weighting on the effectiveness of incentives to subsequently structure a monotonic contract that ensures participation and incentive

[8]If the agent were sophisticated about his dynamic inconsistency, he would reject any contract that provides clear evidence of his susceptibility to probability weighting. A way to overcome this rejection would be to employ a mixed strategy that makes the agent indifferent to accepting or rejecting the contract (Henderson et al., 2017). I focus on naive agents and pure strategies, making this approach inapplicable to my framework.

[9]This scenario can arise when the agent, prior to receiving the contract, believes that the principal will exploit his probability weighting with probability $s \in [0, 1]$. If this probability is sufficiently high, the agent may be inclined to reject any contract offered. However, after receiving the offer $t(y)$, he might accept the contract if the expected benefits of accepting it, given by $(1 - s) \int_{\underline{y}}^{\bar{y}} \mathrm{d}\,(1 - F(y|\hat{e}))$, are sufficiently overweighted.

compatibility.

## 3.4 The Weighted Monotone Likelihood Ratio Property

Before presenting the optimal contract for RDU agents, I introduce a crucial assumption for the theoretical analyses presented below. Notice that while the agent is aware of the relationship between performance and effort (Assumption 1), his probability weighting leads him to evaluate this relationship differently. Specifically, he adjusts the ratio $\frac{f(q|\underline{e})}{f(q|\bar{e})}$ by the factor $\frac{w'\left(1-F(q|\underline{e})\right)}{w'\left(1-F(q|\bar{e})\right)}$. Therefore, changes in the likelihood ratio $\frac{f(q|\underline{e})}{f(q|\bar{e})}$ might be misperceived, as they are accompanied by changes in the weighting ratio $\frac{w'\left(1-F(q|\underline{e})\right)}{w'\left(1-F(q|\bar{e})\right)}$.

Throughout, it is assumed that regardless of the agent's probability weighting, he still perceives that choosing high effort increases the probability of obtaining a high performance level *more* than the probability of obtaining a low performance level. Hence, the ratio $\frac{w'\left(1-F(q|\underline{e})\right)f(q|\underline{e})}{w'\left(1-F(q|\bar{e})\right)f(q|\bar{e})}$ decreases with performance. I refer to this alternative version of the MLRP as the Weighted Monotone Likelihood Ratio Property.

**Assumption 5.** *The W-MLRP states that:*

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

*for all* $q$.

The W-MLRP rules out situations where, due to probability weighting, the agent erroneously believes that low effort is more likely to lead to high output than high effort. Thus, it ensures that the agent's perceived relationship between effort and performance remains directionally consistent with the standard MLRP. As a result, solutions with significant deviations from standard optimal contracts—such as non-monotonicities that make the M constraint bind—cannot be attributed to probability weighting causing fundamental reversals in perceived signal informativeness.

I assume that the principal internalizes Assumption 5. Hence, not only she knows the actual relationship between effort and outcome probabilities (Assumption 1), but also knows how the agent perceives that relationship. Moreover, she uses this knowledge to design incentives optimally. In particular, the principal avoids implementing incentives at performance levels whose associated probabilities the agent underweights. There, he perceives that high effort only slightly increases the likelihood of achieving a higher outcome. In those cases, signals about the agent choosing high effort are weak and seemingly uninformative, making incentive provision costly. Instead, the principal targets performance levels whose associated probabilities the agent overweights. In these segments, the agent perceives that high effort substantially increases the chances of achieving a high performance, resulting in seemingly informative signals about his action. Therefore, the principal implements incentives at these performance levels to effectively motivate the agent.

The following two lemmas highlight some properties of the W-MLRP. The proofs of the main results of the paper are presented in Online Appendix A.

**Lemma 1.** *The W-MLRP implies:*

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

*(ii) the MLRP if* $w'\Big(1 - F(q|\bar{e})\Big) \geq w'\Big(1 - F(q|\underline{e})\Big)$.

**Lemma 2.** *If the MLRP holds and*

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

*holds for all* $q$*, then the W-MLRP must hold.*

The W-MLRP implies first-order stochastic dominance (Lemma 1 (i)). Also, consistent with the previous explanation of how the principal implements incentives, the W-MLRP is not a generalization nor a special case of MLRP. There are performance segments in which the W-MLRP implies the MLRP, but other segments in which the W-MLRP is implied by the MLRP (Lemma 1 (ii)). Specifically, when the marginal weight of the probability generated by $\bar{e}$ exceeds that of the probability generated by $\underline{e}$, the agent's probability weighting generates the perception that higher performance levels are more informative about high effort than they actually are. At the performance levels in which that condition holds, the W-MLRP implies the MLRP. However, since the derivative of $w$ must be on average one—i.e. $\int_0^1 w'(p)dp = 1$—this condition cannot hold for all probabilities, and thus for all $q$. As a result, there must be performance segments in which the W-MLRP does not imply the MLRP, and the informativeness of higher performance realizations is lower than under accurate perception of probabilities.

Furthermore, Lemma 2 shows that the W-MLRP can be implied by the standard MLRP when the curvature of $w$ is restricted. In particular, when the agent's probability weighting function is more convex at the probability associated with low effort than at the probability associated with high effort, the MLRP can imply the W-MLRP.[10] As will be discussed in the next section, this condition is related to how pessimistic or optimistic the agent is at the probabilities induced by his actions. Accordingly, the

[10]This requirement on the shape of $w$ ensures that the derivative $\frac{\partial\left(\frac{w'\left(1-F(q|\underline{e})\right)}{w'\left(1-F(q|\bar{e})\right)}\right)}{\partial q}$ is nonincreasing. As a result, the agent's probability weighting does not reverse or counteract the MLRP.

agent must be more pessimistic (optimistic) about the probabilities generated by low effort (high effort), and such pessimism about low effort implies the MLRP. Throughout, I will assume that the W-MLRP holds. However, note that thanks to this lemma, the results of the model can be obtained under more standard assumptions, provided this restriction on the convexity of the probability weighting function is satisfied.

### 3.5 The General Solution to the Principal's Problem

The results presented in this section are essential for understanding the solution to the principal's problem when the agent suffers from probability weighting. They show that the shape of the probability weighting function crucially influences the maximization problem, and, consequently, the type of solution that applies. The following proposition states that the monotonicity constraint is slack at the optimum when the probability weighting function is everywhere concave.

**Proposition 1.** *Suppose that Assumptions 1-5 hold and that the PC and IC constraints bind. The M constraint is slack for all $q$ if and only if $w''(p) < 0$ for all $p \in [0, 1]$. The same result is obtained if the IC is not included in the principal's program.*

The result in Proposition 1 implies, via complementary slackness, that the optimal contract is strictly increasing for all $q$ if and only if the weighting function is everywhere concave. What remains to be established are the conditions under which the monotonicity constraint binds. These conditions are presented in the following corollary.

**Corollary 1.** *Suppose that Assumptions 1-5 hold and that the PC and IC constraints are binding. The M constraint binds if*

$$\frac{w''(1-F(q|\bar{e}))}{w'(1-F(q|\bar{e}))}f(q|\bar{e}) > -\mu w'(1-F(q|\bar{e}))u'(t(q))\frac{d}{dq}\left(\frac{w'\big(1-F(q|\underline{e})\big)f(q|\underline{e})}{w'\big(1-F(q|\bar{e})\big)f(q|\bar{e})}\right),$$

*where $\mu$ is the lagrangian multiplier of the IC constraint. Moreover, if the IC is not included in the principal's program, this condition becomes $\frac{w''(1-F(q|\bar{e}))}{w'(1-F(q|\bar{e}))}f(q|\bar{e}) > 0$.*

When the IC constraint binds ($\mu > 0$) and the agent's probability weighting function is convex ($w'' > 0$), the M constraint binds at performance levels where the marginal weight $w'(p)$ becomes sufficiently small. At those performance levels, the right-hand side of the inequality stated in Corollary 1 becomes negligible, while the left-hand side remains large.[11] Consequently, by complementary slackness, the optimal contract must flatten out at those output realizations.

[11]Lemma A4 in the Online Appendix D shows that when the weighting function is convex, then $\lim_{q\to\bar{q}} \frac{w''(1-F(q|\bar{e}))}{w'(1-F(q|\underline{e}))} = +\infty$.

We are now in a position to describe how the problem presented in Eq. (4) is solved using Proposition 1 and Corollary 1. The solution proceeds in the following steps. First, I compute the solution from the first-order approach to the principal's problem without the M constraint, which I refer to as the solution to the "unconstrained" problem. Second, Proposition 1 establishes that this unconstrained solution is optimal when the weighting function $w(p)$ is everywhere concave, or when that function is convex but $w'(p)$ remains sufficiently large. However, if $w(p)$ is convex and $w'(p)$ is sufficiently small, Corollary 1 states that the optimal contract must be flat over the performance segment where these conditions met. In such cases, the solution to the relaxed problem is replaced by a constant transfer equal to the payment specified by the unconstrained solution at the lower boundary of the segment.

Importantly, Proposition 1 and Corollary 1 also apply when the IC constraint is not included in the principal's program. Therefore, this way of finding a solution can be used to characterize optimal contracts when the IC does not bind at the optimum.

# 4 Pessimism and Optimism

This section examines the optimal design of contracts when the principal faces two specific types of RDU agents: pessimists and optimists. These agents exhibit probability weighting generated by *motivational factors* (Wakker, 2010).

## 4.1 Definitions

Pessimists pay more attention to unfavorable outcomes (i.e., low performance levels) than to favorable ones (i.e., high performance levels) (Wakker, 2001).[12] This pessimistic attitude toward risk is included into the model by means of a convex probability weighting function, which assigns greater weight to a probability if its associated with a lower performance level.

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

Pessimism generates probabilistic risk aversion, as it leads the agent to assign excessive weight to the probabilities of worst outcomes and insufficient weight to those of higher outcomes.

[12] According to (Wakker, 2010), this attitude may stem from an *irrational* belief that unfavorable outcomes are more likely to occur, leading to an unrealistic overweighting of their probabilities. Alternatively, if probability weighting is interpreted normatively, Wakker (2010) notes that this attitude can also be due to a conscious and deliberate focus on unfavorable outcomes in decisions, without overestimating their likelihood.

By contrast, optimists pay more attention to favorable than to unfavorable outcomes (Wakker, 2001). This attitude is reflected in a concave probability weighting function, which assigns greater weight to a probability if it is associated with a higher performance level.

**Definition 2.** *Optimism is characterized by a probability weighting function $w(p)$, defined in Assumption 4, with the additional restriction that $\tilde{p} = 1$. Therefore, $\lim_{p \to 0^+} w'(p) = +\infty$ and $\lim_{p \to 1^-} w'(p) = 0$.*

Optimism generates probabilistic risk seeking, as it leads the agent to assign excessive weight to the probabilities of best outcomes and insufficient weight to those of lower outcomes. However, this optimistic attitude does not necessarily imply that the agent is risk seeking overall, since the curvature of the utility function $u$ may induce sufficient risk aversion so as to offset the risk-seeking behavior generated by the probability weighting function $w$.

We are interested in investigating how more severe pessimism and optimism influence the optimal contract. The following definition from Yaari (1987) provides a formal basis for understanding varying degrees of pessimism and optimism.

**Definition 3.** *Agent $i$ is more pessimistic (resp. 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 a twice continuously differentiable, strictly increasing, and convex (resp. concave) function.*

A probability weighting function that is more convex than another generates stronger pessimism because it causes the agent to assign larger weights to the probabilities of lower performance and smaller weights to the probabilities of higher performance. The reasoning is analogous for concave probability weighting functions.

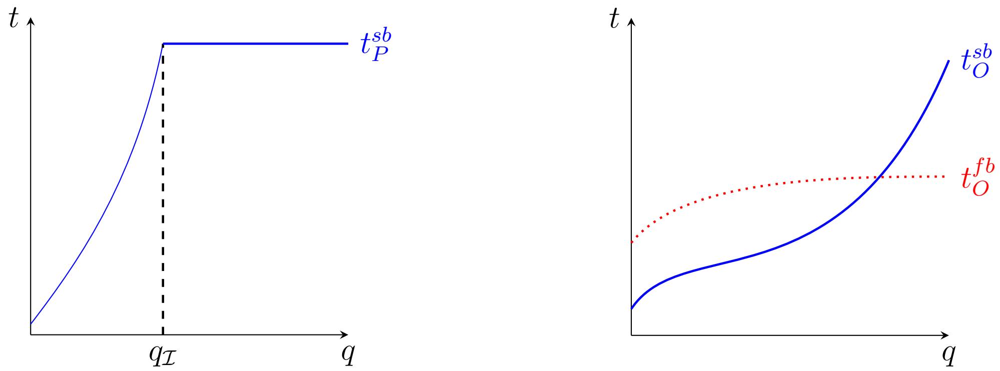
## 4.2 Optimal contracts under Pessimism

We can now proceed to analyze the optimal contracts given to RDU agents with pessimism or optimism. The following proposition describes the properties of the contract that solves the principal's problem when the agent is a pessimist.

**Proposition 2.** *Suppose that Assumptions 1-5 and pessimism hold. The optimal contract, $t_P^{sb}(q)$, is increasing in $q$ up to some threshold $q_{\mathcal{I}} \in (\underline{q}, \bar{q})$ after which payment is constant in $q$.*

When the agent is pessimistic, the optimal contract includes incentives only at low performance levels. Specifically, the contract specifies larger transfers to reward higher performance at the lower end of the output interval, while being performance-insensitive elsewhere. These incentives, along with the agent's overweighting of the

Figure 2: Illustration of Propositions 2 and 4.




(a) Optimal contract under pessimism

(b) Optimal contracts under optimism

Note: The red dotted lines represent contracts when the IC constraint is slack at the optimum and blue solid lines represent contracts when the IC is binding.

probabilities of low performance levels, motivate him to choose a high level of effort. An example of this contract's shape is presented in Figure 2a.

I now use Definition 3 to examine the effect of stronger pessimism on the contract from Proposition 2.

**Corollary 2.** *An increase in the degree of pessimism, in the sense of Definition 3, enlarges the segment $q \in [q_{\mathcal{I}}, \bar{q}]$ for which the contract $t_P^{sb}(q)$ (Proposition 2) is constant in $q$.*

Stronger pessimism leads the principal to increasingly concentrate incentives at lower performance levels while avoiding them in the remainder of the performance space. Consequently, the optimal contract depends less on performance. Importantly, this increased performance insensitivity of the optimal contract entails a cost for the principal, as she must pay a high and fixed transfer over a larger performance segment. The following result formalizes this observation.

**Corollary 3.** *The principal experiences a profit loss from offering a contract that is monotonic and incentive-compatible to the pessimist, and this loss increases with the agent's degree of pessimism.*

An implication of Corollary 2 is that an overly pessimistic agent requires an optimal contract in which rewards and punishments are concentrated in a narrow output subset in the neighborhood of $\underline{q}$. According to Corollary 3, such a contract may be overly costly because a large and fixed transfer must be paid in the remainder of the output space. The question then arises as to whether the principal can afford to implement incentive-compatible contracts for all levels of pessimism. The following proposition demonstrates that severe pessimism can undermine incentive compatibility.

**Proposition 3.** *Suppose that Assumptions 1-5 and pessimism hold. Then, there exists a level of pessimism at which the principal chooses not to implement a high level of effort and the optimal contract is a salary that guarantees participation. Agents who are more pessimistic than that level, in the sense of Definition 3, are also offered such a contract.*

As mentioned above, agents with acute pessimism require excessively costly incentives to be motivated. Proposition 3 demonstrates that these costs cannot be outweighed by the benefits of implementing high effort, leading the principal to optimally abandon incentive compatibility. If she still chooses to contract under these conditions, because the benefits of participation counteract the costs, the optimal contract would consist of a fixed payment to ensure participation; in other words, a salary.[13]

Propositions 2 and 3, and Corollary 2 offer an explanation for the prevalence of performance-insensitive contracts in organizations based on pessimism. Specifically, Proposition 3 suggests that severe pessimism can account for the prevalence of salaries in organizations (Lazear and Oyer, 2007; Salanié, 2003).[14] Moreover, these theoretical results are consistent with two key empirical findings. First, that highly risk-averse individuals are more likely to work under contracts that are not contingent on performance (Dohmen et al., 2011; Grund and Sliwka, 2010), and tend to select occupations with lower earnings variability (Bonin et al., 2007). Second, that individuals who are more risk-averse and pessimistic about their own performance opt for contracts with lower-powered incentives (Larkin and Leider, 2012).

### 4.3 Optimal Contracts under Optimism

Let us now turn to the case of optimism. The following proposition characterizes the solution to the principal's problem when the agent is an optimist.

**Proposition 4.** *Suppose that Assumptions 1-5 and optimism hold. There exists a unique threshold cost level $\hat{c}_O > 0$, such that:*

*i) If $c < \hat{c}_O$, the IC constraint is slack at the optimum and the optimal contract $t_O^{fb}(q)$ is everywhere increasing in $q$.*

*ii) If $c \geq \hat{c}_O$, the IC constraint binds at the optimum and the optimal contract $t_O^{sb}(q)$ is everywhere increasing in $q$ but exhibits higher (resp. lower) transfers relative to $t_O^{fb}(q)$ at high (resp. low) output levels.*

In the case of an optimistic agent, the IC constraint may be slack at the optimum. This occurs because the optimal contract without the IC constraint specifies transfers

[13] Notice that if the principal decides against contracting under such conditions, due to the high costs associated with failing to elicit high effort, the model characterizes the interesting scenario in which the agent's probability weighting leads to the collapse of a contracting relationship that would have otherwise occurred.

[14] In Section 7, I discuss how this result can be further reinforced using ambiguity aversion, a form of pessimism towards events of unknown probability relative to events in which probabilities are known.

that increase in performance, and although these incentives are weak, they suffice to motivate the agent. Specifically, the relatively high transfers offered for the highest performance levels, combined with the agent's tendency to overweight the probability of achieving those outcomes, convince him that exerting high effort is profitable. In this case, the optimal contract is given by Proposition 4 i).

However, when the cost of exerting high effort is sufficiently high, the incentives provided by the contract from Proposition 4 i) become insufficient. In this case, the optimistic agent requires a contract with higher-powered incentives, and the IC constraint binds at the optimum. A contract with stronger rewards for achieving high performance levels, when combined with the agent's optimism, increases the perceived benefits of exerting high effort and thereby offsets its elevated cost. Hence, in such a case, the resulting optimal contract is characterized by Proposition 4 ii). An illustrative example of the two contracts described by Proposition 4 is presented in Figure 2b.

The following comparative static result is useful in understanding the effect of stronger optimism on the optimal contract.

**Corollary 4.** *An increase in the degree of optimism, in the sense of Definition 3, reduces the set of costs $(0, \hat{c}_O]$ for which the contract, $t_O^{fb}(q)$ (Proposition 4 i)) is optimal.*

Stronger optimism makes the lower-powered contract from Proposition 4 i) less often optimal. The agent's increased overweighting of the probability of high performance levels can lead him to erroneously conclude that such outcomes can be achieved even with low effort. That mistaken conviction makes it necessary for the principal to offer stronger incentives, such as those included in the contract described in Proposition 4 ii), to ensure that the perceived benefits of exerting high effort remain sufficiently attractive to the agent.

Proposition 4 and Corollary 4 offer an explanation for the empirical finding that risk-loving individuals tend to choose higher-powered incentive schemes, even when doing so reduces their expected earnings (Bellemare and Shearer, 2010; Larkin and Leider, 2012). Furthermore, Corollary 4 correctly predicts the empirical regularity that the power of incentives will increase with the individuals' optimism (Humphery-Jenner et al., 2016; Larkin and Leider, 2012). Moreover, as it will be further explained below, these results account for the tendency of optimistic individuals to accept contracts with lower average pay (Sautmann, 2013). Otto (2014) shows that such exploitation of optimism occurs in the contracts of CEOs, thus providing empirical evidence of the model's predictions outside the lab.

Furthermore, the theoretical results under optimism are also consistent with the findings of De La Rosa (2011) and Santos-Pinto (2008), who incorporate overconfidence and positive self-image, respectively, into principal-agent frameworks. Thus, the present analysis demonstrates that the results of those models extend to a setting in which the agent exhibits probability weighting due to optimism. This is not a trivial finding, as probability weighting and overconfidence represent different psychological phenomena (Abdellaoui et al., 2023). Probability weighting reflects risk attitudes, whereas overconfidence may stem from biases in beliefs due to incorrect updating—the explanation given by De La Rosa (2011) and Santos-Pinto (2008)) for biased beliefs in their models. In a setting with given probabilities, such as the one that is considered in this paper, beliefs about probabilities do not play any role. This distinction underscores that similar contractual outcomes can emerge from different behavioral assumptions.

To conclude this section, I compare the transfers specified in the contracts from Propositions 2 and 4 ii) to the classical solution of Holmstrom (1979), which I denote by $t^{sb}_{EU}$. Recall that this classical solution to the standard principal-agent problem specifies transfers that strictly increase everywhere in performance.

**Corollary 5.** *Let $q^* \in (\underline{q}, \bar{q})$ be the unique output level that satisfies $w'\big(1 - F(q^*|\bar{e})\big) = 1$. Then,*

*i) $t^{sb}_O(q)$ (Proposition 4 ii)) offers lower transfers in $q \in [\underline{q}, q^*)$ and higher transfers in $q \in (q^*, \bar{q}]$ relative to $t^{sb}_{EU}$.*

*ii) $t^{sb}_P(q)$ (Proposition 2) offers higher transfers in $q \in [\underline{q}, q^*)$ and lower transfers in $q \in (q^*, \bar{q}]$ relative to $t^{sb}_{EU}(q)$.*

While the contracts given to optimistic and EUT agents are qualitatively similar, i.e. both contracts specify transfers that increase in performance, they impart incentives differently. In particular, the contract targeting the optimist offers lower transfers at low performance levels and higher transfers at high performance levels relative to the classical solution. Hence, the principal exploits optimism by overpaying performance levels whose probabilities the agents overweights and by underpaying lower but likely realizations of output. As discussed previously, this feature of the contract is essential to motivate the optimist.

# 5 Likelihood insensitivity

This section investigates the optimal design of contracts when the principal faces agents who suffer from probability weighting due to *likelihood insensitivity*.

## 5.1 Definitions

Likelihood insensitivity represents the cognitive and perceptual limitations that prevent individuals from accurately distinguishing between probabilities (Tversky and Wakker, 1995; Wakker, 2001, 2010). An agent with likelihood insensitivity exhibits extremity-oriented behavior; he assigns small weights to probabilities of intermediate outcomes and large weights to the probabilities of extreme outcomes (highest and lowest performance levels) (Baillon et al., 2023). This yields an inverse S-shaped probability weighting function, which is depicted in Figure 1c.

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

According to Definition 4, likelihood insensitivity is characterized by a probability weighting function with an interior fixed point at $\hat{p} = 0.5$.[15] This implies that the agent tends to evaluate probabilities of intermediate outcomes simplistically, treating them as roughly "50-50" (i.e. either the event happens or it won't) (Wakker, 2010). As a result, the probabilities near the extremes—very high or very low performance levels—are systematically overweighted.

The following definition, based on Baillon et al. (2025), provides a formal basis for understanding varying degrees of likelihood insensitivity. It will enable us to investigate how a contract must be adjusted as a result of more severe likelihood insensitivity.

**Definition 5.** *Agent $i$ is more likelihood-insensitive than agent $j$ if $w_i = \phi \circ w_j$ where $w_i$ and $w_j$ are their respective probability weighting functions, and $\phi : [0,1] \to [0,1]$ is a probability weighting function with likelihood insensitivity in the sense of Definition 4.*

A probability weighting function with a more pronounced inverse-S shape, i.e. shallower at intermediate probabilities and steeper at extreme probabilities, generates stronger likelihood insensitivity. Such a function causes the agent to assign less weight to probabilities of intermediate output levels while assigning more weight to extreme probabilities.

## 5.2 Optimal Contracts under Likelihood Insensitivity

The following proposition characterizes the optimal contract offered to the likelihood-insensitive agent.

**Proposition 5.** *Suppose that Assumptions 1-5 and likelihood insensitivity hold. Then there exists a unique threshold cost level, $\hat{c}_L > 0$, such that:*

*i) If $c < \hat{c}_L$, the IC constraint is slack at the optimum and the optimal contract, $t_L^{fb}(q)$, is constant up to the unique threshold $\tilde{q} \in (\underline{q}, \bar{q})$ after which pay is increasing in $q$.*

*ii) If $c \geq \hat{c}_L$, the IC constraint is binding, and the optimal contract, $t_O^{sb}(q)$, is either everywhere increasing in $q$ and increases more steeply at extreme values of $q$; or*

[15] Recall from Assumption 4 that $\hat{p}$ is the interior fixed point that results from the inflection point $\tilde{p}$ being interior.

*iii) it pays a constant amount for some finite, fixed, and compact interval, but above and below that interval pay is steeply increasing in $q$.*

Proposition 5 i) states that the IC constraint may be slack at the optimum. The rationale for this result is analogous to that given for Proposition 4 i); the agent's overweighting of the probabilities of high performance levels, combined with a contract that provides low-powered incentives at those levels, suffice to convince him that exerting high effort is profitable. Moreover, because the agent is relatively insensitive to changes in the probabilities of intermediate performance levels, the resulting contract takes the form of an "option-like" incentive scheme: it is flat over low and intermediate performance levels but increases with performance at the high end of the output space.[16]

This result echoes the finding of Spalt (2013), according to which an option-like contract is *first-best optimal* under prospect theory preferences. However, Proposition 5 i) offers new insights. First and more importantly, it extends Spalt (2013) by showing that this result can hold in a moral hazard setting. Second, it shows that such a contract shape is due to the overweighting of small and large probabilities caused by likelihood insensitivity. Third, because loss aversion has so far not been introduced into the model, I demonstrate that this bias is not responsible for the option-like contract shape described in Proposition 5 i). Finally, this contract emerges without making assumptions about the functional form of probability weighting or consumption utility.

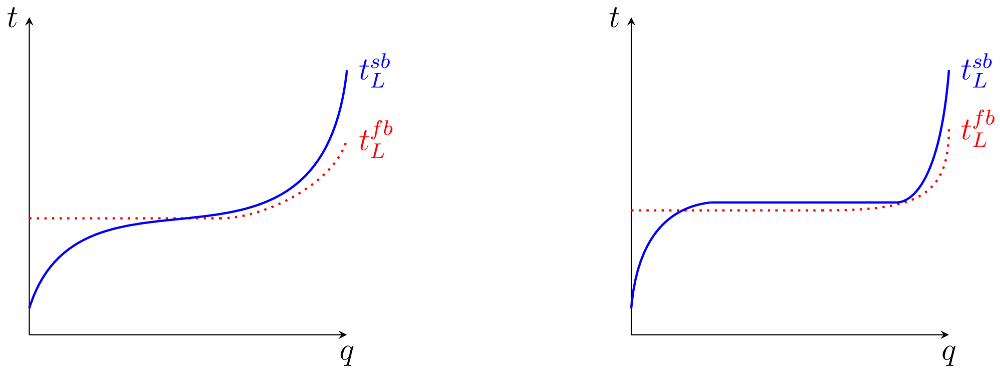
Proposition 5 ii) and iii) show that when the cost of choosing the high level of effort is sufficiently high, i.e. $c \geq \hat{c}_L$, the incentive compatibility constraint binds at the optimum. The principal must offer a contract with higher-powered incentives, compared to those specified in the contract from Proposition 4 i), to convince the agent that exerting a high level of effort pays off despite its elevated costs. Given the agent's probability weighting, the most cost-effective way to achieve this motivation is by concentrating these incentives at extreme output levels—the performance levels whose probabilities the agent overweights. Thus, the resulting contract can be either increasing everywhere, like the contract described by Proposition 5 ii) and depicted in Figure 3a, or performance-insensitive at intermediate performance levels, like the contract with bonuses at extreme performance levels described by Proposition 5 iii) and depicted in Figure 3b. The principal's choice between these two types of contracts depends on the agent's degree of likelihood insensitivity, as will be discussed below.

Let us turn to examine the influence of stronger likelihood insensitivity on the second-best contract. The focus is on the case in which the incentive-compatibility constraint is binding at the optimum.

**Corollary 6.** *Assume that the IC constraint is binding. If agent $i$ is more likelihood insensitive*

[16] The term "option-like" was first used by De Meza and Webb (2007) to describe contracts with this shape, as they resemble the profit profile of a "*call option*."

Figure 3: Illustration of Proposition 5.




(a) Contracts under moderate insensitivity (b) Contracts under strong insensitivity

Note: The red dotted lines represent first-best contracts and blue solid lines represent second-best contracts.

*than agent $j$, then the contract, $t_L^{sb}$ (Proposition 5), offered to $i$ has a larger performance segment in which pay does not depend on performance.*

The principal responds to the agent's reduced sensitivity to the probabilities of intermediate output levels by increasingly minimizing the implementation of incentives in that segment. Consequently, incentives are more concentrated at extreme output levels. This creates incentive schemes with a performance-insensitive segment at intermediate output levels, as in the contract from Proposition 5 iii). Such a contract, which offers bonuses at extreme performance levels, is well-suited for agents with strong likelihood insensitivity, as their distorted perception of probabilities renders incentives at intermediate performance levels inefficient and wasteful for the principal.

Corollary 6 also implies that the performance-insensitive segment of the optimal contract must shrink as the agent becomes less likelihood-insensitive. Thus, an incentive scheme that is everywhere increasing would be effective in the case of an agent with modest likelihood insensitivity, since his slight probability weighting does not prevent him from responding to incentives over the entire output space. He is thus offered the contract described by Proposition 5 ii).

Proposition 5 and Corollary 6 provide a novel justification for contracts with performance-insensitive segments based on likelihood insensitivity. This theoretical prediction is supported by empirical findings. Previous studies have found that likelihood insensitivity is negatively correlated with measures of cognitive ability such as IQ and GPA (Choi et al., 2022; Dimmock et al., 2021; l'Haridon and Vieider, 2019). Moreover, (Curme and Stefanec, 2007) and (Dohmen et al., 2011) find that individuals with lower cognitive ability are less likely to be remunerated with performance-contingent incentives.

These results can also be related to the classical literature on contract theory. The flat segment of the contract from Proposition 5 iii) is reminiscent of the pooling that occurs in adverse selection models with countervailing incentives (Jullien, 2000; Lewis and Sappington, 1989; Maggi and Rodriguez-Clare, 1995), though it stems from different factors in this framework. In those models, this pooling arises because more efficient agents have higher outside options and require stronger incentives. As a result, intermediate types simultaneously wish to mimic an efficient type, in order to receive a large transfer, and an inefficient type, in order to save on costs. This behavior makes pooling optimal for the principal when dealing with intermediate types (Laffont and Martimort (2002)[Ch.3.3]). In this framework, however, the flat segment at intermediate output levels arises due to likelihood insensitivity, which makes implementing incentives at those output levels ineffective.

# 6 Empirical Evidence

The theoretical framework developed in the previous sections yields more specific predictions when combined with stylized facts about probability weighting. The common finding from controlled laboratory experiments is that most individuals exhibit probability weighting due to pessimism and likelihood insensitivity (Fehr-Duda and Epper, 2011; Wakker, 2010). Accordingly, Propositions 2, 3, and 5 predict that agents are given contracts that are either fully performance-insensitive or include substantial segments in which pay does not depend on performance. Moreover, the model also predicts that the extent to which the offered contracts are performance-insensitive increases with the worker's likelihood insensitivity and pessimism (Corollaries 2 and 6).

In the following, I use data from the American Life Panel (ALP) to evaluate the model's prediction that stronger probability weighting due pessimism or likelihood insensitivity, leads to compensation modalities that depend less on performance. The ALP is a representative sample of American households that is regularly administered surveys. I primarily use data from two surveys. First, the incentivized experiment conducted by Dimmock et al. (2021), who elicited the probability weighting functions of respondent.s The original goal of this experiment was to analyze the relationship between household portfolio diversification and probability weighting. Second, I draw on the data from the survey on American Working Conditions which was conducted in 2015 by Maestas et al. (2023). These data contain, among others, information about the type of compensation given to respondents at their current job.

Dimmock et al. (2021) elicited probability weighting functions of each respondent using the method of Abdellaoui (2000). That method has the ability to elicit the utility and probability weighting functions in a non-parametric way. This is achieved by implementing a set of binary lotteries that keep probabilities fixed, in order to elicit utility function curvature, and another set of binary lotteries that keep outcomes fixed and vary probabilities, in order to elicit probability weighting function curvature. Therefore, these data successfully identify these two components of risk attitude in the case of RDU preferences.

A disadvantage of Dimmock et al. (2021)'s elicitation in the context of the present study, is that it confounds probability weighting due to likelihood insensitivity with probability weighting due to pessimism/optimism. To deal with this limitation, I fit the respondent's answers to the questions designed to elicit probability weighting functions to the parametric from of probability weighting due to Chateauneuf et al. (2007). Formally, for each respondent $i$, the following function is estimated:

$$w(p_{ij}) = \begin{cases} 0 & \text{if } p = 0, \\ c_i + s_i \cdot p_{ij} & \text{if } p \in (0,1), \\ 1 & \text{if } p = 1. \end{cases} \tag{5}$$

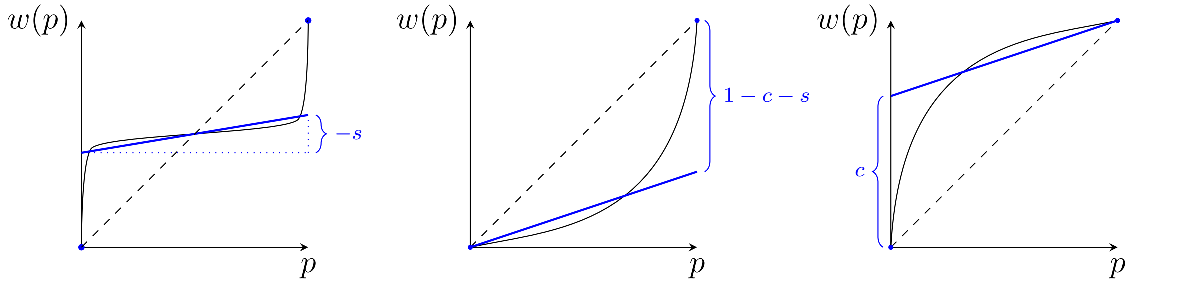
where the index $j$ represents the questions designed to elicit probability weighting.

Equation (5) demonstrates that Chateauneuf et al. (2007)'s proposal is a linear approximation of probability weighting, where deviations from EUT arise from the weights assigned to extreme events. Although this function does not satisfy all the properties stated in Assumption 4, it is recommended over other parametric forms of probability weighting for constructing indexes of likelihood insensitivity and pessimism/optimism because its parameters have a clean and simple interpretation (Wakker, 2010). Specifically, Chateauneuf et al. (2007)'s proposal cannot comply with Assumption 4 iii) since it exhibits $w''(p) = 0$ for all $p$. Consequently, pessimism and likelihood insensitivity are not captured through a convex, or an inverse-S shape of the weighting function. Instead, these phenomena are directly reflected in the magnitudes of $s_i$ and $c_i$, as will be explained below.

To estimate the parameters $c_i$ and $s_i$ in (5) I used non-linear least squares, a method that has been widely used to estimate the parameters of the probability weighting function (Abdellaoui et al., 2011; Baillon et al., 2018a; Dimmock et al., 2021). Importantly, the resulting estimate $\hat{s}_i$ in (5) captures the respondent $i$'s likelihood insensitivity (Wakker, 2010). In particular, the closer $\hat{s}_i$ is to $0$, the more insensitive the respondent is to changes in interior probabilities, and, conversely, a value of $\hat{s}_i$ closer to $1$ implies a perception of probabilities closer to EUT. Thus, I use throughout $-\hat{s}_i$ (if $\hat{s}_i < 1$) as a continuous index of likelihood insensitivity that I refer to as "Insensitivity." Notice that higher values of this index imply stronger likelihood insensitivity. Figure 4a provides an illustrative example of how this index captures likelihood insensitivity.

Furthermore, the expression $\frac{2\hat{c}_i + \hat{s}_i}{2}$, which depends both on $c_i$ and $s_i$, measures the

Figure 4: Examples of neo-additive probability weighting functions




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

Note: Black curves represent probability weighting functions. Blue lines represent their corresponding linear approximations. Dashed black lines represent accurate perception of probability.

extent to which $i$ exhibits pessimism (Wakker, 2010). That expression compares the extent to which a respondent overweights probabilities associated to the worst outcome (due to pessimism) to the extent to which they overweight probabilities associated to the best outcome (due to optimism).[17]. Figures 4c and 4b illustrate how Chateauneuf et al. (2007)'s proposal captures the overweighting of those probabilities. Consequently, higher values of $\frac{2\hat{c}_i+\hat{s}_i}{2}$ reflect stronger pessimism. Throughout, I refer to this continuous index of pessimism as "Pessimism."

Apart from probability weighting, I also estimate each respondent's utility function. The survey questions designed to elicit utility curvature are used to estimate the following functional:

$$u(x_{ik}) = x_{ik}^{1-\gamma_i}, \tag{6}$$

where the index $k$ represents the questions designed to elicit utility curvature. The parameter $\gamma_i$ is estimated using non-linear least squares, and is performed jointly, i.e. as part of the same procedure, as the estimation of parameters of probability weighting.

Table 1 presents descriptive statistics of $\hat{c}_i$ and $\hat{s}_i$. The results indicate that respondents exhibit likelihood insensitivity on average since the mean value of $\hat{s}_i$ is less than 1. They also exhibit pessimism on average, since $1-\hat{c}_i-\hat{s}_i < 1$ holds for the mean values of $\hat{c}_i$ and $\hat{s}_i$. Figure 5 illustrates the average probability weighting function and corroborates that the average respondent suffers from probability weighting due to pessimism and insensitivity.

These findings are further corroborated by analyzing the estimates at the individual level. A majority of respondents, 2012 out of 2640 (approximately 76% of respondents),

[17]As $p$ approaches 1 the weighting function becomes $w(p) \approx 1-c-s$. In contrast, as $p$ approaches 0, it becomes $w(p) \approx c$. Hence, comparing a respondent's level of pessimism to his level of optimism is equivalent to computing the difference $(1-c-s)-c$. Under EUT, that difference is equal to 0. Thus, $(1-c-s)-c = 0 \Leftrightarrow 2c+s = 1$. The index $\frac{2c+s}{2}$ is a linear transformation of that equality

Table 1: Estimates of the Probability Weighting Function

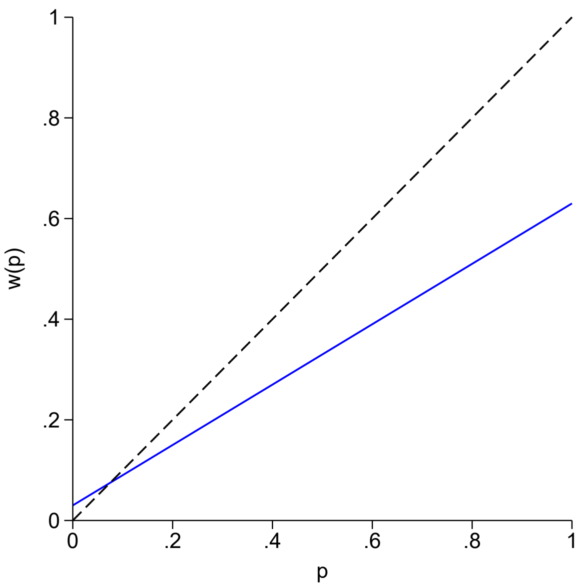
| | **Chateauneuf et al. (2007)** | |
|---|---|---|
| | $\hat{s}_i$ | $\hat{c}_i$ |
| **Mean** | 0.594 | 0.028 |
| **25th perc.** | 0.257 | -0.118 |
| **50th perc.** | 0.611 | 0.001 |
| **75th perc.** | 0.891 | 0.056 |
| **St. Dev.** | 0.358 | 0.067 |
| **N** | 2640 | 2640 |

This table presents the descriptive statistics for estimates of probability weighting obtained at the respondent level when the parametric form due to Chateauneuf et al. (2007), is assumed:

$$w(p_{ij}) = \begin{cases} 0 & \text{if } p = 0, \\ c_i + s_i \cdot p_{ij} & \text{if } p \in (0,1), \\ 1 & \text{if } p = 1. \end{cases}$$

The coefficients $c_i$ and $s_i$ were estimated using non-linear least squares.

Figure 5: Average Probability Weighting Function using Chateauneuf et al. (2007)




Note: The blue lines represent the average probability weighting function in the sample while the dashed lines represent the accurate perception of probabilities benchmark.

exhibit $\hat{s}_i < 1$, which indicates likelihood insensitivity. Furthermore, a majority of respondents, 1872 out of 2640 (approximately 70% of respondents), exhibit $1-\hat{c}_i-\hat{s}_i < 1$, which indicates pessimism. These results are consistent with previous experimental findings (Abdellaoui, 2000; Abdellaoui et al., 2011; Bruhin et al., 2010; l'Haridon and Vieider, 2019).

A subset of these respondents also answered questions regarding the type of compensation received in their current job. These questions allow me to investigate the relationship between incentives and the previously described indices of pessimism and likelihood insensitivity. Specifically, the analyses include the following variables: "Base Salary" measuring if the respondent's compensation is a fixed amount regardless of performance; "Piece rate or Productivity", which captures whether the respondent is paid according to a piece rate or productivity payments; "Commissions", which indicates if the respondent reports being paid commissions for achieving a quota; "Company Performance", representing whether the respondent receives payments based on the performance of the company; and "Performance Pay", capturing if the respondent's compensation consisting of piece rates, productivity payments, commissions, or payments based on the company's performance. This last variable summarizes the extent to which the respondent receives compensations based on performance.

Table 2 presents the descriptive statistics of these variables. The results indicate that a majority of respondents receive earnings that are not based on performance. Specifically, 78 % of respondents report receiving a compensation consisting exclusively of a fixed salary, while only 18% report having earnings based either on piecerates, productivity payments, commissions, or their company's performance. These results are consistent with previous findings in the literature, which suggest that individuals are frequently compensated through simple contracts that are far less performance-sensitive than standard theory would predict, often taking the form of fixed salaries (Lazear and Oyer, 2007; Prendergast, 1999; Salanié, 2003).

Table 2: Descriptive Statistics of Self-Reported Compensation Types

| **Variable** | **Type** | **Mean** | **Median** | **St. Dev.** | **N** |
|---|---|---|---|---|---|
| Base Salary | Binary | 0.785 | 1 | 0.411 | 1313 |
| Performance Pay | Binary | 0.180 | 0 | 0.385 | 1308 |
| Piece rate or Productivity | Binary | 0.036 | 0 | 0 .186 | 1256 |
| Commissions | Binary | 0.017 | 0 | 0.129 | 1300 |
| Company Performance | Binary | 0.147 | 0 | 0.354 | 1275 |

This table presents descriptive statistics for the variables that capture the respondents' self-reported compensation modality.

Each variable representing the type of compensation received by the respondents is regressed on the indexes of probability weighting. The advantage of running separate regressions, where each variable presented in Table 2 serves as the dependent variable in turn, is that they capture different dimensions of performance-based pay. Hence, this approach offers insight into the specific contexts where the predictions of the theory can be validated. For example, Piece rate or Productivity and Company Performance capture different ways of compensating performance; the former focuses on an individual and absolute measure of performance while the latter focuses on the organization's performance, which might be less relevant for the purposes of this paper.

In all regression specifications, I control for the respondents' utility curvature to isolate the effect of probabilistic risk attitudes on the type of compensation. This approach is consistent with the model's prediction, as an empirically relevant relationship between probability weighting and incentives must emerge above and beyond the average influence of utility curvature. Moreover, in some specifications, I include other control variables that might moderate the relationship between probability weighting and performance compensation, such as the respondent's age, gender, ethnicity, level of education, income, and type of work.

Table 3 reports the marginal effects of Probit regressions. The results presented in columns (1) and (2) indicate that stronger pessimism is associated with a significantly higher average probability of being compensated with a salary. In particular, a small increase in the pessimism index is associated with a 37% average increase in the probability of receiving only a salary. Consistent with this finding, the estimates in columns (3) and (4) show that stronger pessimism significantly reduces the average probability of being compensated according to performance. Moreover, columns (5) to (10) show that this negative effect of pessimism on performance-based compensation arises both for compensation tied to individual performance and for compensation contingent on the company performance. Importantly, these relationships remain statistically significant even after the inclusion of control variables. Thus,, the results are overall consistent with the model's prediction that greater pessimism is associated with a broader use of performance-insensitive contracts.

Furthermore, the estimates reported in Table 3 indicate that greater likelihood insensitivity is associated with a significantly lower average probability of being compensated based on performance. Specifically, small increase in the index of likelihood insensitivity is associated with a 26% average decrease in the probability that compensation depends on performance. Notably, this effect holds for the variables Piece Rate or Productivity and Commissions, which capture payment according to individual performance, as well as for Company Performance, which captures payment based on organizational performance. Hence, when interpreted as affecting the extent to which the contract of the average respondent depends on performance, these results are consistent with the model's predictions and suggest that such a contract becomes considerably less performance-contingent as the agent exhibits more severe likelihood insensitivity.

Table 3: The Relationship between Chateauneuf et al. (2007)'s Probability Weighting Function and Type of Compensation

| | (1) Base Salary | (2) Base Salary | (3) Performance Pay | (4) Performance Pay | (5) Piece rate | (6) Piece rate | (7) Commissions | (8) Commissions | (9) Company Performance | (10) Company Performance |
|---|---|---|---|---|---|---|---|---|---|---|
| Insensitivty | 0.231** | 0.242** | -0.240** | -0.257*** | -0.076* | -0.100** | -0.079** | -0.112** | -0.185** | -0.197** |
| | (0.101) | (0.100) | (0.096) | (0.094) | (0.046) | (0.045) | (0.040) | (0.052) | (0.091) | (0.091) |
| Pessimism | 0.336* | 0.368** | -0.359** | -0.414** | -0.175** | -0.223*** | -0.105 | -0.155* | -0.289* | -0.336** |
| | (0.178) | (0.179) | (0.169) | (0.169) | (0.079) | (0.077) | (0.073) | (0.094) | (0.161) | (0.166) |
| S-shaped | -0.170* | -0.192** | 0.193** | 0.221*** | 0.091** | 0.113*** | 0.055 | 0.074 | 0.139* | 0.162* |
| | (0.092) | (0.091) | (0.087) | (0.085) | (0.039) | (0.039) | (0.037) | (0.046) | (0.083) | (0.083) |
| U. Curv | 0.003 | 0.001 | -0.003 | -0.001 | 0.003** | 0.004** | -0.001 | -0.001 | -0.002 | -0.001 |
| | (0.004) | (0.003) | (0.003) | (0.003) | (0.002) | (0.002) | (0.001) | (0.001) | (0.003) | (0.003) |
| Age | | -0.001 | | -0.001 | | -0.006** | | 0.000 | | -0.003 |
| | | (0.007) | | (0.007) | | (0.003) | | (0.003) | | (0.006) |
| $Age^2$ | | 0.023 | | -0.002 | | 0.058* | | -0.010 | | 0.016 |
| | | (0.077) | | (0.071) | | (0.034) | | (0.036) | | (0.066) |
| Female | | 0.072*** | | -0.080*** | | -0.021 | | 0.010 | | -0.070*** |
| | | (0.025) | | (0.023) | | (0.013) | | (0.011) | | (0.021) |
| # Members | | -0.008 | | 0.009 | | 0.001 | | -0.002 | | 0.012* |
| | | (0.009) | | (0.008) | | (0.005) | | (0.004) | | (0.008) |
| Income | | -0.009* | | 0.013*** | | -0.004** | | 0.004 | | 0.014*** |
| | | (0.005) | | (0.005) | | (0.002) | | (0.002) | | (0.005) |
| Low Education | | -0.008 | | 0.002 | | -0.009 | | 0.001 | | 0.044 |
| | | (0.095) | | (0.097) | | (0.041) | | (0.001) | | (0.083) |
| Ethnicity | NO | YES | NO | YES | NO | YES | NO | YES | NO | YES |
| Type of Work | NO | YES | NO | YES | NO | YES | NO | YES | NO | YES |
| Log-Likelihood | -674.558 | -615.953 | -608.706 | -543.036 | -187.202 | -166.948 | -107.658 | -92.238 | -525.399 | -460.092 |
| N | 1301 | 1267 | 1296 | 1261 | 1246 | 1132 | 1288 | 875 | 1264 | 1221 |

This table presents the marginal effects of a probit regression of the model $y_i = b_0 + b_1\text{Inverse-S}_i + b_2\text{ Opt./Pess.}_i + b_3\text{U.curv}_i.\text{+Controls}_i'\Gamma + \varepsilon_i$. The variable $y_i$ captures the respondent's $i$ self-reported type of compensation. It can be one of the following variables: " Base Salary" ,"Performance Pay", "Piece rate", "Commissions", or "Company Performance'. "Insensitivity" is the respondent $i$'s index of likelihood insensitivity obtained from an estimation of Chateauneuf et al. (2007)'s probability weighting function. "Pessimism" is the respondent's $i$'s index of optimism and pessimism obtained from an estimation of Chateauneuf et al. (2007)'s probability weighting function. "U.curv" is the respondent $i$'s curvature of the utility function obtained from an estimation of a CRRA utility. Robust standard errors are presented in parentheses. *** denotes significance at the 0.01 level, ** denotes significance at the 0.05 level, * denotes significance at the 0.1 level.

It should be emphasized that another interpretation of these results is that stronger likelihood insensitivity leads to a higher proportion of respondents receiving contracts that are entirely performance-insensitive. However, the theoretical framework cannot accommodate this interpretation, as the optimal contracts under likelihood insensitivity depend on performance at least at the most extreme performance levels. Thus, alternative explanations may be required to reconcile this interpretation of the empirical findings with the theory. For example, greater insensitivity might be associated with other behavioral biases, such as ambiguity aversion and loss aversion, which—as shown in Online Appendix C—can rationalize this result. Alternatively, greater insensitivity might strengthen limited liability constraints, inducing the principal to forgo implementing incentives even at extreme performance levels.[18]

All in all, the empirical evidence is consistent with the predictions of the model. The empirically-relevant components of probability weighting, likelihood insensitivity and pessimism, significantly decrease the probability that compensation depends on performance. Thus, probability weighting emerges as a novel and empirically validated explanation for the widespread usage of performance-insensitive and simple contracts by firms and organizations.

# 7 Extensions

In this section, I briefly discuss some extensions of the theoretical framework and describe how they can easily be accommodated in the model. The section has two parts: the first primarily focuses on preferences and incorporates other deviations from expected utility, such as loss aversion and ambiguity attitudes, into the model; the second examines robustness.

## 7.1 Preferences

### 7.1.1 Reference Dependence and Loss Aversion

I enrich the agent's risk preferences by considering Cumulative Prospect Theory (CPT henceforth; Tversky and Kahneman, 1992). The main departure of CPT from RDU is that the agent evaluates the transfers in a contract relative to a reference point $R \geq 0$, which can change across different decisions (Wakker, 2010). Transfers that fall below this reference point are considered *losses*, while those above it are considered as *gains*.

[18] Another possible explanation is that, when insensitivity is sufficiently pronounced, the principal may attempt to concentrate incentives around the performance thresholds $\underline{q}$ and $\bar{q}$. However, since achieving these performance levels is highly unlikely, these incentives may appear excessively extreme—leading to the abandonment of incentive compatibility altogether.

For the sake of brevity, the detailed results of the analyses are relegated to Online Appendix B. The main finding of this extension is that the optimal contract consists of two parts. The first, located at the lower end of the output space, involves a fixed transfer equal to the agent's reference point. This payment protects the agent from losses when poor performance is realized. The second, which applies in the remainder of the output space, is fully determined by the shape of the probability weighting function, and will have the same shape as one of the contracts presented in Sections 4 and 5. As a result, the optimal contracts derived from this extension are similar to those in the standard model, with the key difference that they contain a larger performance-insensitive segment. A feature that make these theoretical predictions even more realistic and closer to the contracts used by organizations.

#### 7.1.2 Ambiguity

The theoretical framework can be easily extended to capture deviations from EUT due to attitudes toward ambiguity. This can be achieved by considering a setting in which the distribution $F(q|e)$ is unknown to the agent because, for example, he has limited experience with the delegated task. I characterize the agent's preferences in this setting of ambiguous probabilities with *Source Theory* (Abdellaoui et al., 2011; Baillon et al., 2025), which, broadly speaking, states that under ambiguity the phenomena of risk are amplified because there is "additional probability weighting." Importantly, this additional probability weighting captures ambiguity attitudes. For example, when it enhances the weight given to obtaining the lowest performance level and decreases the weights given to all other events, the individual exhibits ambiguity aversion.

The full discussion of the model with ambiguity is given in Online Appendix C. The main conclusion of that discussion is that ambiguity attitudes due to ambiguity aversion or a-insensitivity (the analogue of likelihood insensitivity under ambiguity) generate larger segments in which the optimal contract is performance-insensitive. Again, this feature makes the theoretical predictions more realistic and closer to the contracts used by organizations in real-life.

### 7.2 Robustness

#### 7.2.1 Continuous effort

In Online Appendix D , I solve a version of the model in which effort is assumed to be continuous. Specifically, effort is assumed to be the variable $e \in [0, \bar{e}]$. The main results of this paper emerge in that more general setting under two conditions. First, the density function must be convex, i.e. $F_{ee}(q|e) > 0$, a well-known requirement for guaranteeing the validity of the first-order approach (Mirrlees, 1999; Rogerson, 1985).

Second, the probability weighting function must be everywhere convex, and is due to González-Jiménez (2024a). Under these two conditions, the contracts characterized in Section 4 are optimal and can be obtained with the first-order approach.

The second condition discussed in the previous paragraph is stringent. It implies that the findings of this paper do not hold when effort is continuous and the agent is optimistic or likelihood-insensitive. Therefore, I consider weaker requirements in order to validate the findings of the standard model with binary effort. Lemma D1, presented in Online Appendix D, shows that a sufficiently convex cost function $c(e)$ can ensure the validity of the first-order approach even when the probability weighting function is not convex everywhere. If that is the case, the optimal contracts characterized in Section 4 and 5 remain valid. Thus, a steeply convex cost function, representing a demanding task to the agent, validates the results of this paper.

#### 7.2.2 Stochastic contracts

The previous subsection concluded that relaxing the assumption of binary effort can render the contracts characterized in Sections 4 and 5 suboptimal. In Online Appendix D, I show that the optimal contracts in this case are stochastic, meaning they include risk other than that inherent in performance. Those contracts are desirable because the agent's probability weighting function is concave over a range of probabilities $p \in (0, 1)$, which induces risk-seeking attitudes. As a result, the optimal contract must incorporate greater risk than that implicit in the solutions presented in sections 4 and 5. This result is consistent with the findings of González-Jiménez (2024a).

In that Appendix, I extend and complement the findings of González-Jiménez (2024a) by fully characterizing the optimal stochastic contract $(r, t(q); 1 - r, 0)$. This is accomplished in two different ways. First, I use the results from a standard model in which the contract is restricted to be non-stochastic to determine the outcome $t(q)$. For instance, when the agent is optimistic, this outcome corresponds to a contract similar to that in Proposition 4. Second, I characterize the optimal probability, $r^*$ to be included in the contract.

#### 7.2.3 Adverse Selection and Moral Hazard

In Online Appendix E, I consider an extension of the model in which the assumption that the principal knows the agent's risk attitudes is relaxed. In that extension, the principal's goal is dual: to screen agents according to their risk preferences and to motivate them to exert a high level of effort. I model that setting as a framework of adverse selection followed by moral hazard.

The analyses in Online Appendix E show that the solution to that more convoluted problem consists of a menu of contracts with the following properties. First, to incentivize high effort, the contracts in the menu should adopt the shapes described in Section 5. These incentive schemes ensure high effort when the agent suffers from probability weighting due to likelihood insensitivity. Second, the menu should include as many contracts as there are types of agents. In the extension, this amounts to two contracts: one targeting the EUT agent and the other targeting the RDU agent with likelihood insensitivity. Lastly, the contract that targets the more efficient type, i.e. the EUT agent, is enriched with an additional transfer. This monetary incentive ensures that the EUT agent does not mimic the RDU agent, so as not to miss out on that payment. A menu of contracts with these properties ensures that agents self-select into the contract that best fits their risk preference, and that high effort is exerted.

Notice that while this extension only considers two types, its underlying rationale can be generalized to more complex settings. For example, suppose that the principal contracts with a continuum of agents who vary in their degree of likelihood insensitivity, while the principal remains uncertain about each agent's specific level of insensitivity.[19] The optimal solution in this case involves offering a menu of contracts, each tailored to a distinct type of agent. Moreover, the incentives in these contracts should follow the principles outlined in Proposition 5 and Corollary 6. Specifically, contracts designed for agents with greater likelihood insensitivity should increasingly concentrate incentives at extreme output levels. Furthermore, to ensure self-selection, these contracts must include lump-sum transfers that become larger as the insensitivity of the agent for whom the contract is intended decreases.

# 8 Conclusion

This paper examined the impact of probability weighting on optimal contracting in moral hazard setting. Its main finding is that probability weighting yields optimal contracts with substantial performance-insensitive segments, and in some cases, contracts that are entirely performance-insensitive. Consequently, the model predicts that the principal must implement simple, low-powered incentive schemes that resemble those commonly used by organizations, such as fixed salaries, lump-sum bonuses, and high-performance commissions. Using data from a representative sample of American households, I provide empirical support for the model's prediction that probability weighting increases the extent to which contracts become performance-insensitive. Taken together, these analyses highlight probability weighting as a preference-based explanation for the widespread use of simple contracts.

[19] To make this example more concrete, suppose that the agents' weighting function can be characterized by Prelec (1998)'s functional with $w(p) = \exp\big(-0.86(-(ln(p))^{\alpha}\big)$. The principal knows that the agents' weighting function have that shape and that she contracts with a continuum of agents in the support $\alpha \in [0, 1]$.

# A Proofs of Theoretical Results
# [FOR ONLINE PUBLICATION ONLY]

## A.1 Preliminary Results

### A.1.1 Results on the Properties of the Probability Weighting Function at Extremes

**Lemma A1.** *If* $\lim_{p\to 0^+} w'(p) = +\infty$, *then* $\lim_{p\to 0^+} w''(p) = -\infty$ *and* $\lim_{p\to 0^+} \frac{w''(p)}{w'(p)} = -\infty$.

*Proof.* Suppose that $\lim_{p\to 0^+} w'(p) = +\infty$ but, to set up the contradiction, also 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 $B > 0$, then $w''(p) > -B$. 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 that $\lim_{p\to 0^+} w'(p) = +\infty$ but that $\lim_{p\to 0^+} \frac{w''(p)}{w'(p)} \neq -\infty$. So for $p \in [0,\bar{p}]$ and $B > 0$, then $\frac{w''(p)}{w'(p)} > -B$. Integrating over $[p_0, p_1] \subseteq [0, \bar{p}]$ yields:

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

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 A2.** *If* $\lim_{p\to 1^-} w'(p) = +\infty$, *then* $\lim_{p\to 1^-} w''(p) = +\infty$ *and* $\lim_{p\to 1^-} \frac{w''(p)}{w'(p)} = +\infty$.

*Proof.* Suppose that $\lim_{p\to 1^-} w'(p) = +\infty$ but, to set up the contradiction, also that $\lim_{p\to 1^-} w''(p) \neq +\infty$. Hence, there exists $\underline{p} \in (0,1)$ such that, for $p \in \Big[\underline{p}, 1\Big]$ and $B > 0$, then $w''(p) < B$. 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 that $\lim_{p\to 1^-} w'(p) = +\infty$ but that $\lim_{p\to 1^-} \frac{w''(p)}{w'(p)} \neq +\infty$. Thus, for $p \in [\underline{p}, 1]$ and $B > 0$, then $\frac{w''(p)}{w'(p)} < B$. 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 A3.** *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 that $\lim_{p\to 1^-} w'(p) = 0$ but, to set up the contradiction, also that $\lim_{p\to 1^-} w''(p) \geq 0$. For$p \in [\underline{p}, 1]$ and $B \geq 0$, then $w''(p) \geq B$. 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 that $\lim_{p\to 1^-} w'(p) = 0$ but that $\lim_{p\to 1^-} \frac{w''(p)}{w'(p)} \neq -\infty$. Thus, for $p \in \left[\underline{p}, 1\right]$, then $\frac{w''(p)}{w'(p)} > -B$. 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 A4.** *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 that $\lim_{p\to 0^+} w'(p) = 0$ but, to set up the contradiction, also that $\lim_{p\to 0^+} w''(p) \leq 0$. Hence, for $p \in \left[0, \bar{p}\right]$ and $B \geq 0$, then $w''(p) \leq -B$. 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 also that $\lim_{p\to 0^+} \frac{w''(p)}{w'(p)} \neq +\infty$. So, for $p \in \left[0, \bar{p}\right]$ and $B > 0$, then $\frac{w''(p)}{w'(p)} \leq B$. 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 $\lim_{p\to 0^+} \frac{w''(p)}{w'(p)} = +\infty$. ■

### A.1.2 Results on Comparative Degrees of Optimism/Pessimism and Insensitivity

**Lemma A5.** *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

$$\frac{w_i''(p)}{w_i'(p)} = \frac{\theta''\Big(w_j(p)\Big)}{\theta'\Big(w_j(p)\Big)} w_j'(p) + \frac{w_j''(p)}{w_j'(p)}. \tag{A.1}$$

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

$$-\frac{w_i''(p)}{w_i'(p)} > -\frac{w_j''(p)}{w_j'(p)}. \tag{A.2}$$

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 (A.2) over $[p_0, p_1]$ to obtain:

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

Integrating the resulting expression over the range of $p_0$ gives:

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

Integrating again but now over the range of $p_1$ gives:

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

The step in the last equivalence is due to the fact that $p_0$ 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 4, $w_i(0) = w_j(0)$ and $w_i(1) = w_j(1)$. Hence, $\int_0^1 w_j'(p)\mathrm{d}p = w_j(1) - w_j(0) = 1 > \int_0^1 w_j'(p)\mathrm{d}p$. Contradicting the assumption that $w_i(1) = 1$. A similar rationale disregards $w_i'(p) > w_j'(p)$ for all $p \in (0,1)$. Hence, if $w_i'(p) \leq w_j(p)$ holds, it must do so for some segment in $p \in (0,1)$.

Let $w_J(p) := \eta\big(w_j(p)\big)$ where $\eta$ is a concave, increasing, and continuous probability weighting function. Lemma A1 and Lemma A3 show that $\lim_{p\to 0^+} w'(p) = +\infty$ and $\lim_{p\to 1^-} w'(p) = 0$ for a generic weighting function $w$. The first part of this Lemma implies that $-\frac{w_J''(p)}{w_J'(p)} > -\frac{w_j''(p)}{w_j'(p)}$ $\forall p \in (0,1)$. Therefore, $w_J'(p)$ tends to infinity faster than $w_j'(p)$ as $p \to 0^+$.

Assumption 4 states that, under optimism, $w'(p)$ is decreasing and continuous. These properties together with $-\frac{w_J''(p)}{w_J'(p)} > -\frac{w_j''(p)}{w_j'(p)}$ $\forall p \in (0,1)$, that $w_J'(p)$ tends to infinity faster than $w_j(p)$ as $p \to 0^+$, and the fact that $\lim_{p\to 1^-} w'(p) = 0$, imply that there exists a unique point $p_k \in (0,1)$ such that $w'_J(p_k) = w'_j(p_k)$. For $p < p_k$ then $w'_J(p) > w'_j(p)$ but instead $w'_J(p) < w'_j(p)$ if $p > p_k$.

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

**Lemma A6.** *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 $p < 0.5$. Since $w_i(p) = \phi\big(w_j(p)\big)$,

$$\frac{w''_i(p)}{w'_i(p)} = \frac{\phi''\big(w_j(p)\big)}{\phi'\big(w_j(p)\big)} w'_j(p) + \frac{w''_j(p)}{w'_j(p)}. \tag{A.3}$$

Due to $\phi'' < 0$ in $p < 0.5$, 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 $p > 0.5$.

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

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

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

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

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

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

A similar procedure gives that when $i$ is more likelihood insensitive than $j$ in the sense of Definition 5, then $w_i(p) < w_j(p)$ in $p > 0.5$.

**Part 3.** Suppose that $w_i'(p) > w_j'(p)$ for all $p < 0.5$. Note that $\int_0^{p_1} w_i'(p)\mathrm{d}p > \int_0^{p_1} w_i'(p)\mathrm{d}p \Leftrightarrow w_i(p_1) > w_j(p_1)$ for arbitrary $p_1 \in (0, 0.5)$, which is in line with the first part of the Lemma. However, 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 $p < 0.5$ is first assumed. Hence, it must be that $w_i'(p) > w_j'(p)$ *for some* $p \in (0, 0.5)$.

Assumption 4 states that $w'(p)$ is decreasing in $p < 0.5$. Moreover, Lemma A1 shows that $\lim_{p \to 0^+} w'(p) = +\infty$. Let $w_J(p) := \eta\big(w_j(p)\big)$, where $\eta$ is an inverse-S probability weighting function. Accordingly, $-\frac{w_J''(p)}{w_J'(p)} > -\frac{w_j''(p)}{w_j'(p)}$ in $p < 0.5$ as shown in the first part of this Lemma. Therefore, the function $w'_J(p)$ tends to infinity faster than $w_j'(p)$ as $p \to 0^+$.

Due to the continuity of $w'(p)$, $w'(p)$ being decreasing in $p$ in $p < 0.5$, $-\frac{w_J''(p)}{w_J'(p)} > -\frac{w_j''(p)}{w_j'(p)}$ in $p < 0.5$, $w'_J(p)$ tending to infinity faster than $w_j'(p)$ as $p \to 0^+$ and the fact that $\lim_{p \to 0.5} w'(p) = \min\{w'(p)\}$, there exists a unique point $p_k \in (0, 0.5)$ 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 $0.5 > p > p_k$.

Next, let $w_i := \phi\big(w_J(p)\big)$ where $\phi$ is an inverse-S probability weighting function. Then, $w_i'(p)$ is a decreasing function in $p < 0.5$, exhibits $-\frac{w_i''(p)}{w_i'(p)} > -\frac{w_J''(p)}{w_J'(p)}$ for all $p < 0.5$, tends to infinity faster than $w'_J(p)$ as $p \to 0^+$, and converges to $\min\{w'(p)\}$ as $p \to 0.5$. Therefore, the point $p_l$ such that $w_i'(p_l) = w'_J(p_l)$ is such that $p_l < p_k < 0.5$.

**Part 4.** Suppose that $w_i'(p) < w_j'(p)$ for all $p > 0.5$. For arbitrary $p_1 \in (\tilde{p}, \bar{p})$, I obtain that $\int_{p_1}^1 w_i'(p)\mathrm{d}p < \int_{p_1}^1 w_i'(p)\mathrm{d}p \Leftrightarrow w_i(p_1) < w_j(p_1)$ which corroborates the first part of the Lemma. However, we also have $\int_{\tilde{p}}^{p_1} w_i'(p)\mathrm{d}p > \int_{\tilde{p}}^{p_1} w_i'(p)\mathrm{d}p \Leftrightarrow w_i(p_1) < w_j(p_1)$, contradicting the first part of the lemma. We find a similar contradiction $w_i'(p) > w_j'(p)$ is assumed. Hence, it must be that $w_i'(p) < w_j'(p)$ holds for some $p > \tilde{p}$.

Assumption 4 states that $w'(p)$ is increasing in $p > \tilde{p}$. Moreover, Lemma A2 shows that $\lim_{p \to 1^-} w'(p) = +\infty$. Let $w_J(p) := \eta\big(w_j(p)\big)$ where $\eta$ is an inverse-S probability weighting function. Accordingly, $\frac{w_J''(p)}{w_J'(p)} > \frac{w_j''(p)}{w_j'(p)}$ in $p > \tilde{p}$ as shown in the first part of this Lemma. Therefore, $w_j'(p)$ tends to infinity faster than $w_J(p)$ as $p \to 1^-$.

Due to the continuity of $w'(p)$, $w'(p)$ being increasing in $p$ in $p > \tilde{p}$, $\frac{w_J''(p)}{w_J'(p)} > \frac{w_j''(p)}{w_j'(p)}$ in $p > \tilde{p}$, $\lim_{p \to \tilde{p}} w'(p) = \min\{w'(p)\}$, and the fact that $w_j'(p)$ tends to infinity faster than $w_J(p)$ as $p \to 1^-$, there exists a unique point $p_m \in (0, 1)$ such that $w'_J(p_m) = w_j'(p_m)$. If $p < p_m$, then $w'_J(p) < w_j'(p)$. Instead, if $p > p_m > 0.5$, then $w'_J(p) > w_j'(p)$.

Next, Let $w_i := \phi\big(w_J(p)\big)$ where $\phi$ is an inverse-S probability weighting function. Then, $w_i'$ is an increasing function in $p > 0.5$, exhibits $\frac{w_i''(p)}{w_i'(p)} > \frac{w_J''(p)}{w_J'(p)}$ for all $p > 0.5$, $w'_J(p)$ tends to infinity faster than $w_i'(p)$ as $p \to 1^+$, and converges to $\min\{w'(p)\}$ as $p \to 0.5$. Hence, the point $p_n$ such that $w_i'(p_n) = w'_J(p_n)$ is such that $p_n > p_m > 0.5$.

■

## A.2 Proofs of the Main Results

### Lemma 1

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

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

$$\Leftrightarrow w'\big(1-F(q_1|\underline{e})\big)f(q_1|\underline{e})w'\big(1-F(q_0|\bar{e})\big)f(q_0|\bar{e}) \leq w'\big(1-F(q_0|\underline{e})\big)f(q_0|\underline{e})w'\big(1-F(q_1|\bar{e})\big)f(q_1|\bar{e}). \tag{A.4}$$

Integrating both sides of the inequality with respect to $q_0$, which takes values from $\underline{q}$ to $q_1$, gives:

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

After rearranging and using $\int_{\underline{q}}^{q_1} w'\big(1-F(q_0|e)\big)f(q_0|e)\,\mathrm{d}q_0 = 1 - w\big(1-F(q_1|e)\big)$, we obtain:

$$\frac{w'\big(1-F(q_1|\underline{e})\big)f(q_1|\underline{e})}{w'\big(1-F(q_1|\bar{e})\big)f(q_1|\bar{e})} \leq \frac{1-w\big(1-F(q_1|\underline{e})\big)}{1-w\big(1-F(q_1|\bar{e})\big)}. \tag{A.5}$$

Integrating (A.4) again, but now with respect to $q_1$, which takes values from $q_0$ to $\bar{q}$, gives:

$$\frac{w\big(1-F(q_0|\underline{e})\big)}{w\big(1-F(q_0|\bar{e})\big)} \leq \frac{w'\big(1-F(q_0|\underline{e})\big)f(q_0|\underline{e})}{w'\big(1-F(q_0|\bar{e})\big)f(q_0|\bar{e})}. \tag{A.6}$$

Letting $q_0 = q_1 = q$ and combining (A.5) and (A.6) yields:

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

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

which proves the first part of the Lemma.

**Part 2**. We have:

$$\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{w'\Big(1-F(q|\underline{e})\Big)f(q|\underline{e})}{w'\Big(1-F(q|\bar{e})\Big)f(q|\bar{e})}\right) = \frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{w'\Big(1-F(q|\underline{e})\Big)}{w'\Big(1-F(q|\bar{e})\Big)}\right)\frac{f(q|\underline{e})}{f(q|\bar{e})} + \frac{w'\Big(1-F(q|\underline{e})\Big)}{w'\Big(1-F(q|\bar{e})\Big)}\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{f(q|\underline{e})}{f(q|\bar{e})}\right). \tag{A.7}$$

The W-MLRP, $\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{w'\big(1-F(q|\underline{e})\big)f(q|\underline{e})}{w'\big(1-F(q|\bar{e})\big)f(q|\bar{e})}\right) < 0$, and equation (A.7) imply:

$$\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{f(q|\underline{e})}{f(q|\bar{e})}\right) < -\frac{\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{w'\big(1-F(q|\underline{e})\big)}{w'\big(1-F(q|\bar{e})\big)}\right)\frac{f(q|\underline{e})}{f(q|\bar{e})}}{\frac{w'\big(1-F(q|\underline{e})\big)}{w'\big(1-F(q|\bar{e})\big)}} \tag{A.8}$$

The MLRP immediately holds if the right-hand side of the inequality in (A.8) is negative. Since $w'\Big(1-F(q|e)\Big) > 0$, this is true when $\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{w'\big(1-F(q|\underline{e})\big)}{w'\big(1-F(q|\bar{e})\big)}\right) < 0$. That condition can be rewritten as:

$$\frac{w''\Big(1-F(q|\bar{e})\Big)}{w'\Big(1-F(q|\bar{e})\Big)}f(q|\bar{e}) \leq \frac{w''\Big(1-F(q|\underline{e})\Big)}{w'\Big(1-F(q|\underline{e})\Big)}f(q|\underline{e}). \tag{A.9}$$

Note that the inequality in (A.9) can be rewritten again as:

$$-\frac{d\ln\Big(w'(1-F(q|\bar{e}))\Big)}{dq} \leq -\frac{d\ln\Big(w'(1-F(q|\underline{e}))\Big)}{dq}. \tag{A.10}$$

Integrating both sides of the inequality in (A.10) with respect to $q$ gives:

$$w'\Big(1-F(q|\bar{e})\Big) \geq w'\Big(1-F(q|\underline{e})\Big). \tag{A.11}$$

■

**Lemma 2**

*Proof.* Equation (A.8) shows that under the MLRP, $\frac{\mathrm{d}}{\mathrm{d}q}\frac{f(q|\underline{e})}{f(q|\bar{e})} \leq 0$, it suffices for the WMLRP to hold if the inequality given in (A.9) holds. The Lemma follows immediately from that inequality.

■

The following Lemma will be useful for the proof of Proposition 2.

**Lemma A7.** *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 equations (A.8) and (A.9) . ■

**Proposition 1**

*Proof.* Denote by $t^r(q)$, the solution to the problem in (4) when the constraint (M) is not taken into account. The proof consists of investigating the conditions under which the solution to that "unconstrained" problem satisfies monotonicity.

Let $\nu$ be the Lagrange multiplier of the agent's PC constraint and $\mu$ that of the IC constraint. The Lagrangian of (4) without the (M) constraint writes as:

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

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

$$\begin{aligned}-f(q|\bar{e})+\mu\Big[u'(t^r(q))(w'(1-F(q|\bar{e}))f(q|\bar{e})-w'(1-F(q|\underline{e}))f(q|\underline{e}))\Big]\\+\nu\Big[u'(t^r(q))w'(1-F(q|\bar{e}))f(q|\bar{e})\Big]=0,\end{aligned}\tag{A.12}$$

and, after re-arranging,

$$\frac{1}{u'\Big(t^r(q)\Big)w'\Big(1-F(q|\bar{e})\Big)}=\nu+\mu\left(1-\frac{w'\Big(1-F(q|\underline{e})\Big)f(q|\underline{e})}{w'\Big(1-F(q|\bar{e})\Big)f(q|\bar{e})}\right).\tag{A.13}$$

Hence, the solution from the first-order approach without the M constraint can be written as:

$$t^r(q)=h\left(\frac{1}{w'\Big(1-F(q|\bar{e})\Big)\left(\nu+\mu\left(1-\frac{w'\big(1-F(q|\underline{e})\big)f(q|\underline{e})}{w'\big(1-F(q|\bar{e})\big)f(q|\bar{e})}\right)\right)}\right),\tag{A.14}$$

where $h$ is the inverse function of $u'$, i.e. $h(t)=u'^{-1}(t)$.

**Concavity of $w(p)$ implies monotonicity.** I first assume that $w''(p) < 0$. Differentiate the expression in (A.13) with respect to $q$, to obtain:

$$t^{r\prime}(q) = \frac{u'(t^r(q))w''(1 - F(q|\bar{e}))}{u''(t^r(q))w'(1 - F(q|\bar{e}))} f(q|\bar{e}) + \mu \frac{w'(1 - F(q|\bar{e}))\, u'(t^r(q))^2}{u''(t^r(q))} \frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{w'\big(1 - F(q|\underline{e})\big) f(q|\underline{e})}{w'\big(1 - F(q|\bar{e})\big) f(q|\bar{e})}\right). \tag{A.15}$$

Since $\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{w'(1-F(q|\underline{e}))f(q|\underline{e})}{w'(1-F(q|\bar{e}))f(q|\bar{e})}\right) < 0$ (Assumption 5), $u'(t^{sb}(q)) > 0$, $u''(t^{sb}(q)) < 0$ (Assumption 3), and $w'(p) > 0$ (Assumption 4), the first and second terms on the right-hand side of (A.15) must be positive. Hence, $w''(p) < 0 \Rightarrow t^{r\prime}(q) > 0$ for all $q$, which, in turn, implies $t^r(q_2) \geq t^r(q_1)$ for any $q_1, q_2 \in [\underline{q}, \bar{q}]$ such that $q_2 > q_1$.

**Monotonicity implies concavity of $w(p)$.** Now assume that $t^r(q_2) > t^r(q_1)$ for any $q_2, q_1 \in [\underline{q}, \bar{q}]$ such that $q_2 > q_1$. Using equation (A.14) we get:

$$t^r(q_2) > t^r(q_1) \Leftrightarrow h\left(\frac{1}{w'\big(1 - F(q_2|\bar{e})\big)\left(\nu + \mu\left(1 - \frac{w'\big(1-F(q_2|\underline{e})\big)f(q_2|\underline{e})}{w'\big(1-F(q_2|\bar{e})\big)f(q_2|\bar{e})}\right)\right)}\right) > h\left(\frac{1}{w'\big(1 - F(q_1|\bar{e})\big)\left(\nu + \mu\left(1 - \frac{w'\big(1-F(q_1|\underline{e})\big)f(q|\underline{e})}{w'\big(1-F(q_1|\bar{e})\big)f(q_1|\bar{e})}\right)\right)}\right). \tag{A.16}$$

Since $h$ is decreasing in $t$, the inequality given above can be rewritten as:

$$\nu\Big(w'\big(1 - F(q_2|\bar{e})\big) - w'\big(1 - F(q_1|\bar{e})\big)\Big) + \mu\left(\left(1 - \frac{w'\big(1 - F(q_2|\underline{e})\big)f(q_2|\underline{e})}{w'\big(1 - F(q_2|\bar{e})\big)f(q_2|\underline{e})}\right) w'\big(1 - F(q_2|\bar{e})\big) - \left(1 - \frac{w'\big(1 - F(q_1|\underline{e})\big)f(q_1|\underline{e})}{w'\big(1 - F(q_1|\bar{e})\big)f(q_1|\underline{e})}\right) w'\big(1 - F(q_1|\bar{e})\big)\right) > 0 \tag{A.17}$$

The inequality in equation (A.17) holds when $w'\big(1 - F(q_2|\bar{e})\big) > w'\big(1 - F(q_1|\bar{e})\big)$ for all $q$ and $e$, which is guaranteed if $w''\big(1 - F(q|e)\big) < 0$ for any $q$ and $e$. In that case, the first expression in equation (A.17) is positive and, because $\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{w'(1-F(q|\underline{e}))f(q|\underline{e})}{w'(1-F(q|\bar{e}))f(q|\bar{e})}\right) < 0$ (Assumption 5), the second expression in that inequality is also positive. Hence, $t^{r\prime}(q) > 0 \Rightarrow w''(p) < 0$.

■

**Corollary 1**

*Proof.* We now examine the conditions under which the (M) constraint binds. We do so by investigating the segments at which $t^r(q)$, the solution to the unconstrained problem, does not comply with monotonicity. To that end, rewrite (A.15) as:

$$t^{r\prime}(q) = \frac{u'(t^r(q))}{u''(t^r(q))}\left(\frac{w''(1-F(q|\bar{e}))}{w'(1-F(q|\bar{e}))}f(q|\bar{e}) + \mu w'(1-F(q|\bar{e}))\,u'(t^r(q))\,\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{w'\big(1-F(q|\underline{e})\big)f(q|\underline{e})}{w'\big(1-F(q|\bar{e})\big)f(q|\bar{e})}\right)\right). \tag{A.18}$$

Since $u' > 0$ and $u'' < 0$ (Assumption 3), $t^{r\prime}(q) < 0$ is obtained when:

$$\frac{w''(1-F(q|\bar{e}))}{w'(1-F(q|\bar{e}))}f(q|\bar{e}) + \mu w'(1-F(q|\bar{e}))\,u'(t^r(q))\,\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{w'\big(1-F(q|\underline{e})\big)f(q|\underline{e})}{w'\big(1-F(q|\bar{e})\big)f(q|\bar{e})}\right) > 0.$$

$$\Leftrightarrow \frac{w''(1-F(q|\bar{e}))}{w'(1-F(q|\bar{e}))}f(q|\bar{e}) > -\mu w'(1-F(q|\bar{e}))\,u'(t^r(q))\,\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{w'\big(1-F(q|\underline{e})\big)f(q|\underline{e})}{w'\big(1-F(q|\bar{e})\big)f(q|\bar{e})}\right). \tag{A.19}$$

The M constraint is violated when the inequality given in the equation above holds.

To consider the first-best case (the case without the IC constraint) let $\mu = 0$. Equation (A.19) becomes:

$$\frac{w''(1-F(q|\bar{e}))}{w'(1-F(q|\bar{e}))}f(q|\bar{e}) > 0. \tag{A.20}$$

■

**Proposition 2 & Proposition 4**

*Proof.* To solve the problem, I characterize the optimal contract as the solution to the "relaxed" problem, given by (A.13), when the (M) constraint does not bind according to Proposition 1 and/or Corollary 1. In contrast, when that constraint binds, the optimal contract must, due to complementary slackness, flatten out, i.e. exhibit $t(q_2) = t(q_1)$ for any $q_2, q_1$ such that $q_2 > q_1$.

**Part 1. Optimal Contract without the M and IC constraints.** Let $\mu = 0$. Denote by $t^{fb}(q)$ the solution to the problem in (4) without the M and IC constraints. This solution is derived from (A.12) and can be written as:

$$\frac{1}{u'\big(t^{fb}(q)\big)w'\big(1-F(q|\bar{e})\big)} = \nu. \tag{A.21}$$

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

To investigate the shape of $t^{fb}(q)$, the restiction $\mu = 0$ is included in (A.15), giving us:

$$t^{fb'}(q) = \frac{u'\big(t^{fb}(q)\big)}{u''\big(t^{fb}(q)\big)} \frac{w''\big(1-F(q|\bar{e})\big)}{w'\big(1-F(q|\bar{e})\big)} f(q|\bar{e}). \tag{A.22}$$

If the agent exhibits optimism, we have $w'(p) > 0$ and $w''(p) < 0$ for all $p$ (Assumption 4 and Definition 2). Under these conditions, the right-hand side of (A.22) is positive, implying that the first-best contract given to the optimist, $t_O^{fb}(q)$, is everywhere increasing in $q$. This is in line with Proposition 1.

If the agent exhibits pessimism, we have $w'(p) > 0$ and $w''(p) > 0$ for all $p$ (Assumption 4 and Definition 1). According to the condition in (A.20) from Corollary 1, the M constraint binds at the optimum for all probabilities. This is evident from equation (A.22) which is strictly negative, implying that the solution to the relaxed problem is strictly decreasing in $q$ for all performance levels. Therefore, by complementary slackness it must be that the optimal contract exhibits $t'(q) = 0$ for all $q$.

**Part 2. Incentive constraint can be Slack at the Optimum under Optimism** Denote by $t^{sb}(q)$ the solution described by (A.13). I first show that $\mu > 0$ might not always hold the optimum. Suppose that $\mu = 0$. Then $t^{sb}(q) = t^{fb}(q)$, where $t^{fb}(q)$ is the contract satisfying (A.21).

**Optimism** Consider the case of an agent with optimism in the sense of Definition 2. From the complementary slackness condition from $\mu = 0$, we get:

$$\int_{\underline{q}}^{\bar{q}} u\big(t_O^{fb}(q)\big) w'(1-F(q|\bar{e})) f(q|\bar{e})\, \mathrm{d}q - c > \int_{\underline{q}}^{\bar{q}} u\big(t_O^{fb}(q)\big) w'(1-F(q|\underline{e})) f(q|\underline{e})\, \mathrm{d}q. \tag{A.23}$$

Integration by parts of (3) gives:

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

which is used to rewrite (A.23) as:

$$\int_{\underline{q}}^{\bar{q}} u'\big(t_O^{fb}(q)\big) t_O^{fb'}(q) \Big( w\big(1-F(q|\bar{e})\big) - w\big(1-F(q|\underline{e})\big)\Big)\, \mathrm{d}q > c. \tag{A.24}$$

According to Lemma 1, the W-MLRP (Assumption 5) implies $w\Big(1-F(q|\bar{e})\Big) \geq w\Big(1-F(q|\underline{e})\Big)$, which, together with $t_O^{fb\prime}(q) > 0$ for the optimist (equation (A.22)), imply that the left-hand side of (A.24) is weakly positive and that the inequality in that equation can hold.

The remainder of this part of the proof varies $c$. Notice that these changes affect $t_O^{fb}(q)$ through the P constraint. Accordingly, smaller values of $c$ leads to higher levels of $t_O^{fb}(q)$. Notably, in this case, the right-hand side of (A.24) can become larger even though $t_O^{fb}(q)$ becomes smaller because $u'$ is a decreasing function ($u'' < 0$). .

Hence, there must exist a threshold cost $\hat{c}_O > 0$ such that if $c \leq \hat{c}_O$, then the inequality in (A.24) holds and $t_O^{sb}(q) = t_O^{fb}(q)$. On the other hand, if $c > \hat{c}_O$, then the inequality in (A.24) does not hold and it must be that $\mu > 0$. In the latter case, $t_O^{sb}(q)$ satisfies (A.13).

**Pessimism** Now consider the case of an agent with pessimism in the sense of Definition 1. From the complementary slackness condition corresponding to $\mu = 0$ we get:

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_P^{fb}(q)\Big) w'(1-F(q|\bar{e})) f(q|\bar{e})\,\mathrm{d}q - c > \int_{\underline{q}}^{\bar{q}} u\Big(t_P^{fb}(q)\Big) w'(1-F(q|\underline{e})) f(q|\underline{e})\,\mathrm{d}q$$
$$\Leftrightarrow u\Big(t_P^{fb}(q)\Big) - c > u\Big(t_P^{fb}(q)\Big)$$
$$\Leftrightarrow -c > 0.$$

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

**Part 3. Shape of the second-best contract** The second part of the proof examines the shape of $t^{sb}(q)$. Consider the derivative $t^{sb\prime}(q)$ given in equation (A.15). 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 5), $u'(t^{sb}(q)) > 0$, $u''(t^{sb}(q)) < 0$ (Assumption 3), and $w'(p) > 0$ (Assumption 4), so the second term on the right-hand side of (A.15) is always positive. The first term on the right-hand side of (A.15) is identical to the right-hand side of (A.22), which determined the shape of $t^{fb}(q)$ and crucially depends on the shape of $w(p)$.

**Optimism** Proposition 1 shows that when the agent exhibits optimism (Definition 2), then $t_O^{sb\prime}(q) > 0$ for all $q$. Let us also study the behavior of the contract at the extremes. From Definition 2 and Lemma A1 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 2 and Lemma A3 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 $t_O^{sb}$ has high-powered incentives at extremes.

**Pessimism** Corollary 1 demonstrates that when the agent exhibits pessimism, the sign of $t_P^{sb\prime}(q)$ depends on the size of $w'(1-F(q|\bar{e}))$. When $q$ decreases, $w'(1-F(q|\bar{e}))$ increases; the second term of (A.15) becomes larger and $t_P^{sb\prime}(q)$ is more likely to be positive. The opposite happens when $q$ increases and $w'(1-F(q|\bar{e}))$ decreases; $t_P^{sb\prime}(q)$ is more likely to be negative and the (M) constraint is more likely to bind.

To further formalize that contract shape, I study (A.15) at the extremes. I start with $q \to \bar{q}$. From Definition 1 and Lemma A4, 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 (A.15) goes to $-\infty$ while the second goes to 0. Therefore, $\lim_{q\to\bar{q}^-} t_P^{sb\prime}(q) = -\infty$.

Let us now examine whether $t_P^{sb}$ ever increases with output; that is, whether $t_P^{sb\prime}(q) > 0$ for any segment in $[\underline{q}, \bar{q}]$, or equivalently, using (A.15), whether

$$\frac{w''\big(1-F(q|\bar{e})\big)f(q|\bar{e})}{w'\big(1-F(q|\bar{e})\big)} \frac{1}{\mu u'(t^{sb}(q)) w'\big(1-F(q|\bar{e})\big)} \leq -\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{w'\big(1-F(q|\underline{e})\big)f(q|\underline{e})}{w'\big(1-F(q|\bar{e})\big)f(q|\bar{e})}\right) \tag{A.25}$$

holds for some $q$. I use the fact that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{w'\big(1-F(q|\underline{e})\big)f(q|\underline{e})}{w'\big(1-F(q|\bar{e})\big)f(q|\bar{e})}\right) = \frac{w'\big(1-F(q|\underline{e})\big)}{w'\big(1-F(q|\bar{e})\big)}\Bigg[\Bigg(&\frac{w''\big(1-F(q|\bar{e})\big)}{w'\big(1-F(q|\bar{e})\big)}f(q|\bar{e}) \\ &- \frac{w''\big(1-F(q|\underline{e})\big)}{w'\big(1-F(q|\underline{e})\big)}f(q|\underline{e})\Bigg)\frac{f(q|\underline{e})}{f(q|\bar{e})} \\ &+ \frac{\mathrm{d}}{\mathrm{d}q}\frac{f(q|\underline{e})}{f(q|\bar{e})}\Bigg]\end{aligned} \tag{A.26}$$

to rewrite the inequality above as:

$$\begin{aligned}\frac{f(q|\underline{e})}{f(q|\bar{e})}\Bigg[\frac{w''\big(1-F(q|\bar{e})\big)}{w'\big(1-F(q|\bar{e})\big)}f(q|\bar{e})\left(\frac{1}{\mu u'(t^{sb}(q))w'\big(1-F(q|\underline{e})\big)\frac{f(q|\underline{e})}{f(q|\bar{e})}}+1\right) \\ - \frac{w''\big(1-F(q|\underline{e})\big)}{w'\big(1-F(q|\underline{e})\big)}f(q|\underline{e})\Bigg] \leq -\frac{\mathrm{d}}{\mathrm{d}q}\frac{f(q|\underline{e})}{f(q|\bar{e})}.\end{aligned} \tag{A.27}$$

From Assumption 4 we know that under pessimism $\lim_{q\to\underline{q}^+} w'\big(1-F(q|\underline{e})\big) = +\infty$. Further, the MLRP states that $\frac{f(q|\underline{e})}{f(q|\bar{e})}$ increases as $q$ decreases. Therefore, the quantity

$$\frac{1}{\mu u'(t^{sb}(q))w'\big(1-F(q|\underline{e})\big)\frac{f(q|\underline{e})}{f(q|\bar{e})}} \tag{A.28}$$

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

$$\left(\frac{w''\Big(1-F(q|\bar{e})\Big)}{w'\Big(1-F(q|\bar{e})\Big)}f(q|\bar{e})-\frac{w''\Big(1-F(q|\underline{e})\Big)}{w'\Big(1-F(q|\underline{e})\Big)}f(q|\underline{e})\right)\frac{f(q|\underline{e})}{f(q|\bar{e})}\leq-\frac{\mathrm{d}}{\mathrm{d}q}\frac{f(q|\underline{e})}{f(q|\bar{e})}, \tag{A.29}$$

which we know to hold from Lemma A7.

Therefore, there exists an output level $q_{\mathcal{I}} \in (\underline{q}, \bar{q})$ such that $t_P^{sb\prime}(q) \geq 0$ if $q < q_{\mathcal{I}}$ and $t_P^{sb\prime}(q) < 0$ otherwise. This result is in line with Corollary 1, which states that the M constraint is binding when $w''(p) > 0$ and $w'(p)$ is sufficiently small. In this specific case of pessimism, that constraint is binding for the segment $q > q_{\mathcal{I}}$. Due to complementary slackness, the optimal contract must exhibit $t'(q) = 0$ in that segment.

Denote by $\tilde{t}_P^{sb}$ the optimal solution when the agent is pessimistic. That solution can be characterized as

$$\tilde{t}_P^{sb}(q) = \begin{cases} t_P^{sb}(q) \text{ from (A.13)} & \text{if } q \in [\underline{q}, q_{\mathcal{I}}), \\ t_P^{sb}(q) \text{ from (A.13) evaluated at } q = q_{\mathcal{I}} & \text{if } q \in [q_{\mathcal{I}}, \bar{q}]. \end{cases} \tag{A.30}$$

■

**Corollary 2**

*Proof.* Let agent $i$ be more pessimistic than $j$. According to Lemma A5 iii), the probability $p_l \in (0,1)$ such that $w_i'(p_l) = w_j'(p_l)$ increases as $i$'s pessimism intensifies. Thus, the segment of probabilities $p \in (p_l, 1]$ such that the inequality $w_i'(p) > w_j'(p)$ holds becomes smaller.

Let $w_j'(p_l) = \varepsilon > 0$ for arbitrary small $\varepsilon > 0$. The reasoning given in the previous paragraph implies that the output level $q_\varepsilon \in [\underline{q}, \bar{q}]$ such that $w_i'\Big(1-F(q_\varepsilon|e)\Big) = \varepsilon$, is located at a lower output level as $i$ becomes more pessimistic. As a result, the performance segment $q \in [q_\varepsilon, \bar{q}]$, in which the weighting function exhibits $w_i'\Big(1-F(q|e)\Big) < \varepsilon$ becomes larger. Equation (A.28) shows that this tendency of $w_i'(p)$ to have a smaller value as $i$'s pessimism intensifies, enlarges the segment in which the solution to the relaxed problem, $t_P^{sb}$, exhibits $\frac{\mathrm{d}t_P^{sb}(q)}{\mathrm{d}q} < 0$. Therefore, the segment $q \in [q_{\mathcal{I}}, \bar{q}]$ over which the M constraint is binding becomes larger.

■

**Corollary 3**

*Proof.* Corollary 2 shows that the segment $q \in [q_{\mathcal{I}}, \bar{q}]$ where the contract $\tilde{t}_P^{sb}$ from Proposition 2 is constant becomes larger with stronger pessimism. The constant transfer given to the agent in that segment is equal to $t_P^{sb}(q_{\mathcal{I}})$ (the solution to the first-order condition in (A.13) evaluated at $q = q_{\mathcal{I}}$).

Notice that at $q = q_{\mathcal{I}}$, the contracts $\tilde{t}_P^{sb}$ from Proposition 2 and $t_P^{sb}$ from (A.13) are identical. However, because $\frac{\mathrm{d}t_P^{sb}(q)}{\mathrm{d}q} < 0$, the solution to the relaxed problem $t_P^{sb}$ becomes smaller for larger values of $q$. This means that this contract is less costly to the principal relative to the contract from Proposition 2. Hence, the cost incurred by the principal from complying with the M constraint is the difference:

$$\int_{\underline{q}}^{\bar{q}} t_P^{sb}(q) - \tilde{t}_P^{sb}(q)\mathrm{d}q = \int_{q_{\mathcal{I}}}^{\bar{q}} t_P^{sb}(q) - \tilde{t}_P^{sb}(q)\mathrm{d}q, \tag{A.31}$$

where in the equality in the equation above follows from the fact that $\tilde{t}_P^{sb}(q) = t_P^{sb}(q)$ for all $[\underline{q}, q_{\mathcal{I}}]$. The right-hand side of (A.31) shows that higher pessimism, and thus a lower value of $q_{\mathcal{I}}$, increases the average cost faced by the principal.

**Proposition 3**

*Proof.* The objective function of the principal when offering $t_P^{sb}$, the solution to the unconstrained problem, given in (A.13), is:

$$\int_{\underline{q}}^{\bar{q}} S(q) - t_P^{sb}(q)f(q|e)\mathrm{d}q = S(\bar{q}) - t_P^{sb}(\bar{q}) - \int_{\underline{q}}^{\bar{q}} S'(q) - t_P^{sb\prime}(q)F(q|e)\mathrm{d}q, \tag{A.32}$$

where the equality in the equation above follows from integration by parts. In the model, it is assumed that the principal is interested in eliciting high effort, $\bar{e}$. That assumption can be written as:

$$-\int_{\underline{q}}^{\bar{q}} S'(q) - t_P^{sb\prime}(q)\Big(F(q|\bar{e}) - F(q|\underline{e})\Big)\mathrm{d}q \geq 0. \tag{A.33}$$

The MLRP implies that $F(q|\bar{e}) \leq F(q|\underline{e})$, which validates the plausibility of the inequality given in the equation above. Equation (A.33) can be rewritten as:

$$-\int_{\underline{q}}^{\bar{q}} S'(q)\Big(F(q|\bar{e}) - F(q|\underline{e})\Big)\mathrm{d}q + \int_{\underline{q}}^{q_{\mathcal{I}}} t_P^{sb\prime}(q)\Big(F(q|\bar{e}) - F(q|\underline{e})\Big)\mathrm{d}q$$
$$+ \int_{q_{\mathcal{I}}}^{\bar{q}} t_P^{sb\prime}(q)\Big(F(q|\bar{e}) - F(q|\underline{e})\Big)\mathrm{d}q \geq 0. \tag{A.34}$$

The first expression in the left-hand side of (A.34) is positive while the second expression there is negative. Moreover, since, $\frac{\mathrm{d}t_P^{sb}(q)}{\mathrm{d}q} < 0$ in $q \in [q_{\mathcal{I}}, \bar{q}]$, the third expression in the left-hand side of (A.34) is positive and makes the inequality in that equation less stringent.

Recall that the contract from Proposition 2 exhibits $\tilde{t}_P^{sb\prime}(q) = 0$ in $q \in [q_{\mathcal{I}}, \bar{q}]$. Thus, when the principal offers that contract, her objective function given in (A.34) becomes:

$$-\int_{\underline{q}}^{\bar{q}} S'(q)\Big(F(q|\bar{e}) - F(q|\underline{e})\Big)\mathrm{d}q + \int_{\underline{q}}^{q_{\mathcal{I}}} \tilde{t}_P^{sb\prime}(q)\Big(F(q|\bar{e}) - F(q|\underline{e})\Big)\mathrm{d}q \geq 0. \tag{A.35}$$

The inequality in (A.35) is more stringent than that in (A.34). Hence, the principal experiences higher costs from implementing $\tilde{t}_P^{sb}$ instead of $t_P^{sb}$.

Next, let $\Omega := \Big\{w_1(p), ..., w_n(p)\Big\}$ be the set of all probability weighting functions with pessimism. Assume that $\Omega$ is a partially ordered set according to increasing degrees of pessimism. Thus, for any two $w_j(p), w_k(p) \in \Omega$ such that $1 \leq j < k \leq n$, it must be that $w_j(p) \geq w_k(p)$, which, in line with Lemma A5, means that $w_j$ exhibits less pessimism than $w_k(p)$. Notably, the cardinality of that set is $n(\Omega) = \infty$, which reflects all the positive values that that the second derivative, $w''$, can attain.

Corollary 2 shows that the segment $q \in [q_{\mathcal{I}}, \bar{q}]$ where the contract $\tilde{t}_P^{sb}$ is constant becomes larger with stronger pessimism. Accordingly, for $w_1(p) \in \Omega$ the performance level $q_{\mathcal{I}}$ is located in the neighborhood of $\bar{q}$. In this case $\tilde{t}_P^{sb}$ an $t_P^{sb}$ are indistinguishable as the segment in which $\tilde{t}_P^{sb}$ is constant is infinitesimal. Thus, in such a case $\tilde{t}_P^{sb}$ is profitable by definition.

In contrast, for $w_n(p) \in \Omega$, the segment in which $\tilde{t}_P^{sb}$ is constant is the largest. In this case, $q_{\mathcal{I}}$ is in the neighborhood of $\underline{q}$. From Definition 1 and Lemma A2 we know that $\lim_{p \to 1^-} \frac{w''(p)}{w'(p)} = +\infty$, so, by equation (A.15), then $\lim_{q \to \bar{q}^-} \tilde{t}_P^{sb\prime}(q) = +\infty$. Hence, the inequality in (A.35) cannot hold when $q_{\mathcal{I}} \to \underline{q}$; it is not profitable for the principal to implement $\tilde{t}_P^{sb}$ when the agent's pessimism is extreme.

The property that $n(\Omega) = \infty$, together with the aforementioned conclusions that $\tilde{t}_P^{sb}$ is profitable for the smallest element of $\Omega$ but not for the largest one, and that the segment in which $\tilde{t}_P^{sb}$ is constant increases with pessimism (Corollary 2), imply the existence of a unique element $w_l(p) \in \Omega$ that makes Eq. (A.35) hold with equality. For all elements $w_i(p) \in \Omega$ such that $w_i(p) \leq w_l(p)$, Eq. (A.35) holds for some $q_{\mathcal{I}} \geq \underline{q}$. Instead, for all elements $w_j(p) \in \Omega$ such that $w_j(p) \geq w_l(p)$, Eq. (A.35) cannot hold. For the partition $[w_{l+1}(p), ..., w_n(p)]$, the principal has no other option than to give up incentive compatibility.

■

**Corollary 4**

*Proof.* The proof consists of two parts. The first one shows that stronger optimism enlarges the segment $[\underline{q}, q^*]$, where $q^* \in (\underline{q}, \bar{q})$ is a performance level that satisfies $w'\Big(1 - F(q^*|\bar{e})\Big) = 1$. The second part shows that if $q \in [\underline{q}, q^*]$, stronger optimism is more likely to reduce the set $(0, \hat{c}_O]$. These two statements amount to stronger optimism decreasing the likelihood that $t_O^{sb}(q) = t_O^{fb}(q)$.

**Part 1.** Let agent $i$ be more optimistic than agent $j$. Accordingly, $w_i(p) = \theta\Big(w_j(p)\Big)$ for all $p \in [0, 1]$ (Definition 3). Let $p_l, p_k \in (0, 1)$ be the probabilities guaranteeing $w_i'(p_l) = 1$ and $w_j'(p_k) = 1$. Notice that $\lim_{p \to 0^+} w_i'(p) = \lim_{p \to 0^+} w_j'(p) = +\infty$. According to Lemma A5 i), $w_i'$ tends to zero faster than $w_j'$ as $p$ increases. Thus, it must be that $p_k < p_l$. This implies that the output level $q^*$ such that $w_j'\Big(1-F(q^*|e)\Big)=1$ takes place at a higher output level when $j$ becomes as optimistic as $i$.

**Part 2.** Proposition 4 demonstrates the existence of a critical cost level $\hat{c}_{O,j}$, such that

$$\hat{c}_{O,j}=\int_{\underline{q}}^{\bar{q}} u'\Big(t_{O,j}^{fb}(q)\Big){t_{O,j}^{fb}}'(q)\bigg(w_j\Big(1-F(q|\bar{e})\Big)-w_j\Big(1-F(q|\underline{e})\Big)\bigg)\,\mathrm{d}q. \tag{A.36}$$

If $i$ were given the same contract as that given to $j$, his critical cost level must satisfy:

$$\hat{c}_{O,i}=\int_{\underline{q}}^{\bar{q}} u'\Big(t_{O,j}^{fb}(q)\Big){t_{O,j}^{fb}}'(q)\bigg(w_i\Big(1-F(q|\bar{e})\Big)-w_i\Big(1-F(q|\underline{e})\Big)\bigg)\,\mathrm{d}q. \tag{A.37}$$

The existence of $\hat{c}_{O,i}$ is guaranteed by $w_i\Big(1-F(q|\bar{e})\Big)\geq w_i\Big(1-F(q|\underline{e})\Big)$ (Lemma 1) and ${t_{O,j}^{fb}}'(q)>0$ (Equation (A.22)), both of which imply that the right-hand side of the inequality in (A.37) is non-negative. As a result, there exists a set $c\in(0,\hat{c}_{O,i})$ in which $t_{O,j}^{fb}$ incentivizes agent $i$.

Throughout, define $w_0,w_1\in[0,1]$ to be the probability weights $w_0=w_j\Big(1-F(q|\underline{e})\Big)$ and $w_1=w_j\Big(1-F(q|\bar{e})\Big)$ for a given output level $q$. Recall that $w_1\geq w_0$ due to Lemma 1. Moreover, notice that in the set $q\in[\underline{q},q^*]$, optimism implies $w_i'\Big(1-F(q|e)\Big)<1$. Recognizing that $w_i'\Big(1-F(q|e)\Big)=\theta'\bigg(w_j\Big(1-F(q|e)\Big)\bigg)w_j'\Big(1-F(q|e)\Big)$, the inequality $w_i'\Big(1-F(q|e)\Big)<1$ is integrated over $[w_0,w_1]$ to obtain:

$$\begin{aligned}\int_{w_0}^{w_1}\theta'(s)\mathrm{d}s<\int_{w_0}^{w_1}\mathrm{d}s\Leftrightarrow\\ w_i\Big(1-F(q|\bar{e})\Big)-w_i\Big(1-F(q|\underline{e})\Big)<&w_j\Big(1-F(q|\bar{e})\Big)-w_j\Big(1-F(q|\underline{e})\Big).\end{aligned} \tag{A.38}$$

Equation (A.38) together with equations (A.36) and (A.37) imply that $\hat{c}_{O,i}<\hat{c}_{O,j}$ for the set $q\in[\underline{q},q^*]$. The first part of this proof demonstrated that the set $q\in[\underline{q},q^*]$ is enlarged with higher optimism. Hence, stronger optimism increases the likelihood that the set $c<\hat{c}_O$ becomes smaller.

■

**Corollary 5**

*Proof.* We start by rewriting the first-order condition in (A.13) 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 of the previous equation 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

$$\nu = \int_{\underline{q}}^{\bar{q}} \frac{1}{u'(t^{sb}(q))} f(q|\bar{e})\, \mathrm{d}q = \mathbb{E}_{\bar{e}}\left(\frac{1}{u'(t^{sb}(q))}\right), \tag{A.39}$$

where $\mathbb{E}_{\bar{e}}$ is the expectation with respect to the probability distribution of $q$ induced by $\bar{e}$. Equation (A.39) shows that $\nu > 0$ is the same for agents with different types of probability weighting functions $w$.

After plugging (A.39) into (A.13) and multiplying by $u(t^{sb}(q))$, we obtain:

$$\begin{aligned} &\mu u(t^{sb}(q))\Big[w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e}) - w'\Big(1 - F(q|\underline{e})\Big) f(q|\underline{e})\Big] \\ &= f(q|\bar{e}) u(t^{sb}(q)) \left[\frac{1}{u'(t^{sb}(q))} - \mathbb{E}_{\bar{e}}\left(\frac{1}{u'(t^{sb}(q))}\right) w'\Big(1 - F(q|\bar{e})\Big)\right]. \end{aligned} \tag{A.40}$$

From the complementary slackness condition associated with $\mu$ we know that:

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

Therefore, we can rewrite (A.40), after integrating with respect to $q$ over $[\underline{q}, \bar{q}]$, as

$$\begin{aligned} \mu c &= \int_{\underline{q}}^{\bar{q}} u(t^{sb}(q)) \left[\frac{1}{u'(t^{sb}(q))} - \mathbb{E}_{\bar{e}}\left(\frac{1}{u'(t^{sb}(q))}\right) w'\Big(1 - F(q|\bar{e})\Big)\right] f(q|\bar{e})\, \mathrm{d}q \\ &= \mathbb{E}_{\bar{e}}\left(\frac{u(t^{sb}(q))}{u'(t^{sb}(q))}\right) - \mathbb{E}_{\bar{e}}\left(\frac{1}{u'(t^{sb}(q))}\right) \int_{\underline{q}}^{\bar{q}} u(t^{sb}(q)) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})\, \mathrm{d}q \\ &= \mathbb{E}_{\bar{e}}\left(\frac{u(t^{sb}(q))}{u'(t^{sb}(q))}\right) - \mathbb{E}_{\bar{e}}\left(\frac{1}{u'(t^{sb}(q))}\right) \tilde{\mathbb{E}}_{\bar{e}}\Big(u(t^{sb}(q))\Big). \end{aligned} \tag{A.41}$$

where $\tilde{\mathbb{E}}_{\bar{e}}$ is the expectation as perceived by an agent who suffers from probability weighting. Since $\tilde{\mathbb{E}}_{\bar{e}}\Big(u(t^{sb}(q))\Big) < \mathbb{E}_{\bar{e}}\Big(u(t^{sb}(q))\Big)$ under pessimism, and the opposite under optimism, equation (A.41) implies that $\mu_P > \mu_{EU} > \mu_O \geq 0$, where $\mu_P$ is the lagrangian multiplier of the pessimist and $\mu_O$ that of the optimist.

We again rewrite (A.13) but this time as:

$$\frac{1}{u'(t^{sb}_{nEU}(q))} = \nu w'\left(1 - F(q|\bar{e})\right) + \mu_{nEU} w'\left(1 - F(q|\bar{e})\right)\left(1 - \frac{w'\left(1 - F(q|\underline{e})\right) f(q|\underline{e})}{w'\left(1 - F(q|\bar{e})\right) f(q|\bar{e})}\right), \quad \text{(A.42)}$$

where $t^{sb}_{nEU} \in \{t^{sb}_O, t^{sb}_P\}$ and $\mu_{nEU} \in \{\mu_O, \mu_P\}$. For the EUT agent, equation (A.42) simplifies to:

$$\frac{1}{u'(t^{sb}_{EU}(q))} = \nu + \mu_{EU}\left(1 - \frac{f(q|\underline{e})}{f(q|\bar{e})}\right). \quad \text{(A.43)}$$

**Part 1.** Equations (A.42) and (A.43) imply that $t^{sb}_{nEU}(q) < t^{sb}_{EU}(q)$ holds if i) $w'\big(1 - F(q|\bar{e})\big) < 1$, ii) $w'\big(1 - F(q|\bar{e})\big) < w'\big(1 - F(q|\underline{e})\big)$, and iii) $\mu_{EU} > \mu_{nEU}$. *For the optimist*, these three conditions hold in $q \in [\underline{q}, q^*)$, where $q^*$ is the output level such that $w'\big(1 - F(q^*|\bar{e})\big) = 1$. In that performance segment, $t^{sb}_O(q) < t^{sb}_{EU}(q)$.

In contrast, in the set $q \in [q^*, \bar{q}]$ the weighting function of the optimist exhibits $w'(1-F(q|\bar{e})) \geq 1$ which contradicts the condition in i). Moreover, since $\lim_{q\to\bar{q}} \frac{f(q|\underline{e})}{f(q|\bar{e})} = 0$ and $\lim_{q\to\bar{q}} w'(1 - F(q|\bar{e})) = +\infty$, it must be that at the highest performance level: $t^{sb}_O(\bar{q}) > t^{sb}_{EU}(\bar{q})$. Furthermore, for output levels lower than $\bar{q}$, the inequality $t^{sb}_O(q) > t^{sb}_{EU}(q)$ can also hold under some conditions. Let $\mu_O \to \mu_{EU}$, which happens in case of moderate optimism. In that case, $t^{sb}_O(q) > t^{sb}_{EU}(q)$ when $\frac{f(q|\underline{e})}{f(q|\bar{e})}$ is sufficiently small, or equivalently, at high enough output levels.

**Part 2.** Comparing equations (A.42) and (A.43) yields that $t^{sb}_{nEU}(q) > t^{sb}_{EU}(q)$ if i) $w'\big(1 - F(q|\bar{e})\big) > 1$, ii) $w'\big(1 - F(q|\bar{e})\big) > w'\big(1 - F(q|\underline{e})\big)$, and iii) $\mu_{nEU} \geq \mu_{EU}$. *For the pessimist*, these conditions hold in $q \in [\underline{q}, q^*)$. In that segment of performance, $t^{sb}_P(\bar{q}) > t^{sb}_{EU}(\bar{q})$.

Instead, in $q \in [q^*, \bar{q}]$, the probability weighting function exhibits $w'\left(1 - F(q|\bar{e})\right) \leq 1$ which contradicts the condition in i). Moreover, since $\lim_{q\to\bar{q}} \frac{f(q|\underline{e})}{f(q|\bar{e})} = 0$ and $\lim_{q\to\bar{q}} w'\big(1 - F(q|\bar{e})\big) = 0$, it must be that at the lowest performance level: $t^{sb}_P(\bar{q}) < t^{sb}_{EU}(\bar{q})$. For output levels lower than $\bar{q}$, the inequality $t^{sb}_P(q) < t^{sb}_{EU}(q)$ can also hold under some conditions. Specifically, it holds for sufficiently large $\frac{f(q|\underline{e})}{f(q|\bar{e})}$, that is, for low enough output levels. ■

**Proposition 5**

*Proof.* The solution to the principal's problem is similar to that presented in Proposition 4 with the difference that $w(p)$ is now inverse-S shaped.

**Part 1. Optimal Contract without the M and IC constraints** The contract satisfying the first-order condition in (A.21) remains to be the solution to the problem in (4) without IC and M constraints. Denote that solution by $t^{fb}_L(q)$.

Equation (A.22) shows that in $q \in [\tilde{q}, \bar{q}]$, the segment in which the weighting function exhibits $w''(p) < 0$, the solution to the problem without the IC and P constraints is increasing in performance, $t_L^{fb'}(q) > 0$. Also, Lemma A2 implies that $\lim_{q \to \bar{q}^-} \frac{w''(1-F(q|\bar{e}))}{w'(1-F(q|\bar{e}))} = +\infty$; at the highest output level, that solution exhibits $\lim_{q \to \bar{q}^-} t_L^{fb'}(q) = +\infty$. Moreover, at the inflection point, $\tilde{q}$, the weighting function exhibits $\lim_{q \to \tilde{q}} w''\big(1 - F(q|\bar{e})\big) = 0$. Thus, it must be that $\lim_{q \to \tilde{q}^-} t_L^{fb'}(q) = 0$.

Consider now the set $q \in [\underline{q}, \tilde{q}]$ where the weighting function exhibits $w''(q) > 0$. According to (A.22), the solution to the unconstrained problem exhibits $t_L^{fb'}(q) < 0$. This is consistent with Corollary 1, which shows that without the IC constraint and when $w(p)$ is everywhere convex, the M constraint must be binding. Hence, because of complementary slackness, the solution to the unconstrained problem without the IC and M constraints must exhibit $t'(q) = 0$ in $q \in [\underline{q}, \tilde{q}]$. Thus, the resulting solution $\tilde{t}_L^{fb}(q)$ is:

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

**Part 2. Incentive constraint can be slack at the optimum** Denote by $t_L^{sb}(q)$ the solution to the unconstrained problem when the M constraint is not taken into account. Suppose that $\mu = 0$.Then $t_L^{sb}(q) = t_L^{fb}(q)$, where $t_L^{fb}(q)$ is the contract presented above. From the complementary slackness condition of $\mu = 0$ we get:

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_L^{fb}(q)\Big) w'(1 - F(q|\bar{e})) f(q|\bar{e}) \,\mathrm{d}q - c > \int_{\underline{q}}^{\bar{q}} u\Big(t_L^{fb}(q)\Big) w'(1 - F(q|\underline{e})) f(q|\underline{e}) \,\mathrm{d}q. \quad \text{(A.44)}$$

Using integration by parts, rewrite (A.44) as:

$$\int_{\underline{q}}^{\bar{q}} u'\Big(t_L^{fb}(q)\Big) \frac{\mathrm{d}t_L^{fb}(q)}{\mathrm{d}q} \Big[w(1 - F(q|\bar{e})) - w(1 - F(q|\underline{e}))\Big] \,\mathrm{d}q > c. \quad \text{(A.45)}$$

Assumption 5 implies $w\,(1 - F(q|\bar{e})) - w\,(1 - F(q|\underline{e})) \geq 0$ which, together with $\frac{\mathrm{d}t_O^{fb}(q)}{\mathrm{d}q} > 0$ in $q > \tilde{q}$ (equation (A.22)) and $u'(t) > 0$ (Assumption 3), imply that the left-hand side of (A.24) is weakly positive.

Because $w$ and $u$ are twice continuously differentiable (Assumptions 3 and 4), and since $c$ is a constant unbounded from above, there exists $\hat{c}_L > 0$ such that if $c < \hat{c}_L$, then the inequality in (A.45) holds. In that case, $\mu = 0$ and $t_L^{sb}(q) = t_L^{fb}(q)$. On the other hand, if $c \geq \hat{c}_L$, then the inequality in (A.45) does not hold. In that case, the IC binds, $\mu > 0$, and $t_L^{sb}(q)$ satisfies (A.13). In the remainder of the proof I consider the latter case.

**Part 3. Shape of $t^{sb}(q)$ when the IC constraint is binding.** The second part of the proof analyzes the shape of $t^{sb}(q)$. To that end, I use equation (A.15) which presents the derivative of $t^{sb}_L(q)$ with respect to $q$. Throughout, let $\tilde{q} \in (\underline{q}, \bar{q})$ be the output level satisfying $w\Big(1-F(\tilde{q}|\bar{e})\Big) = 1-F(\tilde{q}|\bar{e}) = 0.5$, i.e. the output level at which the weighting function transitions from concave to convex.

According to Definition 4, likelihood insensitivity implies $w'' < 0$ for all $q \in (\tilde{q}, \bar{q}]$. Hence, the two terms on the right-hand side of (A.15) are positive and $\frac{\mathrm{d}t^{sb}_L(q)}{\mathrm{d}q} > 0$ in $q \in (\tilde{q}, \bar{q}]$; the solution to the relaxed problem is increasing in $q \in (\tilde{q}, \bar{q}]$. Moreover, since $w''\Big(1-F(\tilde{q}|\bar{e})\Big) = 0$, that solution also displays $\lim_{q\to\tilde{q}^+} \frac{\mathrm{d}t^{sb}_L(q)}{\mathrm{d}q} > 0$.

Likelihood insensitivity also implies that $w'' > 0$ for all $q \in [\underline{q}, \tilde{q}]$ (Definition 4). Therefore, the first term on the right-hand side of (A.15) is negative while the second term there is positive. According to Corollary 1, the magnitude of $w'\Big(1-F(q|\bar{e})\Big)$ determines which of these terms dominates the other, which, in turn, determines the sign of $\frac{\mathrm{d}t^{sb}_L(q)}{\mathrm{d}q}$. When $q$ decreases in the segment $q \in [\underline{q}, \tilde{q}]$, then the expression $w'(1-F(q|\bar{e}))$ increases, making the second term in (A.15) larger, and thus that $\frac{\mathrm{d}t^{sb}_P(q)}{\mathrm{d}q}$ is more likely to be positive. Therefore, the solution to the relaxed problem is increasing in performance as $q \to \underline{q}$.

Let us now investigate the behavior of $\frac{\mathrm{d}t^{sb}_P(q)}{\mathrm{d}q}$ when $q \to \tilde{q}^-$. Notice that at that output level $w''\Big(1-F(\tilde{q}|\bar{e})\Big) = 0$. Hence according to (A.15), $\lim_{q\to\tilde{q}^-} \frac{\mathrm{d}t^{sb}_L(q)}{\mathrm{d}q} > 0$. It remains to be shown whether $t^{sb}$ ever decreases with output; that is whether $\frac{\mathrm{d}t^{sb}_L(q)}{\mathrm{d}q} < 0$ in $(\underline{q}, \tilde{q})$.

Lemma A6 shows that for agent $i$, who is more likelihood insensitive than $j$, the probability $p_l \in (0, 0.5)$ inducing $w'_i(p_l) = w'_j(p_l)$ is smaller as $i$'s insensitivity becomes more severe with respect to that of $j$. That lemma also shows that the probability $p_n \in (0.5, 1)$ such that $w'_i(p_n) = w'_j(p_n)$ becomes larger as $i$ becomes more insensitive with respect to $j$. Consequently, as $i$ becomes more insensitive (with respect to $j$), the output levels $q_1, q_2 \in [\underline{q}, \tilde{q}]$ such that $q_2 > q_1$ that induce $w'_i\Big(1-F(q_1|e)\Big) = w'_i\Big(1-F(q_2|e)\Big) = \varepsilon$ for arbitrary small $\varepsilon > 0$ take place at a lower output level, in the case of $q_1$, and a higher output level, in the case of $q_2$. Thus, stronger likelihood insensitivity expands the segment $q \in [q_1, q_2]$ in which $w'_i\Big(1-F(q|e)\Big) < \varepsilon$. According to Corollary 1, this property facilitates that the solution to the relaxed problem exhibits $\frac{\mathrm{d}t^{sb}_L(q)}{\mathrm{d}q} < 0$ for intermediate values in $q \in (\underline{q}, \tilde{q}]$.

Define the segment of output levels in which $\frac{\mathrm{d}t^{sb}_L(q)}{\mathrm{d}q} < 0$ by $q \in [q_{\mathcal{I}1}, q_{\mathcal{I}2}]$ where $q_{\mathcal{I}1}, q_{\mathcal{I}2} \in (\underline{q}, \tilde{q})$ are output levels such that $q_{\mathcal{I}1} < q_{\mathcal{I}2}$. According to Corollary 1, the M constraint is binding in $q \in [q_{\mathcal{I}1}, q_{\mathcal{I}2}]$ if $i$ becomes suffers from sufficiently strong insensitivity. Thus, due to complementary slackness, the optimal contract must exhibit

$t'(q) = 0$ in $q \in [q_{\mathcal{I}1}, q_{\mathcal{I}2}]$. The optimal contract is therefore:

$$\tilde{t}_L^{sb}(q) = \begin{cases} t_L^{sb}(q) & \text{if } q \in [\underline{q}, q_{\mathcal{I}1}) \cup (q_{\mathcal{I}2}, \bar{q}], \\ t_L^{sb}(q_{\mathcal{I}1}) & \text{if } q \in [q_{\mathcal{I}1}, q_{\mathcal{I}2}]. \end{cases} \tag{A.46}$$

Instead, if likelihood insensitivity is not strong so as to generate $\frac{\mathrm{d}t_L^{sb}(q)}{\mathrm{d}q} < 0$ in $q \in [q_{\mathcal{I}1}, q_{\mathcal{I}2}]$, the optimal contract is equal to $t_L^{sb}(q)$ and is everywhere increasing.

■

**Corollary 6**

*Proof.* Let agent $i$ be more likelihood insensitive than $j$. Proposition 5 shows that the segment $q \in [q_1, q_2]$ in which $w_i'\big(1 - F(q|e)\big) < \varepsilon$ holds is larger for $i$ than for $j$. Corollary 1 shows that this property makes the expression in (A.28) larger, which in turn facilitates $\frac{\mathrm{d}t_L^{sb}(q)}{\mathrm{d}q} < 0$ for intermediate values of the set $q \in (\underline{q}, \tilde{q}]$. Thus, the performance segment in which the (M) constraint binds and the optimal solution exhibits $\tilde{t}^{sb\prime}_L(q) = 0$ becomes larger with stronger likelihood insensitivity.

■

# B Prospect Theory Preferences [FOR ONLINE PUBLICATION ONLY]

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

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

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

**Assumption B1.** *The value function, $V(t, R)$, is a piece-wise function,*

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

*with the following properties:*

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

The agent's utility is convex for losses, generating risk seeking attitudes, and concave for gains, generating risk aversion. Furthermore, Assumption B1 introduces loss aversion. That is, transfers counting as losses loom larger than equally-sized transfers counting as gains. This latter property is captured by the parameter $\lambda > 1$ and expresses a special dislike for losses.

The second ingredient is that the probability weighting function is defined separately over gains and losses. Probabilities associated with gains are transformed by the probability weighting function $w$, introduced in Assumption 4. 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.[20] I simplify the problem by assuming that $z$ adopts the properties of $w$.

**Assumption B2.** *A probability weighting function for losses is a function* $z : [0,1] \to [0,1]$ *satisfying the duality condition* $z(F(q|e)) = 1 - w(1 - F(q|e))$ *for any* $e$.

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

$$CPT(t,e,r) = \int_{\underline{q}}^{\bar{q}} \Big[\zeta v\left(t(q) - R\right) w'\Big(1 - F(q|e)\Big) \\ - \lambda(1-\zeta)v(R - t(q))z'\Big(F(q|e)\Big)\Big] f(q|e)\,\mathrm{d}q - c(e), \qquad \text{(B.1)}$$

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

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

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

$$\text{s.t.} \qquad CPT(t,\bar{e},R) \geq \bar{V},$$

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

$$t(q_2) \geq t(q_1) \text{ for all } q_2, q_1 \in [\underline{q},\bar{q}] \text{ such that } q_2 > q_1.$$

Where the term $\bar{V}$ is the outside option of the agent, which can be negative.

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

**Proposition B1.** *Let Assumptions 1 - 5 and Assumptions B1-B2 hold. There exists a threshold* $\hat{q} \in [\underline{q},\bar{q}]$ *such that the second best-contract,* $t_C^{sb}$*:*

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

*(ii) pays* $R$ *in* $q < \hat{q}$ *and depends on performance as in Proposition 2, 3, 4, or 5 in* $q \geq \hat{q}$ *if* $\hat{q} \in (\underline{q},\bar{q})$*;*

*(iii) depends on performance as in Proposition 2, 3, 4, or 5 if* $\hat{q} = \underline{q}$*.*

*Proof.* Rewrite Eq. (B.1) using Assumption B2 as

$$CPT(t,e,R) = \int_{\underline{q}}^{\bar{q}} \Big[\zeta v\left(t(q) - R\right) w'\Big(1 - F(q|e)\Big) \\ - \lambda(1-\zeta)v(R - t(q))w'\Big(1 - F(q|e)\Big)\Big] f(q|e)\,\mathrm{d}q - c(e), \qquad \text{(B.2)}$$

[20] 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.

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

$$\begin{aligned}\mathcal{L}(q,t) =& \Big(S(q) - t(q)\Big) f(q|\bar{e}) \\ &+ \mu\Bigg( - \lambda v\Big(R - t(q)\Big)\Big(w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e}) - w'\Big(1 - F(q|\underline{e})\Big) f(q|\underline{e})\Big) - c\Bigg) \\ &+ \nu\Bigg( - \lambda v\Big(R - t(q)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e}) - c - \bar{V}\Bigg).\end{aligned} \tag{B.3}$$

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

$$\frac{1}{\lambda v'\Big(R - t\Big)\Big(w'\Big(1 - F(q|\bar{e})\Big)\Big)} = \nu + \mu\left(1 - \frac{w'\Big(1 - F(q|\underline{e})\Big) f(q|\underline{e})}{w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})}\right). \tag{B.4}$$

Denote by $t_C^{sb}(q)$ the transfer satisfying Eq. (B.4). I 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 of performance $q$ such that:

$$\lambda v\Big(R - t_C^{sb}(q)\Big) = \lambda\Big(1 - w(\rho)\Big) v(R). \tag{B.5}$$

Hence $L_\rho := (\rho, R; 1 - \rho, 0)$ leaves the agent's participation and incentive compatibility constraints unchanged. Using the fact that the utility exhibits $v'' < 0$, the above equation implies:

$$\lambda v\Big(R - t_C^{sb}(q)\Big) \leq \lambda v\Big((1 - w(\rho))R\Big). \tag{B.6}$$

Since $v' > 0$, then the previous equation implies $t_C^{sb}(q) > w(\rho)R$. Therefore, the lottery contract $L_\rho$ 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 $L_\rho$ are studied next. Let $\bar{L} = \rho R$. The utility of an agent when given $L_\rho$ is

$$CPT(L_\rho, \bar{e}, R) = -\left(1 - w\left(\frac{\bar{L}}{R}\right)\right) \lambda v(r) - c \tag{B.7}$$

The equation above is not linear in $\bar{L}$ due to $w$ having curvature (Assumption 4). Hence, changes in $\bar{L}$ affect the agent's marginal utility. To further understand how changes in $\bar{L}$ affect the marginal incentives of offering the lottery contract, I compute the first-order condition of (B.7) with respect to $\rho$, which gives:

$$w'(\rho)\lambda v(R) = 0. \tag{B.8}$$

Denote by $\rho^{opt}$ the probability satisfying the condition in (B.8). The second-order condition evaluated at $\rho^{opt}$ is:

$$w''(\rho^{opt})\lambda v(R). \tag{B.9}$$

The second-order condition shows that $\rho^{opt} \in (0,1)$ whenever $w'' < 0$. This holds under optimism or likelihood insensitivity. Moreover, by Assumption 4, $\lim_{\rho \to 1} w'(\rho) = 0$ under optimism. So, in that case the first-order condition in (B.9) holds if $\rho^{opt} \to 1$. Instead, if $w'' > 0$ for any interval in $p \in (0,1)$, then the solution must be $\rho^{opt} \in \{0,1\}$. Comparing the utilities at those extreme levels of probability gives:

$$CPT(L_{\rho=1}, \bar{e}, R) = -c > -\lambda v(R) - c = CPT(L_{\rho=0}, \bar{e}, R). \tag{B.10}$$

Thus, it must be that the solution is $\rho^{opt} = 1$. Regardless of the shape of the probability weighting function, the principal avoids exposing the agent to losses by setting $\rho^{opt} = 1$ and paying $t = R$.

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

$$\begin{aligned}\mathcal{L}(q,t) =& \Big(S(q) - t(q)\Big) f(q|\bar{e}) \\ &+ \mu\bigg(v\Big(t(q) - R\Big)\Big(w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e}) - w'\Big(1 - F(q|\underline{e})\Big) f(q|\underline{e})\Big) - c\bigg) \\ &+ \nu\bigg(v\Big(t(q) - R\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e}) - c - \bar{V}\bigg).\end{aligned} \tag{B.11}$$

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

$$\frac{1}{v'\Big(t - R\Big) w'\Big(1 - F(q|\bar{e})\Big)} = \nu + \mu\left(1 - \frac{w'\Big(1 - F(q|\underline{e})\Big) f(q|\underline{e})}{w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})}\right). \tag{B.12}$$

Since $v' > 0$ and $v'' < 0$ and $w(p)$ is as described by Assumption 4, the solution is similar to that presented in Proposition 2, 3, 4, or 5, except that the initial utility value is now $R > 0$.

To establish the location of the switch from paying the amount $t = R$, given to protect the agent from losses, to paying the agent according to the contracts from Proposition 2, 3, 4, or 5, denote by $\hat{q} \in [\underline{q}, \bar{q}]$ the performance level satisfying:

$$\frac{1}{\frac{\lambda v(R)}{R}} = \nu + \mu\left(1 - \frac{w'\left(1 - F(\hat{q}|\underline{e})\right)f(\hat{q}|\underline{e})}{w'\left(1 - F(\hat{q}|\bar{e})\right)f(\hat{q}|\bar{e})}\right). \tag{B.13}$$

Where the left-hand side of (B.13) captures the marginal incentives of offering the lottery contract $L$ with $\rho = 1$. The existence and uniqueness of $\hat{q}$ is guaranteed by the fact that the left-hand side of Eq. (B.13) of is positive and constant in $q$ while the right-hand side of that equation increases with $q$ (Assumption 5) over $[0, +\infty)$.

There are three cases. When $\frac{\lambda v(R)}{R}$ is small and the right-hand side of (B.13) is large enough, then $\hat{q} \geq \bar{q}$. In that case, the solution is that $t_C^{sb} = R$ is offered everywhere. Alternatively, $\frac{\lambda v(R)}{R}$ can be large so that $\hat{q} \leq \bar{q}$ and the solution is fully characterized by Proposition 4 and Proposition 5, depending on the shape of $w$. Finally, if $\hat{q} \in [\underline{q}, \bar{q}]$ then

$$t_C^{sb}(q) = \begin{cases} R & \text{if } q < \hat{q}, \\ t_P^{sb}(q), t_O^{sb}(q) \text{ (Proposition 4), or } t_L^{sb}(q) \text{ (Proposition 5)} & \text{if } q \geq \hat{q}. \end{cases} \tag{B.14}$$

■

Under CPT preferences, the optimal contract often includes a performance-insensitive segment paying the amount $r$. The reason behind this segments is loss aversion. Exposing the agent to losses by paying amounts lower than $R$ would generate large disutility, eventually leading to the agent rejecting the contract. To prevent this, the principal can either introduce large rewards that compensate the agent for facing this risk, or she can completely eradicate the possibility of losses. The former solution is expensive since losses loom larger than equally sized gains by a factor of $\lambda$. 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 can include incentives that depend on performance in the same way as the contracts described in Proposition 2, 3, 4, or 5. Depending on the agent's probability perception in gains, the shape of one of these contracts applies for the performance levels $q > \hat{q}$. That is because in the domain of gains, the CPT agent exhibits risk attitudes equivalent to those of the RDU agent. Hence, 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 B1, 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 $\lambda > 2$.

# C Ambiguity Attitudes
# [FOR ONLINE PUBLICATION ONLY]

The theoretical framework can be easily extended to capture deviations from EUT due to attitudes toward ambiguity. This can be achieved using *Source Theory* (Abdellaoui et al., 2011; Baillon et al., 2025), which, broadly speaking, states that under ambiguity the phenomena of risk are amplified because there is "additional probability weighting." This additional weighting is included in the function $w_s$, referred to as the *source function*, which is endowed with the properties of Assumption 4. Importantly, the source function might exhibit a different shape than $w$ (the probability weighting function), and this difference in shapes generates ambiguity attitudes.

In order to incorporate ambiguity into the model, consider a setting in which the distribution $F(q|e)$ is unknown to the agent because, for example, he has limited experience with the task. In such a case, it is said that the agent is facing an "ambiguous task." An event in the ambiguous task is defined as a subset $E_i \subset [\underline{q}, \bar{q}]$. The collection of all such events is denoted by $\Sigma$, which I endow with the Borel $\sigma$-algebra.

A widely used framework in decision making under ambiguity is the Choquet Expected Utility model (Schmeidler, 1989). In this model, the agent evaluates events using the following utility function:

$$CEU(t,e) = \int_{[\underline{q},\bar{q}]} u\Big(t(q)\Big)\, \mathrm{d}W - c(e), \tag{C.1}$$

where $W$ is a weighting function, defined over events in $\Sigma$, that satisfies the following properties for a given effort level $e'$: $W(\emptyset|e') = 0$, $W\Big([\underline{q}, \bar{q}]|e'\Big) = 1$, and $W\Big(E_2|e'\Big) > W\Big(E_1|e'\Big)$ for any $E_1, E_2 \in \Sigma$ such that $E_1 \subset E_2$. A distinctive feature of this framework is that it nests the Max-Min Expected Utility model (Gilboa and Schmeidler, 1989) as a special case, arguably the most influential model in decision-making under ambiguity.[21]

The preference representation in (C.1) generalizes subjective expected utility by allowing $W$ to be non-additive. That property introduces ambiguity attitudes into the model. For instance, if $W$ is convex, the decision maker underweights the likelihood of events in the ambiguous task, which makes him averse to betting on these events as compared to betting on *equally likely* events in a risky task—for which the calculation of objective probabilities is possible.[22]

[21]Specifically, the Max-Min Expected Utility model is a special case of the Choquet Expected Utility model in which $W$ is assumed to be convex, that is $W(E_2) + W(E_1) \leq W(E_2 \cap E_1) + W(E_2 \cup E_1)$ holds for all $E_1, E_2 \subset \Sigma$.

[22]More precisely, let the event $E_i$ have a probability $p_i$. The ambiguity averse individual with preferences given in (C.1) perceives that $W(E_i) < p_i$. Therefore, when asked to choose between the bet $(100, E; 0, E^c)$ and the lottery $(100, p; 0, 1-p)$, he prefers the latter.

The main problem with this approach for modeling ambiguity is its high intractability, as the set of weighting functions $W$ that characterize preferences under ambiguity can be extensive (Wakker, 2010). This intractability complicates the main analysis of this paper, which involves comparing the agent's behavior in an ambiguous setting to his behavior in a hypothetical setting where the task is risky (probabilities are known).

To address this problem, I model decision-making under ambiguity with Source Theory. The key component of that theory is the *source of uncertainty*, which refers to a group of events generated by the same random mechanism (Abdellaoui et al., 2011; Tversky and Fox, 1995). Accordingly, the collection of all events in the ambiguous task is one source of uncertainty while the collection of all events in the case of a risky task is another one.

The agent is assumed to exhibit *probabilistic sophistication* within a given source of uncertainty but not necessarily across different sources (Chew and Sagi, 2008). Let $P_s$ denote the probability measure generated by $\Sigma$, the algebra of events in the case of an the ambiguous task. In contrast, when the task is risky, the probability measure in this source is the density function $F(q|e)$. This type of source-dependent sophistication allows for different attitudes toward probabilities arising from sources of uncertainty. In the case of an ambiguous task, there exists a function $w_s$ such that, for any $e$:

$$W(E|e) = w_s(1 - P_s(E)) \text{ for any } E \in [\underline{q}, \bar{q}]. \tag{C.2}$$

The function $w_s$, carries subjective probabilities to decision weights in the model. Importantly, its shape may differ from that of $w$, the probability weighting function which carries objective probabilities to decision weights. This difference in shape between $w_s$ and $w$ generates ambiguity attitudes, as will be discussed below.

Substituting (C.2) into (C.1) shows that in the case of an ambiguous task the agent's preferences are given by:

$$ST(t, e) = \int_{[\underline{q}, \bar{q}]} u\Big(t(q)\Big)\, \mathrm{d}w_s(1 - P_s(q|e)) - c(e). \tag{C.3}$$

Alternatively, when the agent is faced with a risky task, his preferences are given by (3).

Ambiguity attitudes are now incorporated into the model. The most recognized ambiguity attitude is ambiguity aversion, which, in the context of this framework, implies that the agent pays more attention to poor performance realizing in the ambiguous task relative to the same outcome realizing in a risky task. This stronger pessimism about outcomes realizing under ambiguity is captured by the model when $w_s$ is more convex than $w$, leading the agent to overweight worse outcomes more under ambiguity than under risk. In such case,, Proposition 2 characterizes the optimal contract while Corollary 2 demonstrates that stronger ambiguity aversion, i.e. $w_s$ becoming more convex while $w$ is kept fixed, leads to flatter contracts.

This extension of the model can therefore accommodate Ghirardato (1994)'s result that under stronger ambiguity aversion a high level of effort can be elicited with a "uniformly lower incentive scheme." Furthermore, when ambiguity aversion is sufficiently strong, Proposition 3 demonstrates that the principal has no other remedy but to give up incentive compatibility, a result that echoes the finding in Lang (2017) that a constant wage can be optimal.

Recent research shows that ambiguity aversion is not universal because individuals tend to display ambiguity-seeking behavior for extreme events (Abdellaoui et al., 2011; Baillon and Emirmahmutoglu, 2018; Trautmann and van de Kuilen, 2015). This pattern can be captured in the model by assuming that the probability weighting function under ambiguity, $w_s$, exhibitsa more pronounced inverse-S shape relative to $w$. An agent with such preferences is referred to as ambiguity insensitive (Baillon et al., 2018b). In the case of that agent, the optimal contract is characterized by Proposition 5. Furthermore, Corollary 6 implies that, under ambiguity, the optimal contract is flatter at intermediate performance levels and concentrates incentives at the extreme events, compared to the contract that would be optimal under risk.[23]

All in all, this extension reinforces the main finding of this paper. It discusses how in a more realistic setting, in which the agent does not exactly know performance probabilities, simpler contracts, i.e. with less incentives than in the case of risk, must be implemented. This is because ambiguity attitudes are modeled as "extra" probability weighting. Hence, the proposed explanation for the prevalence of simple contracts does not rely solely on probability weighting but can also be partly attributed to ambiguity attitudes.

[23] Note that this framework accommodates the desirability of "bait contracts" (Vierø, 2014). That is, the principal can take advantage of the agent's insensitivity to ambiguity by indexing performance to an ambiguous task rather than a risky one. This is because such a strategy enables her to pay less for likely events and more for unlikely ones.

# D Continuous Effort

# [FOR ONLINE PUBLICATION ONLY]

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

**Assumption D1** (cost of effort). *$c(e) : [\underline{e}, \bar{e}] \rightarrow [0, +\infty)$ is twice continuously differentiable with $c'(e) > 0$ and $c''(e) > 0$.*

Furthermore, I impose the following assumptions on the cumulative distribution function (CDF).

**Assumption D2** (output distribution). *$F(y|e) : [\underline{y}, \bar{y}] \rightarrow [0, 1]$ is twice continuously differentiable with respect to $e$ and $y$, and exhibits $F_{ee}(y|e) > 0$.*

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

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

**Assumption D3** (continuous WMLRP). $\frac{d}{dq}\left(\frac{\frac{d}{de}\left(w'(1-F(q|e))f(q|e)\right)}{w'(1-F(q|e)f(q|e))}\right) > 0$

A central implication of Assumption D3 is that it implies first-order stochastic dominance, $F_e(q|e) \leq 0$.

We are in a position to present a sufficient condition that guarantees the optimality of the contracts characterized in Section 4. The following lemma due to González-Jiménez (2024a) shows that when the agent's weighting function is convex everywhere, those contracts are optimal in a setting in which effort is continuous.

**Lemma D1.** *Assume that $F_{ee}(q|e) > 0$. The candidate solution generated the first-order approach is valid if $w''(p) \geq 0$ for all $p$.*

The condition presented in Lemma D1 is stringent. It implies that our results do not hold when the agent is optimistic or likelihood-insensitive. Therefore, we consider weaker requirements in order to validate the results of our standard model. The following lemma states that a sufficiently convex cost function can ensure the validity of the first-order approach even though the weighting function is not convex everywhere.

**Lemma D2.** *Assume that $F_{ee}(q|e) > 0$. For the first-order approach to be valid, it is necessary that $c''(e) > \mathcal{B}$, where:*

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

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

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

Denote by $t^r$ the solution to the following principal's program:

$$\begin{aligned} \max_{\{t(y)\}} \quad & \int_{\underline{y}}^{\bar{y}} \big(S(y) - t(y)\big) f(y|e)\mathrm{d}y \\ \text{s.t.} \quad & u\big(t(\underline{q})\big) - \int_{\bar{q}}^{\underline{q}} u'\big(t(q)\big)\frac{\mathrm{d}t(q)}{\mathrm{d}y} w\big(1 - F(q|e)\big)\mathrm{d}q - c(e) \geq \bar{U}, \\ & \int_{\bar{q}}^{\underline{q}} u'\big(t(q)\big)\frac{\mathrm{d}t(y)}{\mathrm{d}q} w_e\big(1 - F(q|e)\big)F_e(y|e)\mathrm{d}q - c'(e) \end{aligned} \tag{D.2}$$

In the above program, the incentive compatibility constraint is replaced by the first-order condition of Eq. (D.1) with respect to $e$. This approach is necessary and sufficient if the following condition holds:

$$\int_{\bar{q}}^{\underline{q}} u'\big(t(q)\big)\frac{\mathrm{d}t(q)}{\mathrm{d}q}\Big(w_e\big(1 - F(q|e)\big)F_{ee}(q|e) - w_{ee}\big(1 - F(q|e)\big)\big(F_e(q|e)\big)^2\Big)\mathrm{d}q - c''(e) < 0. \tag{D.3}$$

Since $c''(e) > 0$ (Assumption D1), $u' > 0$ (Assumption 3), $\frac{\mathrm{d}t(q)}{\mathrm{d}q} \geq 0$ (Assumption 1), the following condition suffices for the concavity of $RDU(t,e)$:

$$w'\big(1 - F(q|e)\big)F_{ee}(q|e) - w''\big(1 - F(q|e)\big)\big(F_e(q|e)\big)^2 < 0 \tag{D.4}$$

Due to $F_{ee}(q|e) > 0$ (Assumption D2) and $w'\big(1 - F(q|e)\big) > 0$ (Assumption 4), a probability weighting function that exhibits $w''\big(1 - F(y|e)\big) < 0$ cannot fulfill the condition in Eq. (D.4). Hence, for the optimality of $t^r$ it suffices that $w''\big(1 - F(q|e)\big) > 0$. Letting $p = 1 - F(q|e)$, that condition can be written as $w''(p) > 0$.

■

The Lemma shows that the first-order condition suffices to characterize the IC constraint when either the weighting function is sufficiently convex, so as to guarantee $\frac{w''\left(1-F(q|e)\right)F_e(q|e)}{w'\left(1-F(q|e)\right)} < \frac{F_{ee}(q|e)}{F_e(q|e)}$, or when the cost function, $c(e)$ is sufficiently convex. For simplicity, I assume that when $w''(p)$ is not sufficiently convex, $c(e)$ attains the bound presented in the above Lemma. If that were not the case, the principal might require other means to incentivize the agent. González-Jiménez (2024a) demonstrates that stochastic contracts are optimal when this condition does not hold.

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

**Proposition D1.** *The optimal contract without the IC constraint under optimism or likelihood insensitivity exhibit the shapes of the contracts described in Proposition 4 i) and Proposition 5 i).*

*Proof.* Denoting the Lagrange multiplier of the agent's participation constraint by $\nu$, the Lagrangian of the principal's unconstrained (without the M constraint) problem writes as:

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

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

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

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

The optimal effort level, $e^*$ satisfies

$$\int_{\bar{q}}^{\bar{q}} \Big(S(q) - t^{fb}(q)\Big) f_e(q|e^*)\mathrm{d}q + \\ \nu\left(-\int_{\underline{q}}^{\bar{q}} u'\Big(t^{fb}(q)\Big)\Big(w_e\Big(1 - F(q|e^*)\Big)\Big) F_e(q|e^*)\mathrm{d}q - c'(e^*)\right) = 0. \tag{D.6}$$

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

$$\int_{\bar{q}}^{\bar{q}} \Big(S(q) - t^{fb}(q)\Big) f_e(q|e^*)\mathrm{d}q = 0. \tag{D.7}$$

The solution of the principal's program is thus given by $\{(t^{fb}(q), e^*)\}$, where $t^{fb}(q)$ is the transfer satisfying Eq. (D.5) and $e^*$ is the effort level that satisfies Eq. (D.7).

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

$$t^{fb\prime}(q) = \frac{u'\Big(t^{fb}(q)\Big)}{u''\Big(t^{fb}(q)\Big)} \frac{w''\Big(1 - F(q|e^*)\Big)}{w'\Big(1 - F(q|e^*)\Big)} f(q|e). \tag{D.8}$$

This is exactly the equality in (A.22) when letting $\bar{e} = e^*$. The analyses of the shape of $t^{fb}(q)$ under optimism, and likelihood insensitivity in Propositions 4 i) and 5 i) immediately follow. ■

Consider now a setting of moral hazard. First, I show that when optimism or likelihood insensitivity are moderate, the first-best may suffice to elicit high effort levels. This solution is the analog of Proposition 4 if $c < \hat{c}_O$ and Proposition 5 if $c < \hat{c}_L$. However, as a direct consequence of considering a continuous action space, I condition on the values of $e^{**}$, the optimal effort level implemented by the principal, rather than on $c$.

**Proposition D2.** *Assume Optimism or Likelihood Insensitivity. There exists a unique effort level $\hat{e} \in [\underline{e}, \bar{e}]$ such that if $e^{**}$, the effort level implemented by the principal, is such that $e^{**} < \hat{e}$, the optimal second-best contract is the contract from Proposition D1.*

*Proof.* Denote by $\nu$ the Lagrange multiplier of the agent's participation constraint, and $\mu$, of the incentive compatibility constraint. The Lagrangian of the principal's unconstrained maximization problem (without the M constraint) writes as

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

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

$$\frac{1}{u'\Big(t^{sb}(q)\Big) w'\Big(1 - F(q|e)\Big)} = \nu + \mu \left( \frac{\frac{\mathrm{d}}{\mathrm{d}e}\Big(w'(1 - F(q|e)) f(q|e)\Big)}{w'(1 - F(q|e) f(q|e))} \right) \tag{D.9}$$

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

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

$$\begin{aligned}&\int_{\bar{q}}^{\bar{q}} \Big(S(q) - t^{fb}(q)\Big) f_e(q|e^{**}) \mathrm{d}q + \\ &\mu \left( - \int_{\underline{q}}^{\bar{q}} u'\Big(t^{fb}(q)\Big) \Big( w_e\Big(1 - F(q|e^{**})\Big) F_{ee}(q|e^{**}) - w_{ee}\Big(1 - F(q|e^{**})\Big) F_e(q|e^{**}) \Big) \mathrm{d}q - c''(e^{**}) \right) = 0.\end{aligned} \tag{D.10}$$

The solution of the principal's program is thus given by $\{(t^{sb}(q), e^{**})\}$, where $t^{sb}(q)$ is the transfer satisfying Eq. (D.9) and $e^{**}$ is the effort level that satisfies Eq. (D.10).

I next show that $\mu = 0$ can hold at the optimum under optimism or likelihood insensitivity and the solution to the principal's problem becomes $\{(t^{fb}(q), e^{**})\}$, where $t^{fb}(q)$ is the first-best contract presented in Proposition D1.

**Optimism** Consider the case of an agent with optimism in the sense of Definition 2. From the complementary slackness condition, we get that when $\mu = 0$, then

$$u'\Big(t(q)\Big)\frac{\mathrm{d}t(y)}{\mathrm{d}q}w'\Big(1 - F(q|e)\Big)F_e(y|e)\mathrm{d}q > c'(e) \tag{D.11}$$

Assumption D3 implies $F_e(q|e) < 0$ which, together with $\frac{\mathrm{d}t(q)}{\mathrm{d}q} > 0$ (Proposition D1), $w' > 0$ (Assumption 4) and $u'(t) > 0$ (Assumption 3), imply that the left-hand side of (D.11) is weakly positive, rendering the inequality in (Assumption 4) feasible.

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

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

Hence, for the effort levels $e \in [\underline{e}, \hat{e})$, the inequality in (D.11) holds .

**Likelihood insensitivity** For likelihood insensitivity $\frac{\mathrm{d}t(q)}{\mathrm{d}q} > 0$ (Proposition D1), so the inequality in (Assumption 4) is feasible. Since, $w''(p) < 0$ in $(0, \tilde{p})$, then $w'\Big(1 - F(q|e)\Big)F_e(q|e)$ is decreasing in that probability interval, which guarantees the existence of $\hat{e}$.

■

Second, it is shown that the contract shapes presented in Proposition 2, 4, and 5 continue to hold when effort is continuous.

**Proposition D3.** *The optimal second-best exhibits the shapes of the contracts presented in Proposition 2 under pessimism, or the contracts presented in Propositions 4 and 5 if $e^{**} > \hat{e}$ under either likelihood insensitivity or optimism.*

*Proof.* Assume $\mu > 0$. Differentiate (D.9) with respect to $q$ to obtain:

$$\begin{aligned} t^{sb\prime}(q) =& \frac{u'(t^{sb}(q))w''(1 - F(q|e^{**}))}{u''(t^{sb}(q))w'\left(1 - F(q|e^{**})\right)}f(q|e^{**}) \\ &+ \mu\frac{w'\left(1 - F(q|e^{**})\right)u'(t^{sb}(q))^2}{u''(t^{sb}(q))}\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{\frac{\mathrm{d}}{\mathrm{d}e}\Big(w'(1 - F(q|e))f(q|e)\Big)}{w'(1 - F(q|e)f(q|e))}\right). \end{aligned} \tag{D.12}$$

The above equation and Eq. (A.15) differ only in that $\bar{e}$ is now $e^{**}$ and the discrete MLRP is replaced by its continuous analog. Therefore, the analysis of $t^{sb\prime}(q)$ is similar to that presented in Proposition 4.

Under optimism, $w''(p) < 0$ for all $p \in (0, 1)$ implies that both terms in Eq. (D.12) are positive, implying that $t^{sb}$ is everywhere increasing. Moreover, since $\lim_{q \to \bar{q}} w'(p) = +\infty$

and $\lim_{q\to\underline{q}} w'(p) = 0$, then $t^{sb'}(q) \to +\infty$ at both extremes.

Under pessimism, $w''(p) > 0$ for all $p \in (0,1)$. Hence, the first term in the right-hand side of Eq. (D.12) is negative, while the second one is positive. Due to $\lim_{q\to\bar{q}} w'(p) = 0$, then $\lim_{q\to\bar{q}} \frac{w''(p)}{w'(p)} = +\infty$; the first term in Eq. (D.12) dominates and $\lim_{q\to\bar{q}} t^{sb'}(q) = -\infty$.

Eq. (D.12) implies that $t^{sb'}(q) > 0$ under pessimism requires:

$$-\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{\frac{\mathrm{d}}{\mathrm{d}e}\Big(w'(1-F(q|e))f(q|e)\Big)}{w'(1-F(q|e)f(q|e))}\right) > \frac{w''(1-F(q|e))f(q|e)}{w'(1-F(q|e))}\left(\frac{1}{\mu w'\Big(1-F(q|e)\Big)u'(t^{sb}(q))}\right). \tag{D.13}$$

The W-MLRP gives

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{\frac{\mathrm{d}}{\mathrm{d}e}\Big(w'(1-F(q|e))f(q|e)\Big)}{w'(1-F(q|e)f(q|e))}\right) =& \frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{f_e(q|e)}{f(q|e)}\right)\\ &+\left(-\frac{\Big(w''(1-F(q|e))\Big)^2 F_e(q|e)f(q|e)}{\Big(w'(1-F(q|e))\Big)^2}\right.\\ &\left.-\frac{\Big(w''(1-F(q|e))\Big)f(q|e)}{w'(1-F(q|e))}\right),\end{aligned} \tag{D.14}$$

I use the above expression to rewrite Eq. (D.12) as

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{f_e(q|e)}{f(q|e)}\right) >& \frac{\Big(w''(1-F(q|e))\Big)^2 f(q|e)}{\Big(w'(1-F(q|e))\Big)^2}\left(-F_e(q|e)+\frac{1}{\mu u'(t^{sb}(q))w''(1-F(q|e))}\right)\\ &\frac{-\frac{\mathrm{d}}{\mathrm{d}e}\Big(w''(1-F(q|e))f(q|e)\Big)}{w'(1-F(q|e))}.\end{aligned} \tag{D.15}$$

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

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

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

$$\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{f_e(q|e)}{f(q|e)}\right) > \frac{\Big(w''(1-F(q|e))\Big)^2 f(q|e)}{\Big(w'(1-F(q|e))\Big)^2}\left(-F_e(q|e)+\frac{-\frac{\mathrm{d}}{\mathrm{d}e}\Big(w''(1-F(q|e))f(q|e)\Big)}{w'(1-F(q|e))}\right), \tag{D.16}$$

which holds from the WMLRP (See Eq. (D.14)). Therefore, there exists an output level $q_h \in (\underline{q}, \bar{q})$ such that $t^{sb'}(q) > 0$ if $q \in [\underline{q}, q_h)$ and $t^{sb'}(q) < 0$ otherwise. The method for bunching is the that given by Proposition 1 and Corollary 1.

■

## D.1 Stochastic contracts

Relaxing the assumption of binary effort might render the contracts characterized in Sections 4 and 5 suboptimal. We now study the optimal contracts that emerge in that case.

Lemma D1 shows that a necessary condition for invalidating the results in Sections 4 and 5 is that the agent's probability weighting function be concave for some interval in $p \in (0,1)$. Because that shape of the weighting function implies probabilistic risk-seeking attitudes, the optimal contract should include greater risk than that implicit in the solutions presented in sections 4 and 5. In a related model, González-Jiménez (2024a) shows that when the agent is an RDU maximizer and has probabilistic risk-seeking attitudes, it can be optimal to implement a stochastic contract of the form $L := (r, \frac{t(q)}{p}; 1-r, 0)$ where $r \in (0,1)$ is a probability chosen by the principal.

We complement the findings of González-Jiménez (2024a) by fully characterizing the optimal stochastic contract. This is accomplished in two ways. First, we use the results of our standard model to define the outcome of the stochastic contract $t(q)$. For example, when the agent is optimistic, that outcome is described by Proposition 4 and when the agent is likelihood-insensitive it is described by Proposition 5. Second, we specify the probability that must be included in the optimal contract. The following proposition summarizes that characterization.

**Proposition D4.** *Suppose that the condition presented in Lemma D2 does not hold. Then, the optimal contract is a binary lottery* $(r^*, \frac{t^{sb}(q)}{r^*}; 1-r^*, 0)$, *where:*

*(i)* $r^*$ *satisfies* $w(r^*) = \frac{\int_{\bar{q}}^{q} u(t) dw\left(1-F(q|e)\right)}{\int_{\bar{q}}^{q} u(\frac{t}{r^*}) dw\left(1-F(q|e)\right)}$ *and* $w(r^*) > r^*$*; and*

*(ii)* $t^{sb}$ *depends on performance as in Proposition 4 if the agent is an optimist, or as in Proposition 5 if the agent is likelihood-insensitive.*

*Proof.* The utility of an agent when he is offered the contract $(r, \frac{t}{r}; 1-r, 0)$ is

$$-\int_{\bar{q}}^{q} u\left(\frac{t}{r}\right) \mathrm{d}w\Big(r\big(1-F(q|e)\big)\Big) - c(e). \quad \text{(D.17)}$$

Under the assumption of reduction of compound lotteries, the utility can be rewritten as

$$-w(r)\int_{\bar{q}}^{q} u\left(\frac{t}{r}\right) \mathrm{d}w\big(1-F(q|e)\big) - c(e). \quad \text{(D.18)}$$

The first-order condition of the above equation with respect to $r$ is:

$$-w'(r)\int_{\bar{q}}^{q} u\left(\frac{t}{r}\right)\mathrm{d}w\Big(1-F(q|e)\Big)+w(r)\int_{\bar{q}}^{q} u'\left(\frac{t}{r}\right)\frac{t}{r^2}\mathrm{d}w\Big(1-F(q|e)\Big)=0. \tag{D.19}$$

Some albegraic manipulations yield:

$$\frac{w'(r)}{w(r)}=\frac{\int_{\bar{q}}^{q} u'\left(\frac{t}{r}\right)\frac{t}{r^2}\mathrm{d}w\Big(1-F(q|e)\Big)}{\int_{\bar{q}}^{q} u\left(\frac{t}{r}\right)\mathrm{d}w\Big(1-F(q|e)\Big)}. \tag{D.20}$$

This equality can be written as a the following ordinary differential equation:

$$\frac{\mathrm{d}}{\mathrm{d}r}\Big(\ln(w(r)\Big)=-\frac{\mathrm{d}}{\mathrm{d}r}\ln\left(\int_{\bar{q}}^{q} u\left(\frac{t}{r}\right)\mathrm{d}w\Big(1-F(q|e)\Big)\right) \tag{D.21}$$

Integrating both sides over $[r_0,1]$ gives:

$$-\ln\Big(w(r_0)\Big)=-\ln\left(\int_{\bar{q}}^{q} u\left(t\right)\mathrm{d}w\Big(1-F(q|e)\Big)\right)\Big)+\ln\left(\int_{\bar{q}}^{q} u\left(\frac{t}{r_0}\right)\mathrm{d}w\Big(1-F(q|e)\Big)\right), \tag{D.22}$$

which can be rewritten as

$$w(r_0)=\frac{\int_{\bar{q}}^{q} u\left(t\right)\mathrm{d}w\Big(1-F(q|e)\Big)}{\int_{\bar{q}}^{q} u\left(\frac{t}{r_0}\right)\mathrm{d}w\Big(1-F(q|e)\Big)}. \tag{D.23}$$

Letting $r_0=r^*$. We obtain the first part of the Proposition.

The second order condition of the agent's utility with respect to $r$ is:

$$\begin{aligned}-w''(r^*)\int_{\bar{q}}^{q} u\left(\frac{t}{r^*}\right)\mathrm{d}w\Big(1-F(q|e)\Big)+2\Big(w'(r^*)-\frac{w(r)}{r^*}\Big)\int_{\bar{q}}^{q} u'\left(\frac{t}{r^*}\right)\frac{t}{r^{*2}}\mathrm{d}w\Big(1-F(q|e)\Big)\\ -w(r^*)\int_{\bar{q}}^{q} u''\left(\frac{t}{r^*}\right)\frac{t^2}{r^{*4}}\mathrm{d}w\Big(1-F(q|e)\Big)\end{aligned} \tag{D.24}$$

Since $u''<0$, the last expression in the equation above is negative. The first expression is negative if $w''\Big(r^*\Big)<0$. Moreover, the second expression is positive if

$$w'\Big(r^*\Big)>\frac{w\Big(r^*\Big)}{r^*}\Leftrightarrow\int\frac{\mathrm{d}}{\mathrm{d}r}\ln\Big(w(r)\Big)\mathrm{d}r>\int\frac{\mathrm{d}}{\mathrm{d}r}\ln\Big(r\Big)\mathrm{d}r \tag{D.25}$$

which holds if $w(r^*)>r^*$. ■

The optimal stochastic contract enables the principal to elicit high effort levels by taking advantage of the agent's probability weighting. This is first achieved with a probability that will be overweighted by the agent, i.e. $w(r^*)>r^*$, which guarantees that the agent displays probabilistic risk-seeking behavior and a preference for risky contracts. Moreover, the condition $w(r^*) = \frac{u(t^{sb}(q))}{u\left(\frac{t^{sb}(q)}{r^*}\right)}$ ensures that the agent optimally inflates the perceived benefits from exerting high effort when working under the contract $(r^*, \frac{t^{sb}(q)}{r^*}; 1 - r^*, 0)$.

# E Adverse Selection followed by Moral Hazard [FOR ONLINE PUBLICATION ONLY]

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). I refer to these agents as $L$.

The principal knows that she contracts with a EUT agent with probability $\pi_E$ 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 E1.** *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 $w'\left(1 - F(q|\bar{e})\right) \leq 1$.*

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

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_{EU}(q)\Big) f(q|\bar{e}) \,\mathrm{d}q - c \geq \int_{\underline{q}}^{\bar{q}} u\Big(t_{EU}(q)\Big) f(q|\underline{e}) \,\mathrm{d}q, \tag{E.1}$$

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

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e}) \,\mathrm{d}q - c \geq \int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) w'\Big(1 - F(q|\underline{e})\Big) f(q|\underline{e}) \,\mathrm{d}q. \tag{E.2}$$

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:

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_{EU}(q)\Big) f(q|\bar{e})\, \mathrm{d}q - c \geq \max_{e \in \{\underline{e}, \bar{e}\}} \left\{ \int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) f(q|\bar{e})\, \mathrm{d}q - c(e) \right\}, \tag{E.3}$$

and for the non-EUT agent:

$$\begin{aligned} &\int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})\, \mathrm{d}q - c \\ &\geq \max_{e \in \{\underline{e}, \bar{e}\}} \left\{ \int_{\underline{q}}^{\bar{q}} u\Big(t_{EU}(q)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})\, \mathrm{d}q - c(e) \right\}. \end{aligned} \tag{E.4}$$

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

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_{EU}(q)\Big) f(q|\bar{e})\, \mathrm{d}q - c \geq 0, \tag{E.5}$$

and

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})\, \mathrm{d}q - c \geq 0. \tag{E.6}$$

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:

$$\begin{aligned} \int_{\underline{q}}^{\bar{q}} w' \left(1 - F(q|\bar{e})\right) f(q|\bar{e}) \mathrm{d}q - \int_{\underline{q}}^{\bar{q}} w' \left(1 - F(q|\underline{e})\right) f(q|\underline{e}) \mathrm{d}q = \\ w\left(1 - F(q|\bar{e})\right) - w\left(1 - F(q|\underline{e})\right). \end{aligned} \tag{E.7}$$

Instead, for the EU agent, efficiency amounts to:

$$\int_{\underline{q}}^{\bar{q}} f(q|\bar{e}) \mathrm{d}q - \int_{\underline{q}}^{\bar{q}} f(q|\underline{e}) \mathrm{d}q = \left(1 - F(q|\bar{e})\right) - \left(1 - F(q|\underline{e})\right). \tag{E.8}$$

The W-MLRP (Assumption 5) 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 (E.7) to be larger than (E.8) is $w'(1 - F(q|e)) > 1$ for any $e$. That is because

$$\int_{1-F(q|\underline{e})}^{1-F(q|\bar{e})} w'(s)\mathrm{d}s > \int_{1-F(q|\underline{e})}^{1-F(q|\bar{e})} \mathrm{d}s \Leftrightarrow$$
$$w\left(1-F(q|\bar{e})\right) - w\left(1-F(q|\underline{e})\right) > F(q|\underline{e}) - F(q|\bar{e}) \tag{E.9}$$

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. I first reduce the number of constraints to solve the principal's problem. Equations (E.5) and (E.4) immediately imply (E.6). Hence, at the optimum the participation constraint in (E.5) binds, while the participation constraint in (E.6) slacks.

From equation (E.3) and the constraint in (E.5), which binds at the optimum, we obtain:

$$0 \geq \max_{e\in\{\underline{e},\bar{e}\}} \left\{ \int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) f(q|\bar{e})\, \mathrm{d}q - c(e) \right\}, \tag{E.10}$$

which implies that EUT agents cannot afford to mimic non-EUT agents. Hence, the relevant adverse selection constrain is that in (E.4), which states that the non-EUT agent derives rents from mimicking the EUT agent. In contrast, equation (E.3) slacks at the optimum.

A direct implication that (E.4) binds is $t_L(q) \geq t_{EU}(q)$, which in turn gives

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_{EU}(q)\Big) f(q|\bar{e})\, \mathrm{d}q - c > \int_{\underline{q}}^{\bar{q}} u\Big(t_{EU}(q)\Big) f(q|\underline{e})\, \mathrm{d}q. \tag{E.11}$$

Hence, the moral hazard constraint in (E.1) slacks at the optimum.

Next, from the inequality in (E.6), which slacks at the optimum, along with equation (E.10), which holds with strict inequality, I obtain:

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) w'\Big(1-F(q|\bar{e})\Big) f(q|\bar{e})\, \mathrm{d}q - c > 0 \geq \max_{e\in\{\underline{e},\bar{e}\}} \left\{ \int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) f(q|\bar{e})\, \mathrm{d}q - c(e) \right\}. \tag{E.12}$$

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

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) f(q|\underline{e})\, \mathrm{d}q > \int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) w'\left(1 - F(q|\underline{e})\right) f(q|\underline{e})\, \mathrm{d}q, \tag{E.13}$$

Equations (E.12) and (E.13) imply

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})\, \mathrm{d}q - c > \int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) w'\left(1 - F(q|\underline{e})\right) f(q|\underline{e})\, \mathrm{d}q. \tag{E.14}$$

and equation (E.2) is implied by other constraints in the principal's program.

Hence, at the solution only equations (E.4) and (E.5) 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 (E.5). 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 (E.4).

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.

■

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 the optimal contract from Holmstrom (1979), 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 E1 (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 E1 (1) becomes relevant in this case.

# References


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

_ , "A Genuine Rank-Dependent Generalization of the von Neumann-Morgenstern Expected Utility Theorem," *Econometrica*, 2002, *70* (2), 717–736.

**_ , 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 Cédric Gutierrez**, "Unpacking Overconfident Behavior When Betting on Oneself," *Management Science (forthcoming)*, 2023.

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

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

**_ , _ , Chen Li, and Peter P Wakker**, "Source theory: A tractable and positive ambiguity theory," Technical Report, Working Paper 2023.

**_ , _ , _ , and _** , "Source theory: A tractable and positive ambiguity theory," *Management Science*, 2025.

**_ , _ , Umut Keskin, Olivier l'Haridon, and Chen Li**, "The effect of learning on ambiguity attitudes," *Management Science*, 2018, *64* (5), 2181–2198.

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

**Barberis, Nicholas**, "A model of casino gambling," *Management Science*, 2012, *58* (1), 35–51.

**Bellemare, Charles and Bruce Shearer**, "Sorting, incentives and risk preferences: Evidence from a field experiment," *Economics Letters*, 2010, *108* (3), 345–348.

**Bonin, Holger, Thomas Dohmen, Armin Falk, David Huffman, and Uwe Sunde**, "Cross-sectional earnings risk and occupational sorting: The role of risk attitudes," *Labour Economics*, 2007, *14* (6), 926–937.

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

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

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

**Choi, Syngjoo, Jeongbin Kim, Eungik Lee, and Jungmin Lee**, "Probability weighting and cognitive ability," *Management Science*, 2022, *68* (7), 5201–5215.

**Curme, Michael and Noah Stefanec**, "Worker quality and labor market sorting," *Economics Letters*, 2007, *96* (2), 202–208.

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

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

**Dimmock, Stephen G., Roy Kouwenberg, Olivia S. Mitchell, and Kim Peijnenburg**, "Household portfolio underdiversification and probability weighting: Evidence from the field," *The Review of Financial Studies*, 2021, *34* (9), 4524–4563.

**Dohmen, Thomas, Armin Falk, Klaus Fliessbach, Uwe Sunde, and Bernd Weber**, "Relative versus absolute income, joy of winning, and gender: Brain imaging evidence," *Journal of Public Economics*, 2011, *95* (3-4), 279–285.

**Ebert, Sebastian and Philipp Strack**, "Until the bitter end: On prospect theory in a dynamic context," *American Economic Review*, 2015, *105* (4), 1618–1633.

**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.

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

**Gilboa, Itzhak and David Schmeidler**, "Maxmin expected utility with non-unique prior," *Journal of mathematical economics*, 1989, *18* (2), 141–153.

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

**González-Jiménez, Víctor**, "Incentive contracts when agents distort probabilities," *Quantitative Economics*, 2024, *15* (3), 607–653.

_ , "Incentive design for reference-dependent preferences," *Journal of Economic Behavior & Organization*, 2024, *221*, 493–518.

**Grund, Christian and Dirk Sliwka**, "Evidence on performance pay and risk aversion," *Economics Letters*, 2010, *106* (1), 8–11.

**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.

**Henderson, Vicky, David Hobson, and SL Alex**, "Randomized strategies and prospect theory in a dynamic context," *Journal of Economic Theory*, 2017, *168*, 287–300.

**Herweg, 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.

**Humphery-Jenner, Mark, Ling Lei Lisic, Vikram Nanda, and Sabatino Dino Silveri**, "Executive overconfidence and compensation structure," *Journal of financial Economics*, 2016, *119* (3), 533–558.

**Joseph, Kissan and Manohar U Kalwani**, "The Role of Bonus Pay in Salesforce Compensation Plans," *Industrial Marketing Management*, 1998, *27* (2), 147–159.

**Jullien, Bruno**, "Participation constraints in adverse selection models," *Journal of Economic Theory*, 2000, *93* (1), 1–47.

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

**Karni, Edi and David Schmeidler**, "Atemporal dynamic consistency and expected utility theory," *Journal of Economic Theory*, 1991, *54* (2), 401–408.

**Kilka, Michael and Martin Weber**, "What determines the shape of the probability weighting function under uncertainty?," *Management science*, 2001, *47* (12), 1712–1726.

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

**Laffont, Jean Jacques and David Martimort**, *The theory of incentives: The Principal-Agent Model* number January 2001, Princeton University Press, 2002.

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

**Larkin, Ian and Stephen Leider**, "Incentive schemes, sorting, and behavioral biases of employees: Experimental evidence," *American Economic Journal: Microeconomics*, 2012, *4* (2), 184–214.

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

**Lewis, Tracy R. and David E. Sappington**, "Countervailing incentives in agency problems," *Journal of economic theory*, 1989, *49* (2), 294–313.

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

**Lichtenstein, Sarah and Paul Slovic**, "Reversals of preference between bids and choices in gambling decisions.," *Journal of experimental psychology*, 1971, *89* (1), 46.

**Loomes, Graham and Robert Sugden**, "Regret theory: An alternative theory of rational choice under uncertainty," *The economic journal*, 1982, *92* (368), 805–824.

**Machina, Mark J**, "Dynamic consistency and non-expected utility models of choice under uncertainty," *Journal of Economic literature*, 1989, *27* (4), 1622–1668.

**Maestas, Nicole, Kathleen J Mullen, David Powell, Till Von Wachter, and Jeffrey B Wenger**, "The value of working conditions in the United States and implications for the structure of wages," *American Economic Review*, 2023, *113* (7), 2007–2047.

**Maggi, Giovanni and Andres Rodriguez-Clare**, "On countervailing incentives," *Journal of Economic Theory*, 1995, *66* (1), 238–263.

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

**Otto, Clemens A.**, "CEO optimism and incentive compensation," *Journal of Financial Economics*, 2014, *114* (2), 366–404.

**Oyer, Paul**, "A theory of sales quotas with limited liability and rent sharing," *Journal of labor Economics*, 2000, *18* (3), 405–426.

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

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

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

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

**Rottenstreich, Yuval and Christopher K Hsee**, "Money, kisses, and electric shocks: On the affective psychology of risk," *Psychological science*, 2001, *12* (3), 185–190.

**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.

**Sautmann, Anja**, "Contracts for agents with biased beliefs: Some theory and an experiment," *American Economic Journal: Microeconomics*, 2013, *5* (3), 124–156.

**Schmeidler, David**, "Subjective probability and expected utility without additivity," *Econometrica: Journal of the Econometric Society*, 1989, pp. 571–587.

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

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

**Trautmann, Stefan T. and Gijs van de Kuilen**, "Belief Elicitation: A Horse Race among Truth Serums," *Economic Journal*, 2015, *125* (589), 2116–2135.

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

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

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

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

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

**Wakker, Peter P.**, "Separating Marginal Utility and Probabilistic Risk Aversion," *Theory and Decision*, 1994, *36* (1), 1–44.

_ , "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.

**Zeisberger, Stefan, Dennis Vrecko, and Thomas Langer**, "Measuring the time stability of prospect theory preferences," *Theory and Decision*, 2012, *72*, 359–386.