Online Appendix for “Optimal Incentives without Expected Utility”

Víctor González-Jiménez

January, 2023

Contents

1 Appendix B: Prospect Theory Preferences

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

For simplicity, we assume that the reference point r is assumed to be exogenous to the alternatives faced by the decision-maker. For instance, it can be the agent's current wealth at the moment of making decisions (Kahneman and Tversky, 1979; Tversky and Kahneman, 1981). This reference point rule has been recently validated empirically by Baillon et al. (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 1. *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 1 introduces loss aversion. That is, transfers counting as losses loom larger than equally-sized transfers counting as gains. This latter property is captured by the parameter $\lambda > 1$ and expresses a special dislike for losses.

The second ingredient is that the probability weighting function is defined separately over gains and losses. Probabilities associated with gains are transformed by the probability weighting function w, introduced in Assumption 3. On the other hand, probabilities associated with losses are transformed with a probability weighting function z that applies transformations to cumulative probabilities, $F(q|e)$, rather than to decumulative probabilities.[1]

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

Assumption 2. *A probability weighting function for losses is a function* $z : [0, 1] \to [0, 1]$ *satisfying the duality condition* $z(F(q|e)) = 1 - w(1 - F(q|e))$ *for any* e*.*

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

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

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

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

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

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

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

Proposition 1. *Let Assumptions 3 to 2 hold. There exists a threshold $\hat{q} \in [\underline{q}, \bar{q}]$ such that the second best-contract, t_C^{sb}:*

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

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

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

Proof. Rewrite Eq. (1) using Assumption 2 as

$$CPT(t, e, r) = \int_{\underline{q}}^{\bar{q}} \Big[\theta v\left(t(q) - r\right) w'\Big(1 - F(q|e)\Big) \\ - \lambda(1-\theta)v(r - t(q))w'\Big(1 - F(q|e)\Big)\Big] f(q|e)\,\mathrm{d}q - c(e), \quad (2)$$

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

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

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

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

Denote by $t_C^{sb}(q)$ the transfer satisfying Eq. (4). We show next that a lottery $L = (p, r; 1-p, 0)$ improves upon the solution $t_C^{sb}(q)$ whenever $0 < t_C^{sb}(q) < r$. Since $-\lambda v\big(r - t_C^{sb}(q)\big)$ is increasing in $t_C^{sb}(q)$, there exists a number $\rho \in [0,1]$ for each realization q such that

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

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

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

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

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

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

The above equation is not linear in $\bar{L}$ due to w having curvature (Assumption 3). Hence, changes in $\bar{L}$ affect marginal utility. To understand how changes in $\bar{L}$ affect the marginal incentives of offering the lottery, we compute the first-order condition of (7) with respect to ρ, which gives us

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

Denote by ρ^{opt} the probability satisfying the condition in (8). The second-order condition evaluated at ρ^{opt} is

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

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

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

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

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

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

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

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

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

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

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

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

Where the left-hand side of (13) denote the marginal incentives of offering $L_{\rho=1}$. The existence and uniqueness of $\hat{q}$ is guaranteed by the fact that the left-hand side of Eq. (13) of is positive and constant in q while the right-hand side of that equation increases with q (Assumption 4) over $[0, +\infty)$.

There are three cases. When $\frac{\lambda v(r)}{r}$ is small and the right-hand side of (13) is large enough, then $\hat{q} \geq \bar{q}$. In that case $t_C^{sb} = r$. Alternatively, $\frac{\lambda v(r)}{r}$ can be large so that $\hat{q} \leq \bar{q}$ and the solution is fully given by Proposition 3 and Proposition 6, depending on the shape of w. Finally, if $\hat{q} \in [\underline{q}, \bar{q}]$ then

$$t_C^{sb}(q) = \begin{cases} r & \text{if } q < \hat{q}, \\ t_P^{sb}(q), t_O^{sb}(q) \text{ (Proposition 3), or } t_L^{sb}(q) \text{ (Proposition 6)} & \text{if } q \geq \hat{q}. \end{cases} \tag{14}$$

■

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

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

The contract characterized in Proposition 1, 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 $\lambda > 2$.

2 Appendix C: Continuous Effort

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

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

Assumption 4 (output distribution)**.** *$F(y|e) : [\underline{y}, \bar{y}] \to [0, 1]$ is $\mathcal{C}^2$ with respect to e and y, and exhibits $F_{ee}(y|e) > 0$.*

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

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

Assumption 5 (continuous WMLRP)**.** $\frac{d}{dq}\left(\frac{\frac{d}{de}\Big(w'(1-F(q|e))f(q|e)\Big)}{w'(1-F(q|e)f(q|e))}\right) > 0$

A central implication of Assumption 5 is that it implies first-order stochastic dominance, $F_e(q|e) \leq 0$.

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

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

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

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

$$RDU(t, e) = u\Big(t(\underline{q})\Big) - \int_{\bar{q}}^{\underline{q}} u'\Big(t(q)\Big)\frac{\mathrm{d}t(q)}{\mathrm{d}q}w\Big(1 - F(q|e)\Big)\mathrm{d}y - c(e). \qquad (15)$$

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

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

$$\text{s.t.} \qquad u\Big(t(\underline{q})\Big) - \int_{\bar{q}}^{\underline{q}} u'\Big(t(q)\Big)\frac{\mathrm{d}t(q)}{\mathrm{d}y} w\Big(1 - F(q|e)\Big)\mathrm{d}q - c(e) \geq \bar{U}, \tag{16}$$

$$\int_{\bar{q}}^{\underline{q}} u'\Big(t(q)\Big)\frac{\mathrm{d}t(y)}{\mathrm{d}q} w_e\Big(1 - F(q|e)\Big)F_e(y|e)\mathrm{d}q - c'(e)$$

In the above program, the incentive compatibility constraint is replaced by the first-order condition of Eq. (15) with respect to e. This approach is necessary and sufficient if the following condition holds:

$$\int_{\bar{q}}^{\underline{q}} u'\Big(t(q)\Big)\frac{\mathrm{d}t(q)}{\mathrm{d}q}\bigg(w_e\Big(1-F(q|e)\Big)F_{ee}(q|e) - w_{ee}\Big(1-F(q|e)\Big)\Big(F_e(q|e)\Big)^2\bigg)\mathrm{d}q - c''(e) < 0. \tag{17}$$

Since $c''(e) > 0$ (Assumption 3), $u' > 0$ (Assumption 1), $\frac{\mathrm{d}t(q)}{\mathrm{d}q} \geq 0$ (Assumption 2), the following condition suffices for the concavity of $RDU(t, e)$:

$$w_e\Big(1 - F(q|e)\Big)F_{ee}(q|e) - w_{ee}\Big(1 - F(q|e)\Big)\Big(F_e(q|e)\Big)^2 < 0 \tag{18}$$

Due to $F_{ee}(q|e) > 0$ (Assumption 4) and $w_e\Big(1 - F(q|e)\Big) > 0$ (Assumption 3), a probability weighting function that exhibits $w_{ee}\Big(1 - F(y|e)\Big) < 0$ cannot fulfill the condition in Eq. (18). Hence, for the optimality of t^{fo} it suffices that $w_{ee}\Big(1 - F(q|e)\Big) > 0$. Letting $p = 1 - F(q|e)$, that condition can be written as $w''(p) > 0$.

■

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

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

Proposition 2. *The optimal first-best under optimism, pessimism, or likelihood insensitivity exhibit the shapes of the contracts presented in Proposition 1 and Proposition 5.*

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

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

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

$$\frac{1}{u'\Big(t^{fb}(q)\Big) w'\Big(1 - F(q|e)\Big)} = \nu. \tag{19}$$

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

The optimal effort level, e^* satisfies

$$\begin{aligned}&\int_{\bar{q}}^{\bar{q}} \Big(S(q) - t^{fb}(q)\Big) f_e(q|e^*) \mathrm{d}q + \\ &\qquad \nu \left(- \int_{\underline{q}}^{\bar{q}} u'\Big(t^{fb}(q)\Big) \Big(w_e\Big(1 - F(q|e^*)\Big)\Big) F_e(q|e^*) \mathrm{d}q - c'(e^*) \right) = 0.\end{aligned} \tag{20}$$

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:

$$\int_{\bar{q}}^{\bar{q}} \Big(S(q) - t^{fb}(q)\Big) f_e(q|e^*) \mathrm{d}q = 0. \tag{21}$$

The solution of the principal's program is thus given by $\{(t^{fb}(q), e^*)\}$, where $t^{fb}(q)$ is the transfer satisfying Eq. (19) and e^* satisfies Eq. (21).

To investigate the shape of $t^{fb}(q)$ we differentiate (19) with respect to q, giving us

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

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

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

Proposition 3. *Assume Optimism or Likelihood Insensitivity. There exists a unique effort level $\hat{e} \in [\underline{e}, \bar{e}]$ such that if e^{**}, the effort level implemented by the principal, is such that $e^{**} < \hat{e}$, the optimal contract is the first-best contract from Proposition 2.*

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

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

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

$$\frac{1}{u'\Big(t^{sb}(q)\Big)w'\Big(1-F(q|e)\Big)} = \nu + \mu\left(\frac{\frac{\mathrm{d}}{\mathrm{d}e}\Big(w'(1-F(q|e))f(q|e)\Big)}{w'(1-F(q|e)f(q|e))}\right) \tag{23}$$

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

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

$$\int_{\bar{q}}^{\bar{q}}\Big(S(q)-t^{fb}(q)\Big)f_e(q|e^{**})\mathrm{d}q+$$
$$\mu\left(-\int_{\underline{q}}^{\bar{q}}u'\Big(t^{fb}(q)\Big)\Big(w_e\Big(1-F(q|e^{**})\Big)F_{ee}(q|e^{**})-w_{ee}\Big(1-F(q|e^{**})\Big)F_e(q|e^{**})\Big)\mathrm{d}q-c''(e^{**})\right)=0. \tag{24}$$

The solution of the principal's program is thus given by $\{(t^{sb}(q), e^{**})\}$, where $t^{sb}(q)$ is the transfer satisfying Eq. (23) and e^{**} satisfies Eq. (24).

We next show that $\mu > 0$ might not hold the optimum under optimism or likelihood insensitivity and the solution to the principal's problem becomes $\{(t^{fb}(q), e^{**})\}$. Suppose instead that $\mu = 0$. Accordingly, $t^{sb}(q) = t^{fb}(q)$, where $t^{fb}(q)$ is the first-best contract presented in Proposition 2.

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

$$u'\Big(t(q)\Big)\frac{\mathrm{d}t(y)}{\mathrm{d}q}w_e\Big(1-F(q|e)\Big)F_e(y|e)\mathrm{d}q > c'(e) \tag{25}$$

Assumption 5 implies $F_e(q|e) < 0$ which, together with $\frac{\mathrm{d}t(q)}{\mathrm{d}q} > 0$ (Proposition 2), $w' > 0$ (Assumption 3) and $u'(t) > 0$ (Assumption 1), imply that the left-hand side of (25) is weakly positive, rendering the inequality in (Assumption 3) feasible.

The right-hand side of (25) is increasing because $c'(0) = 0$ and $c''(e) > 0$. Also, because $w_e\Big(1-F(q|e)\Big)F_e(q|e)$ is decreasing (Lemma 1) the left-hand side of (25) is decreasing. Hence, there exists an effort level $\hat{e} \in [\underline{e}, \bar{e}]$ such that

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

Hence, for $e \in [\underline{e}, \hat{e})$, the inequality in (25) holds .

Likelihood insensitivity For likelihood insensitivity $\frac{\mathrm{d}t(q)}{\mathrm{d}q} > 0$ (Proposition 2) so the inequality in (Assumption 3) is feasible. Since, $w''(p) > 0$ in $(0, \tilde{p})$, then $w_e\Big(1-F(q|e)\Big)F_e(q|e)$ is decreasing (Lemma 1) in that probability interval, the existence of $\hat{e}$ is guaranteed.

■

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

Proposition 4. *The optimal second-best exhibits the shapes of the contracts presented in Proposition 3 and Proposition 6 under pessimism, or if $e^{**} > \hat{e}$ and under either likelihood insensitivity or optimism.*

Proof. Assume $\mu > 0$. Differentiate (23) with respect to q to obtain:

$$\begin{aligned} t^{sb'}(q) =& \frac{u'(t^{sb}(q))w''(1-F(q|e^{**}))}{u''(t^{sb}(q))w'\left(1-F(q|e^{**})\right)}f(q|e^{**}) \\ &+ \mu\frac{w'\left(1-F(q|e^{**})\right)u'(t^{sb}(q))^2}{u''(t^{sb}(q))}\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{\frac{\mathrm{d}}{\mathrm{d}e}\Big(w'(1-F(q|e))f(q|e)\Big)}{w'(1-F(q|e)f(q|e))}\right). \end{aligned} \tag{26}$$

The above equation and Eq. (30) differ only in that $\bar{e}$ is now e^{**} and the discrete MLRP is replaced by its continuous analog. Therefore, the analysis of $t^{sb'}(q)$ is similar to that presented in Proposition 3.

Under optimism, $w''(p) < 0$ for all $p \in (0,1)$ implies that both terms in Eq. (26) 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. (26) 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. (26) dominates and $\lim_{q\to\bar{q}} {t^{sb}}'(q) = -\infty$.

Eq. (26) implies that ${t^{sb}}'(q) > 0$ under pessimism requires:

$$-\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{\frac{\mathrm{d}}{\mathrm{d}e}\Big(w'(1-F(q|e))f(q|e)\Big)}{w'(1-F(q|e)f(q|e))}\right) > \frac{w''(1-F(q|e))f(q|e)}{w'(1-F(q|e))}\left(\frac{1}{\mu w'\Big(1-F(q|e)\Big)u'(t^{sb}(q))}\right). \tag{27}$$

The W-MLRP gives

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{\frac{\mathrm{d}}{\mathrm{d}e}\Big(w'(1-F(q|e))f(q|e)\Big)}{w'(1-F(q|e)f(q|e))}\right) =& \frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{f_e(q|e)}{f(q|e)}\right) \\ &+\left(-\frac{\Big(w''(1-F(q|e))\Big)^2 F_e(q|e)f(q|e)}{\Big(w'(1-F(q|e))\Big)^2}\right. \\ &\left.-\frac{\Big(w''(1-F(q|e))\Big)f(q|e)}{w'(1-F(q|e))}\right),\end{aligned} \tag{28}$$

we use the above expression to rewrite Eq. (26) as

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{f_e(q|e)}{f(q|e)}\right) >& \frac{\Big(w''(1-F(q|e))\Big)^2 f(q|e)}{\Big(w'(1-F(q|e))\Big)^2}\left(-F_e(q|e)+\frac{1}{\mu u'(t^{sb}(q))w''(1-F(q|e))}\right) \\ &\frac{-\frac{\mathrm{d}}{\mathrm{d}e}\Big(w''(1-F(q|e))f(q|e)\Big)}{w'(1-F(q|e))}.\end{aligned} \tag{29}$$

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

$$\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{f_e(q|e)}{f(q|e)}\right) > \frac{\Big(w''(1-F(q|e))\Big)^2 f(q|e)}{\Big(w'(1-F(q|e))\Big)^2}\left(-F_e(q|e) + \frac{-\frac{\mathrm{d}}{\mathrm{d}e}\Big(w''(1-F(q|e))f(q|e)\Big)}{w'(1-F(q|e))}\right), \tag{30}$$

which holds from the WMLRP (See Eq. (28)). Therefore, there exists an output level $q_h \in (\underline{q}, \bar{q})$ such that ${t^{sb}}'(q) > 0$ if $q \in [\underline{q}, q_h)$ and ${t^{sb}}'(q) < 0$ otherwise. The method for ironing is the same as in Proposition 3.

■

3 Appendix D: Adverse selection

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). We refer to these agents as L.

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

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

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

Proposition 5. *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

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_{EU}(q)\Big) f(q|\bar{e})\,\mathrm{d}q - c \geq \int_{\underline{q}}^{\bar{q}} u\Big(t_{EU}(q)\Big) f(q|\underline{e})\,\mathrm{d}q, \tag{31}$$

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

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})\,\mathrm{d}q - c \geq \int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) w'\Big(1 - F(q|\underline{e})\Big) f(q|\underline{e})\,\mathrm{d}q. \tag{32}$$

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:

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_{EU}(q)\Big) f(q|\bar{e})\,\mathrm{d}q - c \geq \max_{e\in\{\underline{e},\bar{e}\}} \left\{ \int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) f(q|\bar{e})\,\mathrm{d}q - c(e) \right\}, \tag{33}$$

and for the non-EUT agent:

$$\begin{aligned} &\int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})\,\mathrm{d}q - c \\ &\geq \max_{e\in\{\underline{e},\bar{e}\}} \left\{ \int_{\underline{q}}^{\bar{q}} u\Big(t_{EU}(q)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})\,\mathrm{d}q - c(e) \right\}. \end{aligned} \tag{34}$$

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

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_{EU}(q)\Big) f(q|\bar{e})\,\mathrm{d}q - c \geq 0, \tag{35}$$

and

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})\,\mathrm{d}q - c \geq 0. \tag{36}$$

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

$$\int_{\underline{q}}^{\bar{q}} w'\left(1-F(q|\bar{e})\right) f(q|\bar{e})\mathrm{d}q - \int_{\underline{q}}^{\bar{q}} w'\left(1-F(q|\underline{e})\right) f(q|\underline{e})\mathrm{d}q = \\ w\left(1-F(q|\bar{e})\right) - w\left(1-F(q|\underline{e})\right). \tag{37}$$

Instead, for the EU agent, efficiency amounts to:

$$\int_{\underline{q}}^{\bar{q}} f(q|\bar{e})\mathrm{d}q - \int_{\underline{q}}^{\bar{q}} f(q|\underline{e})\mathrm{d}q = \left(1-F(q|\bar{e})\right) - \left(1-F(q|\underline{e})\right). \tag{38}$$

The W-MLRP (Assumption 4) implies both $F(q|\bar{e}) < F(q|\underline{e})$ and $w\left(1-F(q|\bar{e})\right) > w\left(1-F(q|\underline{e})\right)$.

A sufficient condition for (37) to be larger than (38) is $w'(1-F(q|e)) > 1$ for any e. That is because

$$\int_{1-F(q|\underline{e})}^{1-F(q|\bar{e})} w'(s)\mathrm{d}s > \int_{1-F(q|\underline{e})}^{1-F(q|\bar{e})} \mathrm{d}s \Leftrightarrow \\ w\left(1-F(q|\bar{e})\right) - w\left(1-F(q|\underline{e})\right) > F(q|\underline{e}) - F(q|\bar{e}) \tag{39}$$

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

Suppose the non-EUT agent is more efficient. As shown above, this mainly happens when the agent's possible actions generate probabilities that are located at extremes of the output interval. We first reduce the number of constraints to solve the principal's problem. Equations (35) and (34) immediately imply (36). Hence, at the optimum the participation constraint in (35) binds, while the participation constraint in (36) slacks.

From equation (33) and the constraint in (35), which binds at the optimum, we obtain:

$$0 \geq \max_{e\in\{\underline{e},\bar{e}\}} \left\{ \int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) f(q|\bar{e})\,\mathrm{d}q - c(e) \right\}, \tag{40}$$

which implies that EUT agents cannot afford to mimic non-EUT agents. Hence, the relevant adverse selection constrain is that in (34), which states that the non-EUT agent derives rents from mimicking the EUT agent. In contrast, equation (33) slacks at the optimum.

A direct implication that (34) binds is $t_L(q) \geq t_{EU}(q)$, which in turn gives

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_{EU}(q)\Big) f(q|\bar{e})\,\mathrm{d}q - c > \int_{\underline{q}}^{\bar{q}} u\Big(t_{EU}(q)\Big) f(q|\underline{e})\,\mathrm{d}q. \tag{41}$$

Hence, the moral hazard constraint in (31) slacks at the optimum.

Next, from the inequality in (36), which slacks at the optimum, along with equation (40), which holds with strict inequality, we obtain:

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) w'\Big(1 - F(q|\bar{e})\Big) f(q|\bar{e})\,\mathrm{d}q - c > 0 \geq \max_{e\in\{\underline{e},\bar{e}\}} \left\{ \int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) f(q|\bar{e})\,\mathrm{d}q - c(e) \right\}. \tag{42}$$

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

$$\int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) f(q|\underline{e})\,\mathrm{d}q > \int_{\underline{q}}^{\bar{q}} u\Big(t_L(q)\Big) w'\left(1 - F(q|\underline{e})\right) f(q|\underline{e})\,\mathrm{d}q, \tag{43}$$

Equations (42) and (43) imply

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

and equation (32) is implied by other constraints in the principal's program.

Hence, at the solution only equations (34) and (35) 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 (35). 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 (34).

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 Proposition 2, or as described by Proposition 6. This guarantees that incentives are given according to they way in which each type perceives output realizations. Importantly, to guarantee self-selection into the right contract, informational rents are included in one of the contracts. Specifically, the contract that targets the most efficient type is embellished with rents to discourage mimicking.

So far this solution seems standard. However, whether one agent is more efficient than the other crucially depends on probability weighting. When the agent's actions yield high and/or low probability, the agent suffering from likelihood insensitivity inflates the impact of his action on the probability of obtaining higher output levels. In that case, this irrational agent is more efficient; he is more likely to exert high effort with lower pay. In this situation, the menu in Proposition 5 (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 5 (1) becomes relevant in this case.

References

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

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

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

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

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

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

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

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