Online Appendix for “Incentive design under reference dependence”

Victor Gonzalez-Jimenez

January. 2023

Contents

Appendix B: Additional Proofs Section 4

Corollary 3.

Let $\eta = 0$ and $\phi = 1$. Under these restrictions, $U(e_H, w(\tilde{y}), r)$ is concave for any output realization $\tilde{y} \in \left[\underline{y}, \overline{y}\right]$ since $-\frac{u''(w(\tilde{y}))}{u'(w(\tilde{y}))} \geq 0$. Contradicting the necessary and sufficient condition given in Eq. $(A3)$. Consequently, the contracts satisfying Eq. $(A6)$ and Eq. $(A7)$ are sufficient to solve the principal's maximization problem.

Under the assumed restrictions, $(A6)$ and $(A7)$ become:

$$\frac{1}{u'(w_s^*(y))} = \mu + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right). \qquad \text{(A40)}$$

Rearranging the above expression leads to:

$$w_s^*(y) = h'\left(\frac{1}{\mu + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right)}\right). \qquad \text{(A41)}$$

Moreover, equation (A8) becomes:

$$\frac{dw_s^*(y)}{dy} = \frac{\gamma \frac{d}{dy}\left(\frac{f(y|e_L)}{f(y|e_H)}\right)\left(u'(w_s^F(y))\right)^2}{u''\left(w_s^F(y)\right)}. \qquad \text{(A42)}$$

Since $\frac{d}{dy}\left(\frac{f(y|e_L)}{f(y|e_H)}\right) \leq 0$ (Assumption 2) and $u'' < 0$, then $\frac{dw_s^*(y)}{dy} \geq 0$. The optimal second-best contract, given in $(A41)$, is everywhere increasing in performance. This proves the second part of the corollary.

We turn to study the first-best optimal contract. Therefore, in addition to $\eta = 0$ and $\phi = 1$, let $\gamma = 0$. In that case, $(A31)$ becomes

$$w_f^*(y) = h'\left(\frac{1}{\mu}\right). \qquad \text{(A43)}$$

Equation $(A43)$ shows that $\frac{dw_f^*(y)}{dy} = 0$. Full insurance is given to the agent with a performance-insensitive contract. This proves the first part of the corollary. ■

Corollary 4.

Let $\eta = 1$, $\phi = 0$, and $u' = 1$. Under these restrictions, $U(e_H, w(\tilde{y}), r)$ is S-shaped for any realization $\tilde{y} \in [\underline{y}, \overline{y}]$. In the domain of losses, that is when $\theta_{\mathbb{I}} = 0$, $-\frac{v''\left(u(r)-u(w(\tilde{y}))\right)}{v'\left(u(r)-u(w(\tilde{y}))\right)} \geq 0$, corroborating the necessary and sufficient condition in Eq. $(A3)$. Hence, only the solution from the first-order approach for the domain of gains is sufficient and necessary to solve the principal's problem. Under the assumed restrictions that solution, given by Eq. $(A6)$, becomes

$$\frac{1}{v'(w_s^*(y) - r)} = \mu + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right). \qquad \text{(A44)}$$

Rearranging the above expression, the following closed-form expression is obtained:

$$w_s^*(y) = r + j'\left(\frac{1}{\mu + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right)}\right). \qquad \text{(A45)}$$

Moreover, Eq. (A8) becomes

$$\frac{dw_s^*(y)}{dy} = \frac{\gamma \frac{d}{dy}\left(\frac{f(y|e_L)}{f(y|e_H)}\right)\left(v'\ (w_s^F(y) - r)\right)^2}{v''(w_s^F(y) - r)}. \qquad (A46)$$

Since $\frac{d}{dy}\left(\frac{f(y|e_L)}{f(y|e_H)}\right) \leq 0$ (Assumption 2) and $v'' < 0$, then $\frac{dw_s^*(y)}{dy} \geq 0$ if $\theta_{\mathbb{I}} = 1$.

Eq. $(A21)$ becomes:

$$\frac{1}{\frac{\lambda v(u(r))}{r}} = \mu + \gamma\left(1 - \frac{f(\hat{y}_{ps}|e_L)}{f(\hat{y}_{ps}|e_H)}\right). \qquad (A47)$$

So, the transition from losses to gains is given by the $\hat{y}_{ps}$ that satisfies $(A47)$. The existence and uniqueness of that output level is guaranteed by $\frac{1}{\frac{\lambda v(u(r))}{r}} > 0$, $\frac{1}{\frac{\lambda v(u(r))}{r}}$ being constant in performance, and $\frac{d}{dy}\left(\frac{f(y|e_L)}{f(y|e_H)}\right) \leq 0$ (Assumption 2).

Furthermore, Theorem 1 shows that it is optimal to pay the lowest possible in the domain of losses, that is $w_s^*(y) = 0$ in $y < \hat{y}_{ps}$.

All in all, the optimal contract is given by $w_s^*(y) = \begin{cases} 0 & if\ y < \hat{y}_{ps}, \\ r + j'\left(\frac{1}{\mu+\gamma\left(1-\frac{f(y|e_L)}{f(y|e_H)}\right)}\right) & if\ y \geq \hat{y}_{ps}. \end{cases}$ This proves the second part of the corollary.

We turn to study the first-best contract. Therefore, in addition to $\eta = 1$, $\phi = 0$, and $u' = 1$, let $\gamma = 0$. In that case, Eq. $(A45)$ becomes

$$w_s^*(y) \; = r + j'\left(\frac{1}{\mu}\right). \qquad \text{(A48)}$$

Eq. $(A48)$ shows that $\frac{dw_f^*(y)}{dy} = 0$ if $\theta_{\mathbb{I}} = 1$. Moreover, under the considered restrictions, Eq. $(A30)$ becomes:

$$\int_{\hat{y}_{p1}}^{\overline{y}} v\big(w_f^*(y) - r\big) f(y|e_H)\, \mathrm{d}y - \lambda \int_{\underline{y}}^{\hat{y}_{pf}} v(r) f(y|e_H) \mathrm{d}y - c \; = \overline{U}. \qquad (A49)$$

The transition from losses to gains is given by the $\hat{y}_f$ that satisfies $(A49)$. The existence of $\hat{y}_{pf} \in (\underline{y}, \overline{y})$ is guaranteed by the fact that $\int_{\hat{y}_{pf}}^{\overline{y}} v\big(w_f^*(y) - r\big) f(y|e_H)\, \mathrm{d}y$ is positive, while $-\lambda \int_{\underline{y}}^{\hat{y}_{pf}} v(r) f(y|e_H) \mathrm{d}y$ is negative. The lower and upper limit of those integrals, $\hat{y}_{pf}$, can be adjusted to obtain a positive expression in the left-hand side of $(A49)$ equal to $\overline{U} \geq 0$. Also, that both $w_f^*(y)$ satisfying Eq. (A48) and the lottery payment L make the agent's participation constraint bind, as shown by Eqs. (A33), (A34), and (A27), guarantees the existence of $\hat{y}_{pf}$. Finally, Theorem 1 shows that $w_s^*(y) = 0$ in $y < \hat{y}_{pf}$.

Therefore, the optimal contract is given by $w_f^*(y) = \begin{cases} 0 & if \; y < \hat{y}_{pf}, \\ r + j'\left(\frac{1}{\mu}\right) & if \; y \geq \hat{y}_{pf}. \end{cases}$ This proves the first part of the corollary. ■

Corollary 5.

The agent makes, at most, two choices: accepting or rejecting the contract and choosing an effort level. These choices generate the set of candidates for max-min reference point $r = \max\{\min\{\overline{U}\}, \min\{w_s^*(y)\}\}$. Intuitively, rejecting the contract yields welfare $\overline{U}$ and the minimum that the contract pays, regardless of effort level, is $\min\{w_s^*(y)\}$.

Since the participation constraint binds at the optimum, then $\mathbb{E}\big(U(e, w_s^*(y), r)\big) = \overline{U}$. Also, because the incentive compatibility constraint binds, it must be that $\mathbb{E}\big(U(e, w_s^*(y), r)\big) > \mathbb{E}\big(U(r, \min\{w_s^*(y)\}, r)\big)$. Otherwise, the second-best contract would not implement punishments for low performance and thus would not incentivize high effort. Hence, $r = w_s^*(y)$

Corollary 4 presented the solution to the principal's problem when the agent exhibits prospect theory preferences with an exogenous r. Since the agent's preferences are still characterized by prospect theory, the solution presented therein remains optimal once $r = w_s^*(y)$ is accounted for.

To that end, I first define the output level after which the agent is awarded the bonus. Let $\hat{y}_u$ be the output level satisfying:

$$\frac{1}{\frac{\lambda v\left(w_s^*(y)\right)}{w_s^*(y)}} = \mu + \gamma\left(1 - \frac{f(\hat{y}_u|e_L)}{f(\hat{y}_u|e_H)}\right). \qquad (A50)$$

The above condition is analog to that presented in Eq. $(A47)$ when $r = w_s^*(y)$. The existence and uniqueness of that output level is guaranteed by $\frac{1}{\frac{\lambda v\left(w_s^*(y)\right)}{w_s^*(y)}} > 0$, $\frac{d}{dy}\left(\frac{1}{\frac{\lambda v\left(w_s^*(y)\right)}{w_s^*(y)}}\right) = \frac{v(w_s^*(y)) - w_s^*(y)v'(w_s^*(y))}{\lambda\left(v(w_s^*(y))\right)^2} < 0$ due to Taylor's theorem around zero, and $-\frac{d}{dy}\left(\frac{f(y|e_L)}{f(y|e_H)}\right) \geq 0$ (Assumption 2).

According to Corollary 4 if $y \geq \hat{y}_u$, the agent must be transitioned to the domain of gains with a contract satisfying Eq. (A44). Adapting that incentive schedule to account for the considered reference point rule gives:

$$\frac{1}{\int_{\underline{y}}^{\overline{y}} v'\left(w_s^*(y) - w_s^*(\tilde{\tilde{y}})\right) f\left(\tilde{\tilde{y}}\middle|e\right) d\tilde{\tilde{y}}} = \mu + \gamma\left(1 - \frac{f(y|e_H)}{f(y|e_L)}\right), \qquad (A51)$$

for $\tilde{\tilde{y}} \in [\underline{y}, \overline{y}]$. Also, Corollary 4 states that $w_s^*(y) = 0$ if $y < \hat{y}_u$. The reference point rule does not affect that level of payment. Therefore, the optimal contract is given by:

$$w_s^*(y) = \begin{cases} 0 & if\ y < \hat{y}_u, \\ w_s^*(y) \text{ satisfying } (A51) & if\ y \geq \hat{y}_u. \end{cases}$$

■

To prove Proposition 1, we first show that, under certain conditions on the agent's incentive scheme, goals motivate high effort at the cost of generating disutility.

Lemma 2. *Under A1-A4 and that the agent's preferences are characterized by Eq.* (4)*, higher goals:*

i. *generate disutility if* $w'(g) > 0$*;*
ii. *motivate high effort if* $w'(g) > 0$*,* $w'(y) > 0$ *for* $y > g$ *, and* $w'(y) \leq 0$ *for* $y \leq g$*.*

Proof. Compute the derivative of Eq. (4), with respect to g to obtain:

$$\frac{\mathrm{d}U(e,w(y),g)}{\mathrm{d}g} = -\int_{g}^{\overline{y}} w'(g)v'\big(w(y)-w(g)\big)f(y|e)dy - \lambda\int_{\underline{y}}^{g} w'(g)v'\big(w(g)-w(y)\big)f(y|e)\mathrm{d}y. \qquad (A52)$$

Since $v' > 0$ and $\lambda > 1$, Eq. $(A52)$ is weakly negative if $w'(g) \geq 0$. In that case, higher goals induce disutility. This proves the first part of the lemma.

Integration by parts applied to Eq. (4) gives:

$$U(e,w(y),g) = v\big(w(\overline{y})-w(g)\big) - \int_{g}^{\overline{y}} w'(y)v'\big(w(y)-w(g)\big)F(y|e)\mathrm{d}y - \lambda\int_{\underline{y}}^{g} w'(y)u'\big(w(g)-w(y)\big)\,F(y|e)dy - c(e). \qquad (A53)$$

The agent's incentive compatibility constraint can be rewritten using Eq.$(A53)$ as

$$-\int_{g}^{\overline{y}} w'(y)v'\big(w(y)-w(g)\big)\big(F(y|e_H)-F(y|e_L)\big)\mathrm{d}y - \lambda\int_{\underline{y}}^{g} w'(y)v'\big(w(g)-w(y)\big)\,\big(F(y|e_H)-F(y|e_L)\big)\mathrm{d}y \geq c. \quad (A54)$$

To investigate whether higher goals incentivize high effort, derive $(A54)$ with respect to g to obtain:

$$-(\lambda-1)w'(g)v'\big(w(g)-w(g)\big)\big(F(g|e_H)-F(g|e_L)\big) + \int_{g}^{\overline{y}} w'(y)w'(g)v''\big(w(y)-w(g)\big)\big(F(y|e_H)-F(y|e_L)\big)\mathrm{d}y - \lambda\int_{\underline{y}}^{g} w'(y)w'(g)v''\big(w(g)-w(y)\big)\,\big(F(y|e_H)-F(y|e_L)\big)\mathrm{d}y. \qquad (A55)$$

Since $v' > 0$, $v'' < 0$, and $F(y|e_L) \geq F(y|e_H)$, an implication of Assumption 2, Eq.$(A55)$ shows that higher goals generate higher effort only if $w'(g) > 0$, $w'(y) > 0$ in $y \geq g$ and $w'(y) \leq 0$ in $y < g$. ■

Proposition 1.

Part i). Let $\eta = 1$, $\phi = 0$, $u' = 1$, and $r = w(g)$. Under these restrictions, $U(e_H, w(\tilde{y}), w(g))$ is S-shaped for any realization $\tilde{y} \in [\underline{y}, \overline{y}]$ since $-\frac{v''\left(u(w(g))-u(w(\tilde{y}))\right)}{v'\left(u(w(g))-u(w(\tilde{y}))\right)} \geq 0$, corroborating the condition in Eq. ($A3$). Therefore, the solution from the first-order condition is only sufficient and necessary for the domain of gains. That condition, given by Eq. ($A6$), rewrites under the considered restrictions as:

$$\frac{1}{v'(w_S^*(y)-w_S^*(g))}=\mu+\gamma\left(1-\frac{f(y|e_L)}{f(y|e_H)}\right). \qquad \text{(A56)}$$

Rearranging (A56) gives

$$w_S^*(y)=w_S^*(g)+j'\left(\frac{1}{\mu+\gamma\left(1-\frac{f(y|e_L)}{f(y|e_H)}\right)}\right). \qquad \text{(A57)}$$

Eq. ($A57$) shows that the requirement $w'(g)>0$ from Lemma 2, immediately implies $\frac{dw_S^*(y)}{dg}>0$. Moreover, the derivative of Eq. (A56) with respect to y gives $\frac{dw_S^F(y)}{dy}=\frac{\gamma\frac{d}{dy}\left(\frac{f(y|e_L)}{f(y|e_H)}\right)\left(v'\left(w_S^F(y)-w_S^F(g)\right)\right)^2}{v''\left(w_S^F(y)-w_S^F(g)\right)}\geq 0$. Therefore, the second-best optimal contract must increase both in goals and performance in the domain of gains.

To study the location of the bonus, Eq. ($A21$) is rewritten to account for the considered restrictions, yielding:

$$\frac{1}{\frac{\lambda v(w_S^*(g))}{w_S^*(g)}}=\nu+\gamma\left(1-\frac{f(\hat{y}_g|e_L)}{f(\hat{y}_g|e_H)}\right). \qquad (A58)$$

The output level $\hat{y}_g\in\left[\underline{y},\overline{y}\right]$ that satisfies ($A58$) is the critical threshold that transitions the agent from losses to gains. The existence and uniqueness of that output level is guaranteed by $\frac{1}{\frac{\lambda v(w_S^*(g))}{w_S^*(g)}}>0$, $\frac{d}{dy}\left(\frac{1}{\frac{\lambda v(w_S^*(g))}{w_S^*(g)}}\right)=0$, and Assumption 2.

Theorem 1 also demonstrates that, for an agent with S-shaped utility function, paying $w_S^*(y)=0$ in $y<\hat{y}_g$ is second-best optimal. Accordingly, the second-best contract is

$$w_S^*(y)=\begin{cases}0 & if\ y<\hat{y}_g,\\ w_S^*(g)+j'\left(\frac{1}{\mu+\gamma\left(1-\frac{f(y|e_L)}{f(y|e_H)}\right)}\right) & if\ y\geq\hat{y}_g.\end{cases}$$

This solution exhibits a discrete jump at $y = \hat{y}_g$, as $\lim_{y \to \hat{y}_g^+} w_s^*(y) = w_s^*(g)$ and $\lim_{y \to \hat{y}_g^-} w_s^*(y) = 0$. Moreover, a consequence of committing to an incentives scheme with $w_s^{*\prime}(g) > 0$, a condition that Lemma 2 shows necessary for goals to be motivating, is that the bonus must become larger as g increases.

Finally, I demonstrate that $\hat{y}_g = g$. Proceed by contradiction by assuming that $\hat{y}_g < g$. In that case, the principal overinsures the agent in $y \in [\hat{y}_g, g]$, a segment where he is risk seeking due to $U(e_H, w(y), r)$ being S-shaped. The principal could increase her profits by setting $w_s^*(y) = 0$ and, due to the convexity of $U(e_H, w(\tilde{y}), r)$ in losses, the agent would be willing to accept that contract change. Now suppose $\hat{y}_g > g$. The agent is in that case overexposed to risk in $y \in [g, \hat{y}_g]$, where he is risk averse. This potentially leads the agent to reject the contract. Hence, to ensure that the contract is accepted, the principal offers $w_s^*(y)$ given by $(A57)$ for any $y \geq g$. It must be then that $\hat{y}_g = g$.

The result $\hat{y}_g = g$ has three relevant implications. First, since $\lim_{y \to g^+} w_s^*(y) = w_s^*(g)$ and $\lim_{y \to g^-} w_s^*(y) = 0$, a bonus of size $w_s^*(g)$ is given at $y = g$. Second, to fulfill $w_s^{*\prime}(g) > 0$ from Lemma 2, that bonus increases with the size of g. Notice that this is consistent with the result of Corollary 2, obtained when the reference point was assumed to be exogenous. Thirdly, the conditions $w_s^{*\prime}(y) > 0$ if $y \geq g$ and $w_s^{*\prime}(y) = 0$ if $y < g$ from Lemma 2 are met. Consequently, it is second-best optimal to offer

$$w_s^*(y) = \begin{cases} 0 & if\ y < g, \\ w_s^*(g) + j'\left(\dfrac{1}{\mu + \gamma\left(1 - \dfrac{f(y|e_L)}{f(y|e_H)}\right)}\right) & if\ y \geq g. \end{cases}$$

Part ii). Under $\eta = 1$, $\phi = 0$, $u' = 1$, and $r = w(g)$, the derivative of Eq. $(A4)$ with respect to g gives:

$$\begin{aligned} &\mu\Big(-w'(g)\theta_{\mathbb{I}} v'\big(w(y) - w(g)\big) f(y|e) - \lambda(1 - \theta_{\mathbb{I}}) w'(g) v'\big(w(g) - w(y)\big) f(y|e)\Big) \\ &\qquad + \gamma\Big(-w'(g)\theta_{\mathbb{I}} v'\big(w(y) - w(g)\big)\big(f(y|e_H) - f(y|e_L)\big) \\ &\qquad - \lambda w'(g)(1 - \theta_{\mathbb{I}}) v'\big(w(g) - w(y)\big)\big(f(y|e_H) - f(y|e_L)\big)\Big) = 0. \qquad (A59) \end{aligned}$$

Denoting by g^* the goal level that satisfies $(A59)$, the following conditions are obtained:

$$-w'(g)\Big(\mu v'\big(w(y) - w(g^*)\big) f(y|e) + \gamma u'\big(w(y) - w(g^*)\big)\big(f(y|e_H) - f(y|e_L)\big)\Big) = 0, \quad (A60)$$

if $\theta_{\mathbb{I}} = 1$, and

$$-\lambda w'(g)\left(\mu v'\left(w(g^*)-w(y)\right)f(y|e)+\gamma v'\left(w(g^*)-w(y)\right)\left(f(y|e_H)-f(y|e_L)\right)\right)=0. \quad (A61)$$

if $\theta_{\mathbb{I}}=0$. Eqs. $(A60)$ and $(A61)$ imply:

$$\mu+\gamma\left(1-\frac{f(y|e_L)}{f(y|e_H)}\right)=0. \qquad (A62)$$

Eqs. $(A56)$ and $(A62)$ are used to obtain:

$$v'(w_S^*(y)-w_S^*(g^*))=+\infty. \qquad (A63)$$

Since $\lim_{y\to g^+} w_S^*(y)=w_S^*(g)$, then, due to the Inada condition, $\lim_{y\to g^+} v'\left(w_S^*(y)-w_S^*(g^*)\right)=+\infty$. Because the realization of y is ex-ante unknown to the principal, she cannot set a goal that is just met to comply with $(A63)$. However, she can set a goal such that $(A63)$ holds on expectation. Using the fact that $v'(\cdot)$ is concave, then

$$\mathbb{E}\left(v'(w_S^*(y)-w_S^*(g)\right)\leq v'\left(\mathbb{E}\left(w_S^*(y)\right)-w_S^*(g)\right). \quad (A64)$$

Therefore, g^* can be set such that $\mathbb{E}\left(w_S^*(y)\right)-w_S^*(g^*)=\epsilon$ for arbitrarily small $\epsilon>0$. This gives $v'\left(\mathbb{E}\left(w_S^*(y)\right)-w_S^*(g^*)\right)=+\infty.$■

Corollary 6.

Let $\overline{\overline{w}}\coloneqq\mathbb{E}\left(w(y)\right)$. Assume first that $\phi=1$, $u'=1$, and $r=\overline{\overline{w}}$. The agent's utility $U(e_H,w(\tilde{y}),\overline{\overline{w}})$ is S-shaped for any $\tilde{y}\in\left[\underline{y},\overline{y}\right]$ since in the domain of losses$-\frac{v''\left(u(\overline{\overline{w}})-u(w(\tilde{y}))\right)}{v'\left(u(\overline{\overline{w}})-u(w(\tilde{y}))\right)}\geq 0$, corroborating Eq. $(A3)$. Therefore, the solution from the first-order approach is only necessary and sufficient for the domain of gains.

That solution, given in Eq. $(A6)$, becomes:

$$\frac{1}{\left(1+\eta v'\left(w_S^*(y)-\overline{\overline{w}}\right)\right)}=\mu+\gamma\left(1-\frac{f(y|e_L)}{f(y|e_H)}\right). \qquad (A65)$$

Algebraic manipulations of the above equation yield $w_S^*(y)=\overline{\overline{w}}+j'\left(\frac{1}{\eta}\left(\frac{1}{\mu+\gamma\left(1-\frac{f(y|e_L)}{f(y|e_H)}\right)}-1\right)\right)$. Moreover, implicit derivation of Eq. $(A65)$ gives:

$$\frac{dw_S^*(y)}{dy}=\frac{\gamma\frac{d}{dy}\left(\frac{f(y|e_L)}{f(y|e_H)}\right)\left(1+\eta v'\left(w_S^*(y)-\overline{\overline{w}}\right)\right)^2}{\left(\eta v''\left(w_S^*(y)-\overline{\overline{w}}\right)\right)}. \qquad (A66)$$

Since $\frac{d}{dy}\left(\frac{f(y|e_L)}{f(y|e_H)}\right) \leq 0$ and $v'' < 0$, then $\frac{dw_s^*(y)}{dy} \geq 0$. The optimal increases in performance in the domain of gains.

The transition from losses to gains is given by Eq. $(A21)$, which, under the assumed restrictions, is equal to:

$$\frac{1}{\frac{1+\lambda v(\bar{\bar{w}})}{\bar{\bar{w}}}} = \mu + \gamma\left(1 - \frac{f(\hat{y}_{ms}|e_L)}{f(\hat{y}_{ms}|e_H)}\right). \qquad (A67)$$

The bonus is given at $y = \hat{y}_{ms}$ for $\hat{y}_{ms} \in (\underline{y}, \overline{y})$ satisfying Eq. $(A67)$. According to Theorem 1, it is second-best optimal to set $w_s^*(y) = 0$ in $y < \hat{y}_{ms}$ and the solution from the first-order approach elsewhere. Hence, the optimal contract is

$$w_s^*(y) = \begin{cases} 0 & if\ y < \hat{y}_{ms}, \\ \bar{\bar{w}} + j'\left(\frac{1}{\eta}\left(\frac{1}{\mu + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right)} - 1\right)\right) & if\ y \geq \hat{y}_{ms}. \end{cases}$$

This proves the first part of the corollary.

Next, I to study the first-best contract. Consider $\gamma = 0$ in addition to $\phi = 1$, $u' = 1$, and $r = \bar{\bar{w}}$. Eq. $(A65)$ becomes $\frac{1}{\left(1+\eta v'\left(w_f^*(y)-\bar{\bar{w}}\right)\right)} = \mu$, which after some manipulations yields $w_f^*(y) = \bar{\bar{w}} + j'\left(\frac{1}{\eta}\left(\frac{1}{\mu} - 1\right)\right)$. According to Eq. $(A66)$, that transfer exhibits $\frac{dw_f^*(y)}{dy} = 0$. Furthermore, Theorem 1, shows that is first-best optimal to provide this fixed-transfer for high output levels.

Theorem 1 also demonstrates that is first-best optimal to offer $w_f^*(y) = 0$ at the low-end of the output space. The transition from $w_f^*(y) = 0$ to $w_f^*(y) = \bar{\bar{w}} + j'\left(\frac{1}{\eta}\left(\frac{1}{\mu} - 1\right)\right)$ is given by the following equality, which adapts Eq. $(A30)$ to account for the assumed restrictions:

$$\int_{\hat{y}_{mf}}^{\overline{y}} w_f^*(y) f(y|e_H)\, \mathrm{d}y + \eta \int_{\hat{y}_{mf}}^{\overline{y}} v\left(w_f^*(y) - \bar{\bar{w}}\right) f(y|e_H)\, \mathrm{d}y - \lambda \int_{\underline{y}}^{\hat{y}_{mf}} v(\bar{\bar{w}}) f(y|e_H) \mathrm{d}y - c = \overline{U}. \quad (A68)$$

The output level $\hat{y}_{mf} \in [\underline{y}, \overline{y}]$ that satisfies Eq.$(A68)$ provides the transition from losses to gains. Theorem 1 shows that this output level is unique and interior. Hence, the first-best optimal contract is $w_f^*(y) = \begin{cases} 0 & if\ y < \hat{y}_{mf}, \\ \bar{\bar{w}} + j'\left(\frac{1}{\eta}\left(\frac{1}{\mu} - 1\right)\right) & if\ y \geq \hat{y}_{mf}. \end{cases}$ This proves the third part of the corollary.

Consider next $\phi = 1$, $v' = 1$, and $r = \overline{\overline{w}}$. Under these restrictions $U(e_H, w(\tilde{y}), \overline{\overline{w}})$ is concave for each $\tilde{y} \in [\underline{y}, \overline{y}]$ since $-\frac{u''(w(\tilde{y}))}{u'(w(\tilde{y}))} \geq 0$, contradicting the condition in Eq. $(A3)$. Hence, the solutions from the first-order approach are necessary and sufficient to solve the principal's maximization problem.

The first-order conditions from Eqs. $(A6)$ and $(A7)$ become

$$\frac{1}{u'(w_S^*(y))(1+\eta)} = \mu + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right), \qquad (A69)$$

and

$$\frac{1}{u'(w_S^*(y))(1+\eta\lambda)} = \mu + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right), \qquad (A70)$$

respectively. Eq. $(A69)$ can be expressed as $w_S^*(y) = h'\left(\frac{1}{(1+\eta)\left(\mu+\gamma\left(1-\frac{f(y|e_L)}{f(y|e_H)}\right)\right)}\right)$ and Eq. $(A60)$ as $w_S^F(y) = h'\left(\frac{1}{(1+\eta\lambda)\left(\mu+\gamma\left(1-\frac{f(y|e_L)}{f(y|e_H)}\right)\right)}\right)$. These two solutions exhibit $\frac{dw_S^*(y)}{dy} \geq 0$, as shown by Eqs. $(A8)$ and $(A9)$.

The transition from $w_S^*(y)$ satisfying $(A69)$ to $w_S^*(y)$ satisfying $(A70)$ is given by the unique output level $\hat{y}_{ms} \in (\underline{y}, \overline{y})$ that satisfies:

$$\int_{\hat{y}_{ms}}^{\overline{y}} u(W_S^*(y))f(y|e_H)\,\mathrm{d}y + \eta\int_{\hat{y}_{ms}}^{\overline{y}} \left(u(W_S^*(y)) - u(\overline{\overline{w}})\right)f(y|e_H)\,\mathrm{d}y + \int_{\underline{y}}^{\hat{y}_{ms}} u(W_S^*(y)f(y|e_H)\mathrm{d}y$$
$$- \lambda\eta\int_{\underline{y}}^{\hat{y}_{ms}} \left(u(\overline{\overline{w}}) - u\left(w_S^*(y)\right)\right)f(y|e_H)\mathrm{d}y - c = \overline{U}. \qquad (A71)$$

The existence of $\hat{y}_{ms}$ is guaranteed by the fact that the solutions from Eqs. $(A69)$ and $(A70)$ make the participation constraint bind for gains and losses, respectively. This is evident from Eqs. (A33), (A34). As a result, the optimal incentive scheme is given by:

$$w_f^*(y) = \begin{cases} h'\left(\dfrac{1}{(1+\eta)\left(\mu+\gamma\left(1-\dfrac{f(y|e_L)}{f(y|e_H)}\right)\right)}\right) & if\ y \geq \hat{y}_{ms}\ , \\ h'\left(\dfrac{1}{(1+\eta\lambda)\left(\mu+\gamma\left(1-\dfrac{f(y|e_L)}{f(y|e_H)}\right)\right)}\right) & if\ y < \hat{y}_{ms}. \end{cases} \qquad (A72)$$

This proves the second part of the corollary.

To conclude, we analyze the first-best contract when utility is concave. Add $\gamma = 0$ to the considered set of restrictions: $\phi = 1$, $v' = 1$, and $r = \overline{\overline{w}}$. The first-order condition in $(A69)$ becomes $\frac{1}{u'\left(w_f^*(y)\right)(1+\eta)} = \mu$, which after algebraic manipulations gives $w_f^*(y) = h'\left(\frac{1}{(1+\eta)\mu}\right)$. Eq. $(A8)$ shows that $\frac{dw_f^*(y)}{dy} = 0$ under the considered restrictions. Finally, Theorem 1 shows that paying $w_f^*(y) = h'\left(\frac{1}{(1+\eta)\mu}\right)$ everywhere is first-best optimal when $U(e_H, w(y), r)$ is concave. ■

Proposition 2.

i) Lagrangian and first-order conditions

Denote by $\mu \geq 0$ and $\gamma \geq 0$ the Lagrangian multipliers of the agent's participation and incentive compatibility constraints, respectively. The Lagrangian of the principal's maximization problem writes as

$$\begin{aligned}\mathcal{L} = &\big(S(y) - w(y)\big)f(y|e_H) \\ &+ \mu\Bigg(u(w(y))f(y|e_H) + \eta\theta_{\mathbb{I}}\int_{\underline{y}}^{\overline{y}} v\big(u(w(y) - u(w(\tilde{\tilde{y}}))f(y|e_H)f(\tilde{\tilde{y}}|e_H)\mathrm{d}\tilde{\tilde{y}} \\ &- \eta\lambda(1-\theta_{\mathbb{I}})\int_{\underline{y}}^{\overline{y}} v\left(u(w(\tilde{\tilde{y}})) - (u(w(y))\right)f(y|e_H)f(\tilde{\tilde{y}}|e)\mathrm{d}\tilde{\tilde{y}} - c\Bigg) \\ &+ \gamma\Bigg((u(w(y))\big(f(y|e_H) - f(y|e_L)\big) \\ &+ \theta_{\mathbb{I}}\eta\int_{\underline{y}}^{\overline{y}} v\left((u(w(y)) - u(w(\tilde{\tilde{y}})\right)\big(f(y|e_H) - f(y|e_L)\big)f(\tilde{\tilde{y}}|e)\mathrm{d}\tilde{\tilde{y}} \\ &- \lambda\eta(1-\theta_{\mathbb{I}})\int_{\underline{y}}^{\overline{y}} v\left(u(w(\tilde{\tilde{y}}) - (u(w(y))\right)\big(f(y|e_H) - f(y|e_L)\big)f(\tilde{\tilde{y}}|e)\mathrm{d}\tilde{\tilde{y}} - c\Bigg).\end{aligned} \qquad (A73)$$

Pointwise optimization with respect to $w(y)$ gives

$$-f(y|e_H)+\mu\Bigg(u'(w(y))f(y|e_H)+\eta\theta_{\mathbb{I}}\int_{\underline{y}}^{\overline{y}}v'(u(w(y)-u(w(\tilde{\tilde{y}}))u'\,(w(y)f(y|e_H)f(\tilde{\tilde{y}}|e)\mathrm{d}\tilde{\tilde{y}}$$

$$+\eta\lambda(1-\theta_{\mathbb{I}})\int_{\underline{y}}^{\overline{y}}v'\left(u\left(w(\tilde{\tilde{y}})\right)-(u(w(y))\right)u'\,(w(y)f(y|e_H)f(\tilde{\tilde{y}}|e)\mathrm{d}\tilde{\tilde{y}}\Bigg)$$

$$+\gamma\Bigg((u'(w(y))(f(y|e_H)-f(y|e_L))$$

$$+\theta_{\mathbb{I}}\eta\int_{\underline{y}}^{\overline{y}}v'\left((u(w(y))-u(w(\tilde{\tilde{y}})\right)u'(w(y)(f(y|e_H)-f(y|e_L))f(\tilde{\tilde{y}}|e)\mathrm{d}\tilde{\tilde{y}}$$

$$+\lambda\eta(1-\theta_{\mathbb{I}})\int_{\underline{y}}^{\overline{y}}v'\left(u(w(\tilde{\tilde{y}})-(u(w(y))\right)(f(y|e_H)-f(y|e_L))u'(w(y))f(\tilde{\tilde{y}}|e)\mathrm{d}\tilde{\tilde{y}}\Bigg)=0. \tag{A74}$$

Denoting by $w_s^F(y)$ the transfer satisfying $(A74)$, the following expressions are obtained after algebraic manipulations:

$$\frac{1}{u'(w_s^F(y))+\eta\int_{\underline{y}}^{\overline{y}}v'(u(w_s^F(y)-u(w_s^F(\tilde{\tilde{y}}))u'\,(w_s^F(y))f(\tilde{\tilde{y}}|e)\mathrm{d}\tilde{\tilde{y}}}=\mu+\gamma\left(1-\frac{f(y|e_L)}{f(y|e_H)}\right), \tag{A75}$$

if $\theta_{\mathbb{I}}=1$, and

$$\frac{1}{u'(w_s^F(y))+\eta\lambda\int_{\underline{y}}^{\overline{y}}v'\left(u(w_s^F(\tilde{\tilde{y}})-(u(w_s^F(y))\right)u'\,(w_s^F(y))f(\tilde{\tilde{y}}|e)\mathrm{d}\tilde{\tilde{y}}}=\mu+\gamma\left(1-\frac{f(y|e_L)}{f(y|e_H)}\right), \tag{A76}$$

if $\theta_{\mathbb{I}}=0$. Due to $u''<0$, $v''<0$ and $\frac{d}{dy}\left(\frac{f(y|e_L)}{f(y|e_H)}\right)\leq 0$ (Assumption 2), the derivative of $(A75)$ with respect to y exhibits $\frac{dw_s^F(y)}{dy}\geq 0$. Similarly, the derivative of $(A76)$ with respect to y exhibits $\frac{dw_s^F(y)}{dy}\geq 0$ if $U\left(e_H,w(\tilde{y}),w(\tilde{\tilde{y}})\right)$ is concave for any realization $\tilde{y}\in[\underline{y},\overline{y}]$ but becomes decreasing, $\frac{dw_s^F(y)}{dy}\leq 0$, if $U\left(e_H,w(\tilde{y}),w(\tilde{\tilde{y}})\right)$ is S-shaped. As mentioned in Theorem 1, the latter is an undesirable property.

ii) Solution when $U(e_H,w(y),r)$ is S-shaped

Let $U\left(e_H,w(\tilde{y}),w(\tilde{\tilde{y}})\right)$ be S-shaped for each $\tilde{y}\in[\underline{y},\overline{y}]$. If $w_s^F(y)$ satisfying Eq. $(A76)$ exhibits $0<w_s^F(y)<\ w_s^F(\tilde{\tilde{y}})$, the principal is better off offering $L_s:=(p\colon w_s^F(\tilde{\tilde{y}}),1-p\colon 0)$ for given $\tilde{\tilde{y}}$ and $p\in[0,1]$. Since

$$\mathbb{E}_y\left(u\left(w_s^F(y)\right)\right) - \eta\lambda \int_{\underline{y}}^{\tilde{y}} \int_{\underline{y}}^{\overline{y}} v\left(u\left(w_s^F(\tilde{\tilde{y}})\right) - u\left(w_s^F(y)\right)\right) f(\tilde{\tilde{y}}|e)\, \mathrm{d}\tilde{\tilde{y}} f(y|e) \mathrm{d}y$$
$$+ \eta \int_{\tilde{y}}^{\overline{y}} \int_{\underline{y}}^{\overline{y}} v\left(u\left(w_s^F(y)\right) - u\left(w_s^F(\tilde{\tilde{y}})\right)\right) f(\tilde{\tilde{y}}|e)\, \mathrm{d}\tilde{\tilde{y}} f(y|e) \mathrm{d}y,$$

increases in $w_s^F(y)$, there must exist a $p \in [0,1]$ such that:

$$\mathbb{E}_y\left(u\left(w_s^F(y)\right)\right) - \eta\lambda \int_{\underline{y}}^{\tilde{y}} \int_{\underline{y}}^{\overline{y}} v\left(u\left(w_s^F(\tilde{\tilde{y}})\right) - u\left(w_s^F(y)\right)\right) f(\tilde{\tilde{y}}|e)\, \mathrm{d}\tilde{\tilde{y}} f(y|e) \mathrm{d}y$$
$$+ \eta \int_{\tilde{y}}^{\overline{y}} \int_{\underline{y}}^{\overline{y}} v\left(u\left(w_s^F(y)\right) - u\left(w_s^F(\tilde{\tilde{y}})\right)\right) f(\tilde{\tilde{y}}|e)\, \mathrm{d}\tilde{\tilde{y}} f(y|e) \mathrm{d}y$$
$$= p\, \mathbb{E}_y\left(u\left(w_s^F(y)\right)\right) - p(1-p)\eta\lambda\left(\int_{\underline{y}}^{\overline{y}} v\left(-u\left(w_s^F(y)\right)\right) f(y|e) \mathrm{d}y\right)$$
$$+ p(1-p)\eta\left(\int_{\underline{y}}^{\overline{y}} v\left(u\left(w_s^F(y)\right)\right) f(y|e) \mathrm{d}y\right). \qquad (A77)$$

Hence, replacing $w_s^F(y)$ from Eq. $(A76)$ by L_s leaves the agent's participation incentive compatibility constraints unchanged. Since $\mathbb{E}_y\left(u\left(w_s^F(y)\right)\right) - p\, \mathbb{E}_y\left(u\left(w_s^F(y)\right)\right) \geq 0$, Eq. $(A77)$,the concavity of v, and loss aversion, $\lambda > 1$, imply:

$$-\eta \int_{\tilde{y}}^{\overline{y}} \int_{\underline{y}}^{\overline{y}} v\left(u\left(w_s^F(y)\right) - u\left(w_s^F(\tilde{\tilde{y}})\right)\right) f(\tilde{\tilde{y}}|e)\, \mathrm{d}\tilde{\tilde{y}} f(y|e) \mathrm{d}y$$
$$+ \eta\lambda \int_{\underline{y}}^{\tilde{y}} \int_{\underline{y}}^{\overline{y}} v\left(u\left(w_s^F(\tilde{\tilde{y}})\right) - u\left(w_s^F(y)\right)\right) f(\tilde{\tilde{y}}|e)\, \mathrm{d}\tilde{\tilde{y}} f(y|e) \mathrm{d}y$$
$$> -p\eta\left(\left(\int_{\underline{y}}^{\overline{y}} v\left((1-p)u\left(w_s^F(y)\right)\right) f(y|e) \mathrm{d}y\right)\right.$$
$$\left. + \lambda\left(\int_{\underline{y}}^{\overline{y}} v\left((1-p)u\left(w_s^F(y)\right)\right) f(y|e) \mathrm{d}y\right)\right)$$
$$(A78)$$

Some rearranging and using the fact that $\tilde{\tilde{y}} \in [\underline{y}, \overline{y}]$, gives

$$\lambda \int_{\underline{y}}^{\tilde{y}} \int_{\underline{y}}^{\overline{y}} v\left(u\left(w_s^F(\tilde{\tilde{y}})\right) - u\left(w_s^F(y)\right)\right) - pv\left((1-p)\, u\left(w_s^F(\tilde{\tilde{y}})\right)\right) f(\tilde{\tilde{y}}|e)\, \mathrm{d}\tilde{\tilde{y}} f(y|e) \mathrm{dy} >$$
$$\int_{\tilde{y}}^{\overline{y}} \int_{\underline{y}}^{\overline{y}} v\left(u\left(w_s^F(y)\right) - u\left(w_s^F(\tilde{\tilde{y}})\right)\right) - pv\left((1-p)u\left(w_s^F(\tilde{\tilde{y}})\right)\right) f(\tilde{\tilde{y}}|e)\, \mathrm{d}\tilde{\tilde{y}} f(y|e) \mathrm{dy}.$$
$$(A79)$$

The inequality in Eq. $(A79)$ is satisfied by

$$u(w_s^F(y)) > pu(w_s^F(y)). \qquad (A80)$$

Thus, L_s is more cost-effective for the principal than the solution implied by $(A76)$.

Next, I investigate the marginal incentives from offering L_s. Denote by $\overline{L_s}$ its expected value and substitute it in the agent's expected utility to obtain:

$$U\left(e_H, L, w_s^F(\tilde{\tilde{y}})\right) = \left(\frac{\overline{L_s}}{w_s^F(\tilde{\tilde{y}})}\right) u\left(w_s^F(\tilde{\tilde{y}})\right) + \frac{\overline{L_s}}{w_s^F(\tilde{\tilde{y}})}\left(1 - \frac{\overline{L_s}}{w_s^F(\tilde{\tilde{y}})}\right)\eta\left(v\left(u\left(w_s^F(\tilde{\tilde{y}})\right)\right) - \lambda v\left(-u\left(w_s^F(\tilde{\tilde{y}})\right)\right)\right). \qquad (A81)$$

The above expression is not linear in $\overline{L_s}$. Hence, changes in $\overline{L_s}$, via changes in p affect the agent's marginal utility. So unlike Theorem 1, there is an interior probability $p_s^*(y)$ that maximizes the agent's utility. The first-order condition of the utility in Eq. $(A81)$ with respect to p is:

$$u\left(w_s^F(\tilde{\tilde{y}})\right) + (1-2p)\eta\left(v\left(u\left(w_s^F(\tilde{\tilde{y}})\right)\right) - \lambda v\left(-u\left(w_s^F(\tilde{\tilde{y}})\right)\right)\right) = 0. \qquad (A82)$$

The second derivative of Eq. $(A80)$ with respect to p is $-2\eta\left(v\left(u\left(w_s^F(\tilde{\tilde{y}})\right)\right) - \lambda v\left(-u(w_s^F(\tilde{\tilde{y}}))\right)\right) < 0$. Therefore, the optimal probability is interior and has closed-form solution $p_s^*(\tilde{\tilde{y}}) = \frac{1}{2} + \frac{u\left(w_s^F(\tilde{\tilde{y}})\right)}{2\eta\left(v\left(u\left(w_s^F(\tilde{\tilde{y}})\right)\right) - \lambda v\left(-u\left(w_s^F(\tilde{\tilde{y}})\right)\right)\right)}$. Hence, the principal implements the lottery $L_s^* = (p_s^*(\tilde{\tilde{y}}): w_s^F(\tilde{\tilde{y}}), 1 - p_s^*(\tilde{\tilde{y}}): 0)$ if the contract $w_s^F(y)$ satisfying Eq. $(A76)$ exhibits $0 < w_s^F(y) < w_s^F(\tilde{\tilde{y}})$.

It is next shown that L_s^* is implemented for all $y < \overline{y}$. I proceed by contradiction. If there was an interior threshold output, $\hat{y}_{ds}$below which L_s^* is paid, and above which $w_s^F(y)$ from Eq. $(A75)$ is paid, that output level ought to satisfy:

$$\frac{1}{\frac{u(w_s^F(\hat{y}_{ds}))}{w_s^F(\hat{y}_{ds})} + \frac{1}{w_s^F(\hat{y}_{ds})}\left(1 - \frac{2\overline{L}}{w_s^F(\hat{y}_{ds})}\right)\eta\left(v\left(u(w_s^F(\hat{y}_{ds}))\right) - \lambda v\left(-u(w_s^F(\hat{y}_{ds}))\right)\right)} = \mu + \gamma\left(1 - \frac{f(\hat{y}_{ds}|e_L)}{f(\hat{y}_{ds}|e_H)}\right) \qquad (A83)$$

Since p_s^*satisfies Eq. $(A82)$, the numerator of the left-hand side of Eq. $(A83)$ is zero. Due to $\frac{d}{dy}\left(\frac{f(y|e_L)}{f(y|e_H)}\right) \leq 0$ (Assumption 2), the solution to Eq. $(A82)$ is at the boundary of the output space, $\hat{y}_{ds} = \overline{y}$.

iii) Properties of the optimal contract.

Let $U\left(e_H, w(\tilde{y}), w_s^F(\tilde{\tilde{y}})\right)$ be S-shaped for any $\tilde{y} \in [\underline{y}, \overline{y}]$. Denote by $b_s > 0$ the payment of $w_s^F(y)$ satisfying Eq. $(A75)$ evaluated at $y = \overline{y}$. Eq. $(A83)$ shows that is second-best optimal to offer $L_s^* = (p_s^*(\overline{y})\colon b_s, 1 - p_s^*(\overline{y})\colon 0)$ if $y < \overline{y}$ and b_s at $y = \overline{y}$. This proves the first part of the Proposition.

Let $U\left(e_H, w(\tilde{y}), w(\tilde{\tilde{y}})\right)$ be concave for any $\tilde{y} \in [\underline{y}, \overline{y}]$. The optimal contract, $w_s^*(y)$, consists of two components: $w_s^F(y)$ satisfying Eq. $(A75)$, which implies $w_s^F(y) \geq w_s^F(\tilde{\tilde{y}})$, and $w_s^F(y)$ satisfying Eq. $(A76)$, which implies $w_s^F(y) < w_s^F(\tilde{\tilde{y}})$. The optimal contract combines these two first-order conditions. To see how, notice that a contract only paying $w_s^F(y)$ satisfying Eq. $(A76)$ might lead to excess gains. The principal could deviate from this solution by implement higher outcomes that generate losses when incorporated as reference point. Moreover, a contract only paying Eq. $(A76)$ leads to excess losses and would be rejected. Therefore, the optimal contract consists of both components.

The transition from $w_s^F(y)$ satisfying Eq. $(A75)$ to $w_s^F(y)$ satisfying Eq. $(A76)$ is given by $\hat{y}_{ds} \in [\underline{y}, \overline{y}]$ satisfying:

$$\int_{\underline{y}}^{\hat{y}_{ds}} u(w_s^F(y))f(y|e)dy + \int_{\hat{y}_{ds}}^{\overline{y}} u(w_s^F(y))f(y|e)dy + \eta \int_{\hat{y}_{ds}}^{\overline{y}} \int_{\underline{y}}^{\overline{y}} v\left(u(w_s^F(y)) - u\left(w_s^F(\tilde{\tilde{y}})\right)\right) f(\tilde{\tilde{y}}|e)f(y|e)d\tilde{\tilde{y}}dy$$
$$- \eta\lambda \int_{\underline{y}}^{\hat{y}_{ds}} \int_{\underline{y}}^{\overline{y}} v\left(u\left(w_s^F(\tilde{\tilde{y}})\right) - u(w_s^F(y))\right) f(y|e)\, f(\tilde{\tilde{y}}|e)\, d\tilde{\tilde{y}}dy - c(e) = \overline{U} \qquad (A84)$$

The existence of $\hat{y}_{ds}$ is guaranteed inasmuch as the left-hand side of Eq. $(A84)$ can be negative when $\hat{y}_{ds} = \overline{y}$, it increases as $\hat{y}_{ds}$ increases and it becomes positive when $\hat{y}_{ds} = \underline{y}$.

Hence, a candidate for optimal contract is given by

$$w_s^*(y) = \begin{cases} w_s^F(y) & \text{satisfying } (A75) \quad \text{if } y \geq \hat{y}_{ds}, \\ w_s^F(y) & \text{satisfying } (A76) \quad \text{if } y < \hat{y}_{ds}. \end{cases} \qquad (A85)$$

That solution exhibits a discrete jump at $y = \hat{y}_{ds}$ since $\lambda > 1$ appears in the denominator of the right-hand side of $(A76)$ but this coefficient does not enter in $(A75)$. This proves the second part of the theorem.

iv) Optimal first-best contract

Consider now $\gamma = 0$ on top of $\phi = 1$. Denote by $w_f^F(y)$ the candidate solution from the first-order approach under these restrictions. Eq. $(A75)$ collapses to

$$\frac{1}{u'(w(y)) + \eta \int_{\underline{y}}^{\overline{y}} v'(u(w(y) - u(w(\tilde{\tilde{y}}))u'\,(w(y)f(\tilde{\tilde{y}}|e)\mathrm{d}\tilde{\tilde{y}}} = \mu, \qquad (A86)$$

if $\theta_{\mathrm{II}} = 1$, and Eq. $(A76)$ collapses to

$$\frac{1}{u'\big(w(y)\big) + \eta\lambda \int_{\underline{y}}^{\overline{y}} v'\Big(u(w(\tilde{\tilde{y}}) - (u(w(y))\Big) u'\,(w(y) f\big(\tilde{\tilde{y}}\big|e\big) \mathrm{d}\tilde{\tilde{y}}} = \mu, \qquad (A87)$$

if $\theta_{\mathbb{I}} = 0$. Eqs. $(A86)$ and $(A87)$ show that $\frac{dw_f^F(y)}{dy} = 0$ under $\gamma = 0$. Hence, $w_f^F(y)$ is performance insensitive.

As in the derivation of the second-best contract, it can be shown that if $U\big(e_H, w(\tilde{y}), w(\tilde{\tilde{y}})\big)$ is S-shaped for any $\tilde{y}$, the principal is better off paying a lottery $L_f = (p: w_f^F\big(\tilde{\tilde{y}}\big), 1-p: 0)$ instead of $w_f^F(y)$ satisfying Eq. $(A76)$. That lottery, L, can be offered to the agent with a $p \in (0,1)$ that satisfies:

$$\begin{aligned}
\mathbb{E}_y\left(u\left(w_f^F(y)\right)\right) &- \eta\lambda \int_{\underline{y}}^{\tilde{\tilde{y}}} \int_{\underline{y}}^{\overline{y}} v\left(u\left(w_f^F\big(\tilde{\tilde{y}}\big)\right) - u\left(w_f^F(y)\right)\right) f\big(\tilde{\tilde{y}}\big|e\big)\, \mathrm{d}\tilde{\tilde{y}} f(y|e) \mathrm{d}y \\
&+ \eta \int_{\tilde{\tilde{y}}}^{\overline{y}} \int_{\underline{y}}^{\overline{y}} v\left(u\left(w_f^F(y)\right) - u\left(w_f^F\big(\tilde{\tilde{y}}\big)\right)\right) f(\tilde{y}|e)\, \mathrm{d}\tilde{\tilde{y}} f(y|e) \mathrm{d}y \\
&= p\, \mathbb{E}_y\left(u\left(w_f^F\big(\tilde{\tilde{y}}\big)\right)\right) - p(1-p)\eta\lambda\left(\int_{\underline{y}}^{\overline{y}} v\left(-u\left(w_f^F\big(\tilde{\tilde{y}}\big)\right)\right) \mathrm{d}y f(y|e) \mathrm{d}y\right) \\
&+ p(1-p)\eta\left(\int_{\underline{y}}^{\overline{y}} v\left(u\left(w_f^F\big(\tilde{\tilde{y}}\big)\right)\right) f(y|e) \mathrm{d}y\right). \qquad (A88)
\end{aligned}$$

Therefore, replacing $w_f^F(y)$ from Eq. $(A76)$ by L_f leaves the agent's participation constraint unchanged. Eq. $(A88)$ and the convexity of $U\left(e_H, w_f^F(\tilde{y}), w_f^F\big(\tilde{\tilde{y}}\big)\right)$ for $\tilde{\tilde{y}}$ such that $w_f^F\big(\tilde{\tilde{y}}\big) < w_f^F(y)$, imply:

$$\begin{aligned}
\mathbb{E}_y\left(u\left(w_f^F(y)\right)\right) &- p\, \mathbb{E}_y\left(u\left(w_f^F\big(\tilde{\tilde{y}}\big)\right)\right) \\
&\geq \eta\lambda \int_{\underline{y}}^{\tilde{\tilde{y}}} \int_{\underline{y}}^{\overline{y}} v\left(u\left(w_f^F\big(\tilde{\tilde{y}}\big)\right) - u\left(w_f^F(y)\right)\right) f\big(\tilde{\tilde{y}}\big|e\big)\, \mathrm{d}\tilde{\tilde{y}} f(y|e) \mathrm{d}yy \\
&- \eta \int_{\tilde{\tilde{y}}}^{\overline{y}} \int_{\underline{y}}^{\overline{y}} v\left(u\left(w_f^F(y)\right) - u\left(w_f^F\big(\tilde{\tilde{y}}\big)\right)\right) f(\tilde{y}|e)\, \mathrm{d}\tilde{\tilde{y}} f(y|e) \mathrm{d}y \\
&- p\eta\lambda\left(\int_{\underline{y}}^{\overline{y}} v\left(-(1-p) u\left(w_f^F\big(\tilde{\tilde{y}}\big)\right)\right) \mathrm{d}y f\big(\tilde{\tilde{y}}\big|e\big) \mathrm{d}\tilde{\tilde{y}}\right) \\
&+ p\eta\left(\int_{\underline{y}}^{\overline{y}} v\left((1-p) u\left(w_f^F\big(\tilde{\tilde{y}}\big)\right)\right) f\big(\tilde{\tilde{y}}\big|e\big) \mathrm{d}\tilde{\tilde{y}}\right). \qquad (A89)
\end{aligned}$$

Since $\mathbb{E}_y\left(u\left(w_f^F(y)\right)\right) - p\,\mathbb{E}_y\left(u\left(w_f^F(\tilde{\tilde{y}})\right)\right) > 0$ and $v' > 0$, the last inequality implies $w_f^F(y) > pw_f^F(\tilde{\tilde{y}})$. Hence, L_f is more cost-effective for the principal than the candidate solution given by Eq. $(A86)$.

To investigate the incentives of L_f, denote by $\overline{L}_f$ its expected value and substitute it in the agent's expected utility to obtain:

$$U\left(e, L, w_f^F(\tilde{\tilde{y}})\right) = \left(\frac{\overline{L}_f}{w_f^F(\tilde{\tilde{y}})}\right) u\left(w_s^F(\tilde{\tilde{y}})\right) + \frac{\overline{L}_f}{w_f^F(\tilde{\tilde{y}})}\left(1 - \frac{\overline{L}_f}{w_f^F(\tilde{\tilde{y}})}\right)\eta\left(v\left(u\left(w_f^F(\tilde{\tilde{y}})\right)\right) - \lambda v\left(-u\left(w_f^F(\tilde{\tilde{y}})\right)\right)\right), \tag{A90}$$

an expression that is not linear in $\overline{L}_f$. Hence, changes in $\overline{L}_f$ affect the agent's marginal utility. As it was the case with the second-best contract, the probability that maximizes the agent's utility can be found via the first order condition of Eq. $(A90)$ with respect to p. The resulting probability is $p_f^*(\tilde{\tilde{y}}) = \frac{1}{2} + \frac{u\left(w_f^F(\tilde{\tilde{y}})\right)}{2\eta\left(v\left(u\left(w_f^F(\tilde{\tilde{y}})\right)\right) - \lambda v\left(-u\left(w_f^F(\tilde{\tilde{y}})\right)\right)\right)}$. Hence, lottery $L_f^* := (p_f^*(\tilde{\tilde{y}}): w_s^F(\tilde{\tilde{y}}), 1 - p_f^*(\tilde{\tilde{y}}): 0)$ is proposed by the principal.

When $U\left(e_H, w(\tilde{y}), w(\tilde{\tilde{y}})\right)$ is S-shaped for each $\tilde{y}$, the first-best contract, $w_f^*(y)$, consists of L_f^*. Suppose instead that $w_f^F(y)$ satisfying Eq. $(A88)$ is given everywhere. Since $w_f^F(y) > pw_f^F(\tilde{\tilde{y}})$, the principal can profitably deviate from that solution by paying L_f at the lower end of the output space. Because L_f^* is evaluated as a sizeable loss when $w_f^F(y)$ satisfying $(A86)$ is taken as reference point, the principal increases the segment for which L_f^* is the solution until the boundary $\hat{y}_{df} = \overline{y}$ is reached. This strategy is cost-effective. Also, proceeding in such way would not fully locate the agent in losses; gains are experienced when the outcome of the lottery $w = 0$ is adopted as reference point.

Denote by $b_f > 0$ the pay level $w_f^F(y)$ satisfying Eq. $(A75)$ evaluated at $y = \overline{y}$. It is second-best optimal to offer $L_s^* = \left(p_f^*(\overline{y}): b_f, 1 - p_{s=f}^*(\overline{y}): 0\right)$ if $y < \overline{y}$ and b_s paid at $y = \overline{y}$ proving the first part of the Proposition.

When $U\left(e_H, w(\tilde{y}), w(\tilde{\tilde{y}})\right)$ is concave for any $\tilde{y} \in [\underline{y}, \overline{y}]$, the optimal contract also consists of two components: $w_f^F(y)$ satisfying Eq. $(A88)$, which implies $w_f^F(y) \geq w_s^F(\tilde{\tilde{y}})$, and $w_f^F(y)$ satisfying Eq. $(A89)$, which implies $w_f^F(y) < w_s^F(\tilde{\tilde{y}})$. Since the agent is loss averse, $\lambda > 1$, $w_f^F(y)$ satisfying $(A89)$ cannot be a solution on its own as it induces considerable disutility, leading the agent to reject the contract. Moreover, a combination of these two components exposes the agent to the risk of experiencing losses. Since $U\left(e_H, w(\tilde{y}), w(\tilde{\tilde{y}})\right)$ is concave, such a combination does not provide full insurance. Hence, it must be $w_f^F(y)$ satisfying Eq. $(A88)$ is given. This proves the fifth part of the Proposition. ■

Corollary 7.

Let $x := u(w(y)) - u(w(g))$. Replace $u'' < 0$ from Assumption 4 for $u''(x) < 0$ if $x \geq 0$ and $u''(x) > 0$ if $x < 0$. Under such modification, receiving lottery $L_k := (1-p: w_s^*(y), p: 0)$ generates the following utility:

$$KR(L) = (1-p)u\big(w_s^*(y)\big) + p(1-p)\eta\left(v\big(u\big(w_s^*(y)\big)\big) + \lambda v\left(-u\big(w_s^*(y)\big)\right)\right). \qquad (A91)$$

Since the second derivative of Eq. $(A91)$ is $-2\eta\left(v(u(w_s^*(y))) + \lambda v\left(u\big(-w_s^*(y)\big)\right)\right)$, a negative expression, the interior probability that maximizes utility is given by the closed-form solution $p_k^*(y) = \frac{1}{2} - \frac{u(w_s^*(y))}{2\eta(\lambda v(-u(w_s^*(y))) + v(u(w_s^*(y))))}$.

Proposition 2 (i) shows that $\hat{y}_{ds} = \overline{y}$. Hence, it is optimal to offer $L_k := (1 - p_k^*(y): b_s, p_k^*(y): 0)$ for any $y < \overline{y}$. Consider $\eta = 1$ and $v' = 1$. The optimal probability becomes $p_k^* = \frac{1}{2} + \frac{1}{2(\lambda-1)}$, where it is clear that $p^* \in (0,1)$ if $\lambda > 2$ but $p_k^* = 1$ if $\lambda \leq 2$. Consequently, $w_s^*(y) = \begin{cases} 0 \; if \; y < \overline{y} \\ b_s \; if \; y = \overline{y} \end{cases}$ with $b\_s > 0$ is the optimal contract if $\lambda \leq 2$. Instead, the stochastic contract L_k with $p_k^*(y) \in (0,1)$ is given if $\lambda > 2$. ■

Appendix C

C.1 Other salience-based reference points.

Corollary C.1 *Under A1- A4 and the max-min rule, the agent's reference point and the second-best contract are identical to those presented in Corollary 5.*

Proof. The set of candidates for min-max reference point are $r = \min\{\max\{\overline{U}\}, \max\{w_s^*(y)\}\}$. Intuitively, rejecting the contract yields welfare $\overline{U}$ and the maximum that the contract pays, regardless of effort level, is $\max\{w_s^*(y)\}$.

Since the participation constraint binds at the optimum, then $\mathbb{E}\big(U(e, w_s^*(y), r)\big) = \overline{U}$. Also, because the incentive compatibility constraint binds, it must be that $\mathbb{E}\big(U(e, w_s^*(y), r)\big) < \mathbb{E}\big(U(r, \max\{w_s^*(y)\}, r)\big)$. Otherwise, the second-best contract would not implement rewards for high performance and thus would not incentivize high effort. Hence, $r = w_s^*(y)$

Since the agent's preferences are characterized by prospect theory and his reference point is $r = w_f^*(y)$, the optimal incentive scheme is identical to that presented in Corollary 5. ■

Corollary C.2 *Under A1-A4 and the $w(y)$ at max P rule, the agent's reference point is $r = w_s^*(y_p)$, where y_p satisfies $f\big(y_p|e_H\big) = \max\{f(y|e_H)\}$, and the second-best contract, $w_s^*(y)$, pays the lowest possible in $y < \hat{y}_p$, exhibits a bonus at $y = y_p$, and increases in performance in $y > \hat{y}_p$.*

Proof. Let $y_p \in \big[\underline{y}, \overline{y}\big]$ be a performance level satisfying $f\big(y_p|e\big) = \max\{f(y|e)\}$. If $f\big(y_p|e\big)$ is multimodal, define y_p as the smallest output level satisfying $f\big(y_p|e\big) = \max\{f(y|e)\}$. That $\big[\underline{y}, \overline{y}\big] \subseteq \mathbb{R}^+$, implies that the point y_p attains the highest probability as compared to any $y \in \big[\underline{y}, \overline{y}\big] \setminus \{y_p\}$.

Since the agent's preferences are characterized by prospect theory, a contract with the same shape as that described in Corollary 4 remains to be optimal. Denote that contract by $w_s^*(y)$. The $w(y)$ *at max P* reference point rule entails that $r = w_s^*\big(y_p\big)$. Hence, the output level after which the bonus is awarded might be different than that given in Corollary 4. That point is defined next. Let $\hat{y}_p \in [\underline{y}, \overline{y}]$ satisfy:

$$\frac{1}{\frac{\lambda v\left(w_s^*(\hat{y}_p)\right)}{w_s^*(\hat{y}_p)}} = \nu + \gamma\left(1 - \frac{f\big(\hat{y}_p|e_L\big)}{f\big(\hat{y}_p|e_H\big)}\right). \qquad (C1)$$

According to Corollary 4, the optimal contract should pay $w_s^*(y)$ satisfying the following first-order condition

$$\frac{1}{v'\big(w_s^*(y) - w_s^*(y_p)\big)} = \nu + \gamma\left(1 - \frac{f(y|e_H)}{f(y|e_L)}\right), \qquad (C2)$$

if $y > \hat{y}_p$, and $w_s^* = 0$ if $y < \hat{y}_p$.

Finally, I demonstrate that $\hat{y}_p = y_p$. I proceed by contradiction. Suppose instead that $\hat{y}_p < y_p$. In that case, the principal is overinsuring the agent from risk in $y \in [\hat{y}_p, y_p]$ by offering $w_s^*(y)$ satisfying $(C2)$ in a segment where he is risk seeking due to diminishing sensitivity. The principal could increase profits by exposing the agent to large amounts of risk by setting $w_s^*(y) = 0$ for all $y < y_p$, including $y \in [\hat{y}_p, y_p]$, and the agent would accept such contract.

Next, suppose that $\hat{y}_p > y_p$. In that case the agent is being exposed to large amounts of risk in $y \in [y_p, \hat{y}_p]$, a segment where he is risk averse (Assumption 4). This incentivizes the agent to reject the contract. The principal anticipates this and provides insurance offering the payment scheme $w_s^*(y)$ satisfying $(C2)$ for $y \geq y_p$. Hence, it must be that $\hat{y}_p = y_p$. ■

C.2 Gul's (1991) model

The disappointment model of Gul (1991) differs from those of Bell (1985) and Loomes and Sugden (1986) in that the agent's reference point is his certainty equivalent. Importantly, the certainty equivalent includes the agent's psychological utility component.

More formally, consider the general specification of reference dependence given in (2). As with the other previous disappointment models allow for expected consumption utility by letting $\phi = 1$. Under these restrictions, the agent's certainty equivalent is the level $CE \in \mathbb{R}$ that satisfies $U(e_H, w(y), CE) = u(CE)$ for a given incentive scheme $w(y)$.

The first-best and second-best optimal contracts under Gul's (1991) preferences are presented in the following next corollary.

Corollary C.3 *Under assumptions A1-A4,* $\phi = 1$, *and* $r = CE$, *there exist unique output levels* $y_{cf}, y_{cs} \in [\underline{y}, \overline{y}]$ *such that:*

i) *The first-best contract,* $w_f^*(y)$, *is equal to that given in Corollary 6 with* y_{mf} *replaced by* y_{cf}.

ii) *The second-best contract,* $w_s^*(y)$, *is equal to that given in Corollary 6 with* y_{mf} *replaced by* y_{cs}.

Proof. Let $\phi = 1$, $u' = 1$, and $r = CE$. Under the assumed restrictions, $U(e_H, w(\tilde{y}), CE)$ is S-shaped since $-\frac{v''\left(u(r)-u(w(y))\right)}{v'\left(u(r)-u(w(y))\right)} > 0$ in the domain of losses, corroborating equation $(A3)$. According to Theorem 1, it is optimal to set $w_s^*(y) = 0$ for low output levels. Moreover, the first-order condition is necessary and sufficient only in the domain of gains.

The point at which the bonus of the second-best contract is awarded is defined by the following output level, which adapts Eq. $(A14)$ to account for the assumed restrictions,

$$\frac{1}{\frac{1+\lambda v(CE)}{CE}} = \mu + \gamma\left(1 - \frac{f(\hat{y}_{cs}|e_L)}{f(\hat{y}_{cs}|e_H)}\right). \qquad (C3)$$

The bonus is awarded when the unique output level $\hat{y}_{cs}$ that satisfies $(C3)$ is surpassed. Hence, the optimal contract is given by $w_s^*(y) = \begin{cases} 0 & if\ y < \hat{y}_{cs}, \\ CE + j'\left(\frac{1}{\eta}\left(\frac{1}{\mu+\gamma\left(1-\frac{f(y|e_L)}{f(y|e_H)}\right)} - 1\right)\right) & if\ y \geq \hat{y}_{cs}. \end{cases}$ This proves the first part of the corollary.

We turn to study the first-best contract. Hence, consider $\gamma = 0$ in addition to $\phi = 1$, $u' = 1$, and $r = CE$. Eq. $(A55)$ becomes $\frac{1}{\left(1+\eta v'\left(w_f^*(y)-CE\right)\right)} = \mu$, which after some manipulations yields $w_f^*(y) = CE + j'\left(\frac{1}{\eta}\left(\frac{1}{\mu} - 1\right)\right)$. According to Eq. $(A56)$, that transfer exhibits $\frac{dw_f^*(y)}{dy} = 0$.

Also, Theorem 1 shows that in the domain of losses it is also first-best optimal to offer $w_f^*(y) = 0$. The transition from $w_f^*(y) = 0$ to $w_f^*(y) = CE + j'\left(\frac{1}{\eta}\left(\frac{1}{\mu} - 1\right)\right)$ is given by the following equality, which adapts Eq. $(A23)$ to account for the considered restrictions,

$$\int_{\hat{y}_{cf}}^{\overline{y}} w_f^*(y) f(y|e_H)\,\mathrm{d}y + \eta \int_{\hat{y}_{cf}}^{\overline{y}} v\left(w_f^*(y) - CE\right) f(y|e_H)\,\mathrm{d}y - \lambda \int_{\underline{y}}^{\hat{y}_{cf}} v(CE) f(y|e_H)\mathrm{d}y - c = \overline{U}. \quad (C4)$$

Hence, the output $\hat{y}_{cf}$ that satisfies Eq. $(C4)$ provides that transition. Theorem 1 shows that this output level is unique and interior. Hence, the optimal contract is $w_f^*(y) = \begin{cases} 0 & if\ y < \hat{y}_{cf}, \\ CE + j'\left(\frac{1}{\eta}\left(\frac{1}{\mu} - 1\right)\right) & if\ y \geq \hat{y}_{cf}. \end{cases}$ This proves the third part of the corollary.

Next let $\phi = 1$, $v' = 1$, and $r = CE$. Under these restrictions $U(e_H, w(\tilde{y}), CE)$ is concave since $-\frac{u''(w(y))}{u'(w(y))} > 0$, a contradiction of the condition in Eq. $(C3)$. Hence, the solutions from the first-order conditions are necessary and sufficient to solve the maximization problem of the principal.

The first order conditions given by Eqs. $(A59)$and $(A60)$ provide the solution. The transition from $w_s^*(y)$ satisfying $(A59)$ to $w_s^F(y)$ satisfying $(A60)$ is given by the unique output level $\hat{y}_{cs} \in (\underline{y}, \overline{y})$ that satisfies:

$$\int_{\hat{y}_{cs}}^{\overline{y}} u(w_S^*(y)) f(y|e_H)\,\mathrm{d}y + \eta \int_{\hat{y}_{cs}}^{\overline{y}} \big(u(w_S^*(y)) - u(CE)\big) f(y|e_H)\,\mathrm{d}y + \int_{\underline{y}}^{\hat{y}_{cs}} u(w_S^*(y) f(y|e_H)\mathrm{d}y$$
$$- \lambda\eta \int_{\underline{y}}^{\hat{y}_{cs}} \Big(u(CE) - u\big(w_s^*(y)\big)\Big) f(y|e_H)\mathrm{d}y - c \;= \overline{U}. \qquad (C5)$$

The existence of $\hat{y}_{cs}$ is guaranteed by the fact that the solutions from Eqs. $(A59)$ and $(A60)$ make the participation constraint bind for gains and losses, respectively. As a result, the optimal incentive scheme is given by:

$$w_f^*(y) = \begin{cases} h'\left(\dfrac{(1+\eta)}{\mu + \gamma\left(1 - \dfrac{f(y|e_L)}{f(y|e_H)}\right)}\right) & if\ y \geq \hat{y}_{cs}\,, \\[2ex] h'\left(\dfrac{(1+\eta\lambda)}{\mu + \gamma\left(1 - \dfrac{f(y|e_L)}{f(y|e_H)}\right)}\right) & if\ y < \hat{y}_{cs}. \end{cases} \qquad (C6)$$

This proves the second part of the corollary.

To conclude, we analyze the first-best contract. Hence, consider $\gamma = 0$ in addition to $\phi = 1$, $v' = 1$, and $r = CE$. The first-order condition in $(A59)$ becomes $\frac{1}{u'\left(w_f^*(y)\right)(1+\eta)} = \mu$, which after algebraic manipulations gives $w_f^*(y) = h'\left(\frac{(1+\eta)}{\mu}\right)$. Eq. $(A8)$ shows that $\frac{dw_f^*(y)}{dy} = 0$ under the considered restrictions. Finally, Theorem 1 shows that $w_f^*(y) = h'\left(\frac{(1+\eta)}{\mu}\right)$ is first-best optimal when $U(e_H, w(y), r)$ is concave. ■

An agent with disappointment averse preferences and who adopts his certainty equivalent as reference point is insured and motivated with contracts that greatly resemble those described in Corollary 6. Therefore, contracts with a bonus enable the principal to exploit the agent's irrationalities of loss aversion and diminishing sensitivity in an optimal way.

However, the reference point rule specified by Gul (1991)'s mode generates potential differences in the location and magnitude of the bonus. Intuitively, a (globally) risk averse agent must exhibit $CE < \mathbb{E}(w(y))$. Hence, to guarantee that the contract is accepted, the principal protects this agent from risk by awarding the bonus at lower output levels as compared to the hypothetical case in which the agent was risk neutral, $CE = \mathbb{E}(w(y))$. Hence, $y_{cf} < y_{mf}$ and $y_{cs} < y_{ms}$. These more lenient threshold levels come at the cost of the magnitude of the bonus included in each contract, which becomes smaller as compared to the risk neutral case. In that way, the principal keeps the agent just indifferent between accepting or rejecting the contract. A similar intuition leads to the conclusion that for a globally risk seeking agent $y_{c1} > y_{m1}$, $y_{c2} > y_{m2}$, and both contracts include a larger bonus. A result that is consistent with the comparative static presented in Corollary 2.

C.3 Adapting the model to accommodate De Meza and Webb (2007)

The model can be adapted to allow for the results of De Meza and Webb (2007). The following adaptation of Assumption 4 is considered.

Assumption C1 (AC1). *The agent's value function V is the piece-wise function:*

$$V(w,r) = \begin{cases} 0 & \text{if } w(y) > r, \\ -v\left(u(r) - u\big(w(y)\big)\right) & \text{if } r \leq w(y). \end{cases}$$

Where $v: \mathbb{R}^+ \longrightarrow \mathbb{R}^+$, v is C^2, $v(0) = 0, v' > 0 \; \forall \, y \in \left[\underline{y}, \overline{y}\right], v'' \leq 0$ and v has a differentiable inverse $j := v^{-1}$.

There are three key differences between A4 and AB1. First, loss aversion in the usual sense, i.e. sign dependence, is abandoned. In other words, $\lambda = 1$. Second, outcome comparisons relative to the reference point are restricted to the domain of losses. A consequence of that assumption is that diminishing sensitivity applies only to losses. This effect is referred as loss aversion by De Meza and Webb (2007). However, this way of modeling loss aversion contradicts standard definitions when $v'' < 0$. (Tversky and Kahneman, 1992, Köbberling and Wakker, 2005). Third, the transition from gains to losses is not given at the reference point but once that value is surpassed, that is for $w(y) > r$.

Moreover, let $\phi = 1$ and $\eta = 1$. All in all, the decision-maker's preferences are given by

$$U(e, w(y), r) = \int_{\underline{y}}^{\overline{y}} u\big(w(y)\big) f(y|e) \mathrm{d}y - \int_{\underline{y}}^{\overline{y}} \left((1 - \theta_{\mathbb{I}}) v\left(u(r) - u\big(w(y)\big)\right)\right) f(y|e) \mathrm{d}y - c(e) \tag{C7}$$

Under preferences as in Eq. $(C7)$, the results of De Meza and Webb (2007) follow. When $v' = l > 0$, that is when diminishing sensitivity is assumed to be piece-wise linear, then the results of their Proposition 1 follow. Throughout, they interpret $l > 0$ as loss aversion.

In that case, $U(e, w(\tilde{y}), r)$ for any $\tilde{y} \in [\underline{y}, \overline{y}]$ is concave. The first order conditions from Eqs. $(A6)$ and $(A7)$, given in the Proof of Theorem 1, become

$$\frac{1}{u'(w_S^*(y))} = \mu + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right), \qquad (C8)$$

and

$$\frac{1}{u'(w_S^*(y))(1+l)} = \mu + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right), \qquad (C9)$$

respectively. These two equations are increasing in performance. Also, these equations are equal to Eq.(8) in their paper, crucial for their proof of Proposition 1.

When l is small enough, the principal can offer $w_S^*(y)$ satisfying $(C9)$ while still guaranteeing $U(e, w_S^*(y), r) = \overline{U}$. That is because expected consumption utility, $\phi = 1$, ensures that the participation constraint binds even though the contract locates the agent in the domain of losses. Since r is still part of the domain of losses, due to $w < r$, the optimal incentive scheme can be insensitive at high output levels, i.e. $w_S^*(y) = r$. This leads to Proposition 1 (ii) and Figure 1a in De Meza and Webb (2007).

For larger l, the agent needs to be translated to the domain of gains to guarantee $U(e, w_S^*(y), r) = \overline{U}$. In that case, expected consumption utility does not fully outweigh losses in the contract and at high ouput levels $w(y) > r$ must be ensured. Since $\lambda = 1$, there is no jump or kink when the transition from $w_S^*(y) = r$ to $w_S^*(y)$ satisfying Eq. $(C8)$ takes place. This case generates Proposition 1 (iv) and Figure 1c in De Meza and Webb (2007). For even higher l, the exposure of the agent in the domain of losses is reduced by paying $w_S^*(y) = r$ for low output levels and $w_S^*(y)$ satisfying Eq. $(C8)$ being paid at intermediate and high output levels. This covers their Proposition 1 (iii) and Figure 1b.

Finally, when $v'' < 0$, which De Meza and Webb (2007) denote as non-linear loss aversion, the results of their Proposition 2 follow. Let $U(e, w(\tilde{y}), r)$ be S-shaped. The proof of Theorem 1 shows that in the domain of losses either $w_S^*(y) = 0$ or $w_S^*(y) = r$ must be given. Hence, the optimal incentive scheme is given by $w_S^*(y) = 0$, $w_S^*(y) = r$, and $w_S^*(y)$ satisfying Eq. $(C8)$. As above, the magnitude of v determines the shape of the incentive scheme. When v is small enough, then a combination of $w_S^*(y) = 0$ and $w_S^*(y) = r$ is given to the agent. This case is given in Figure 2c in Meza and Webb (2007). When v is larger, the agent needs to be transitioned in the domain of gains for high output levels. Then, the optimal incentive scheme is a combination of $w_S^*(y) = 0$, $w_S^*(y) = r$, and $w_S^*(y)$ satisfying Eq. $(C8)$. Finally, a large enough v leads to an optimal incentive scheme that combines $w_S^*(y) = r$ and $w_S^*(y)$ satisfying Eq. $(C8)$. This case is depicted in Figure 2a.

Appendix D: Results Section 5

This section extends the results presented in Section 3 to further gain generalizability and highlight the significance of Theorem 1. Here, I show that the results in that theorem are valid beyond a set of assumptions made on the principal's preferences and knowledge.

In this section, I return to the assumption that the agent's reference point is exogenous, i.e. $r > 0$, and lay focus on the moral hazard case.

D.1 Principal with Reference Dependent Preferences

A discernible extension considers a setting in which the principal also exhibits reference-dependence. Formally, let the principal's preferences be characterized by

$$\Pi\big(S(y), r_p, w\big) = \begin{cases} S(y) - r_p - w(y) & if \;\; S(y) \geq r_p + w(y), \\ -\lambda_p \left(r_p + w(y) - S(y)\right) & if \; S(y) < r_p + w(y), \end{cases} \qquad (D1)$$

where $r_p \geq 0$ and $\lambda_p > 1$.

According to Eq. $(D1)$, the principal is loss averse. This irrationality applies to both her benefit and cost functions. An assumption consistent with the notion, also present in Assumption 4, that these biases apply to monetary outcomes.[1] Furthermore, she does not suffer from diminishing

[1] Another possible representation of reference dependence is

$$\Pi\big(S(y), r_p, w\big) = \begin{cases} P\big(S(y) - r_p\big) - w(y) & if \;\; S(y) \geq r_p\,, \\ -\lambda_p P\left(r_p - S(y)\right) - w(y) & if \; S(y) < r_p. \end{cases}$$

This representation is consistent with the approach taken throughout the paper to model reference dependence for the agent. Notice that this assumption together with the assumption that the agent's preference is given by Eq. (2) imply that the contracts in Theorem 1 remain optimal. The principal's loss aversion and diminishing sensitivity do not apply to her cost component, $w(y)$, so in that case the principal's problem is unchanged.

sensitivity. This simplifying assumption can be justified on the grounds of the principal being able to pool multiple risks and, as a result, not exhibiting utility curvature.

The solutions to the principal's program when she is loss averse are comparable to those presented in Theorem 1. The only difference appears for the case in which output attains intermediate levels.

Proposition D.1. *Let $\hat{y}_s \in (\underline{y}, \overline{y})$ be the unique output level from Theorem 1. Under A1-A4, and that the principal's preferences are given by* (7), *there exists a unique output level $\hat{y}_p \in \left[\underline{y}, \overline{y}\right]$ such that the second-best contract, $w_s^*(y)$:*

i) Is identical to the contract presented in Theorem 1 (i) and (ii) if $\hat{y}_p < \hat{y}_s$.

ii) Pays the minimum possible if $y < \hat{y}_s$, exhibits a bonus at $y = \hat{y}_s$, increases in performance in $y > \hat{y}_s$ but at a lower rate in the segment $y \in (\hat{y}_s, \hat{y}_p)$ if $\hat{y}_p \geq \hat{y}_s$ and $U(e, w_s^*(\tilde{y}), r)$ is S-shaped.

iii) Exhibits a bonus at $y = \hat{y}_s$, increases in performance in $y < \hat{y}_s$ and $y > \hat{y}_s$, but at a lower rate in the segment $y \in (\hat{y}_s, \hat{y}_p)$ if $\hat{y}_p \geq \hat{y}_s$ and $U(e, w_s^*(\tilde{y}), r)$ is concave.

Proof.

Denote by $\mu \geq 0$ and $\gamma \geq 0$ the Lagrangian multipliers of the agent's participation and incentive compatibility constraints. First, let $S(y) < r_p + w(y)$. In that case, the Lagrangian of the principal's maximization program writes as:

$$\mathcal{L} = \left(-\lambda_p \left(r_p + w(y) - S(y)\right)\right) f(y|e_H)$$
$$+\mu\left[\phi u(w(y) + \theta_{\mathbb{I}} \eta v(u(w(y)) - u(r)) f(y|e_H) - \lambda(1-\theta_{\mathbb{I}})\eta v(u(r) - u(w(y))) f(y|e_H) - c - \overline{U}\right] \quad (D2)$$
$$+\gamma\left[\phi u(w(y)(f(y|e_H) - f(y|e_L)) + \eta\theta_{\mathbb{I}} v\left(u(w(y)) - u(r)\right)\left(f(y|e_H) - f(y|e_L)\right)\right.$$
$$\left. - \lambda(1-\theta_{\mathbb{I}})\eta v(u(r) - u(w(y)))(f(y|e_H) - f(y|e_L)) - c\right].$$

Pointwise optimization with respect to $w(y)$ gives

$$-f(y|e_H)\lambda_p + u'(w(y)\mu\left[\phi + \eta\theta_{\mathbb{I}} v'\left(u(w(y)) - u(r)\right) f(y|e_H) + \eta\lambda(1-\theta_{\mathbb{I}}) v'\left(u(r) - u(w(y))\right) f(y|e_H)\right]$$
$$+ u'(w(y)\gamma\left[\phi + \eta\theta_{\mathbb{I}} v'\left(u(w(y)) - u(r)\right)\left(f(y|e_H) - f(y|e_L)\right)\right.$$
$$\left. + \eta\lambda(1-\theta_{\mathbb{I}}) v'\left(u(r) - u(w(y))\right)\left(f(y|e_H) - f(y|e_L)\right)\right] = 0 \quad (D3)$$

Denote by $w_s^F(y)$ the transfer satisfying $(D3)$. After algebraic manipulations, I find the following expressions

$$\frac{\lambda_p}{u'(w(y))\left(\phi+\eta\theta_{\mathbb{I}}v'\left(u\big(w(y)\big)-u(r)\right)\right)}=\mu+\gamma\left(1-\frac{f(y|e_L)}{f(y|e_H)}\right), \qquad (D4)$$

if $\theta_{\mathbb{I}}=1$, and

$$\frac{\lambda_p}{u'(w(y))\left(\phi+\lambda\eta\theta_{\mathbb{I}}v'\left(u(r)-u\big(w(y)\big)\right)\right)}=\mu+\gamma\left(1-\frac{f(y|e_L)}{f(y|e_H)}\right), \qquad (D5)$$

and $\theta_{\mathbb{I}}=0$. As in Theorem 1, if $U(e_H,w(\tilde{y}),r)$ is concave, the conditions given in $(D4)$ and $(D5)$ are necessary and sufficient to solve the maximization problem. Instead, if $U(e_H,w(\tilde{y}),r)$ for any $\tilde{y}\in[\underline{y},\overline{y}]$ is S-shaped, the principal is better of offering lottery $L=(p{:}\,r,1-p{:}\,0)$ with a probability $p\in(0,1)$ that satisfies:

$$\phi u(w_s^F(y))-\lambda v\left(u(r)-u\big(w_s^F(y)\big)\right)=\phi p u(r)-(1-p)\lambda v\big(u(r)\big). \qquad (D6)$$

Hence, paying L does not change participation and incentive compatibility constraint. Also, offering that lottery is more cost effective for the principal since from $(D6)$:

$$\phi\left(u\big(w_s^F(y)\big)-pu(r)\right)\leq\lambda v\left(u(r)-u\big(w_s^F(y)\big)\right)-\lambda v\big((1-p)u(r)\big),$$

implying $w_s^F(y)>pr$.

Denote by $\hat{y}_s$ the output level satisfying:

$$\frac{1}{\frac{\phi u(r)+\lambda\eta v\big(u(r)\big)}{r}}=\mu+\gamma\left(1-\frac{f(\hat{y}_s|e_L)}{f(\hat{y}_s|e_H)}\right). \qquad (D7)$$

That output level is unique since the left-hand side of Eq.$(D7)$ is constant in y and is positive, while the right-hand side of that equation increases in y over the domain $[0,\infty)$.

When $y<\hat{y}_s$, the scheme pays $w_s^F(y)=0$. That is because when offered L the agent's utility can be expressed as

$$U(e_H,L,r)=-\lambda\left(1-\frac{\overline{L}}{r}\right)u(r)-c, \qquad (D8)$$

where $\overline{L}$ is the expected value of L. Notice that Eq. (8) is linear in $\overline{L}$. Hence, changes in $\overline{L}$ do not affect the agent's marginal utility and the principal can afford to set $p=0$. Instead, if $\hat{y}_s<y$, the agent's payment can be set $p=1$, which brings him to the domain of gains. In that domain, the principal should be paid $w_s^F(y)$ satisfying $(D4)$. Therefore, the solution to the principal's problem is

$$w_{SB}(y) = \begin{cases} 0 & if\ y < \hat{y}_s, \\ w_s^F(y) \text{ from } (C3) & if\ y \geq \hat{y}_s. \end{cases} \quad (D9)$$

If $U(e_H, w(\tilde{y}), r)$ for any $\tilde{y} \in [\underline{y}, \overline{y}]$ is concave $w_s^*(y)$, consists of two components: $w_s^F(y)$ satisfying $(D4)$, which implies $w_s^F(y) \geq r$, and $w_s^F(y)$ satisfying $(D5)$, which implies $w_s^F(y) < r$. Because the agent is loss averse, $\lambda > 1$, $w_f^F(y)$ satisfying $(D5)$ cannot be a solution on its own as it induces considerable disutility, leading the agent to reject the contract. Also $(D4)$ on its own is not optimal, as the principal would be fully protecting the agent from losses, demotivating him to exert high effort to avoid the disutility from experiencing losses. Hence, the optimal contract combines the first-order conditions $(D4)$ and $(D5)$. The transition from $w_s^F(y)$ satisfying $(D4)$ to $w_s^F(y)$ satisfying $(D5)$ is defined next. Let $\hat{y}_s \in (\underline{y}, \overline{y})$ be the output level satisfying:

$$\phi \int_{\hat{y}_s}^{\overline{y}} u(w_s^F(y)) f(y|e_H)\,\mathrm{d}y + \eta \int_{\hat{y}_s}^{\overline{y}} v\left(u(w_s^F(y)) - u(r)\right) f(y|e_H)\,\mathrm{d}y + \phi \int_{\underline{y}}^{\hat{y}_s} u(w_s^F(y) f(y|e_H)\mathrm{d}y$$
$$- \lambda\eta \int_{\underline{y}}^{\hat{y}_s} v\left(u(r) - u(w_s^F(y))\right) f(y|e_H)\mathrm{d}y - c = \overline{U}. \quad (D10)$$

The existence of $\hat{y}_s$ is guaranteed by the fact that the two solutions given by Eqs. $(D4)$ and $(D5)$ make the participation constraint bind for gains and losses. Uniqueness of $\hat{y}_s$ is because the magnitude of the first four expressions depends on $\hat{y}_s$. The first three expressions in the left-hand side of Eq. $(D10)$ are positive and become larger as $\hat{y}_s$ decreases, while the fourth expression is negative and becomes larger as $\hat{y}_s$ increases. Since $\overline{U}$ is constant, there exists a unique $\hat{y}_s$ that satisfies $(D10)$.

As a result, the optimal incentive scheme is given by:

$$w_s^*(y) = \begin{cases} w_s^F(y) & satisfying\ (A6)\ \ if\ y \geq \hat{y}_s, \\ w_s^F(y) & satisfying\ (A7)\ \ if\ y < \hat{y}_s. \end{cases} \quad (D11)$$

Notice that this solution exhibits a discrete jump at $y = \hat{y}_s$ since $\lambda > 1$ appears in the denominator of the right-hand side of $(D5)$ but this coefficient does not enter in $(D4)$.

Now suppose that $S(y) \geq r_p + w(y)$. Since principal's and agent's objective functions are identical to thos studied in Theorem 1, that solution remains optimal.

Denote by $\hat{y}_p \in [\underline{y}, \overline{y}\,]$ the output level satisfying $S(\hat{y}_p) - r_p - w_s^F(\hat{y}_p) = 0$. The existence of that output level is guaranteed by $S' > 0$, $S(0) = 0$, $w_s^{F\prime}(y) > 0$ in $y > \hat{y}_s$, and $w_{SB} = 0$ in $y < \hat{y}_s$. There follow two relevant cases. Namely, $\hat{y}_s < \hat{y}_p$ and $\hat{y}_s > \hat{y}_p$.

Let $\hat{y}_s < \hat{y}_p$. If $y < \hat{y}_s < \hat{y}_p$, both agent and principal are in the domain of losses. Then, $w_s^*(y) = 0$ is given when $U(e_H, w(\tilde{y}), r)$ is S-shaped while $w_s^*(y)$ satisfying $(D5)$ is given when $U(e_H, w(y), r)$ is concave. If $\hat{y}_s < y < \hat{y}_p$, the principal is in the domain of losses, while the agent is in the domain of gains. In that case, the principal offers insurance to the agent by paying $w_s^F(y)$

satisfying $(D4)$. Finally, for $\hat{y}_s < \hat{y}_p < y$ both principal and agent are in the domain of gains, so the principal offers $w_s^F(y)$ satisfying $(A5)$. Since $\lambda_p > 2$ is absent in $(A5)$ but present in $(D4)$, the agent's compensation exhibits a kink at $y = \hat{y}_p$.

Let $\hat{y}_p < \hat{y}_s$. If $y < \hat{y}_p < \hat{y}_s$, both agent and principal are in the domain of losses. Again, $w_s^*(y) = 0$ is given when $U(e_H, w(\tilde{y}), r)$is S-shaped and $w_s^*(y)$ satisfying $(D5)$ is given when $U(e_H, w(\tilde{y}), r)$ is concave. If $\hat{y}_p < y < \hat{y}_s$, the agent is in the domain of losses, while the principal is in the domain of gains. The solution is in that case identical to Theorem 1. Namely, the principal offers $w_s^*(y) = 0$ when $U(e_H, w(\tilde{y}), r)$is S-shaped and $w_s^*(y)$ satisfying $(D5)$ is given when $U(e_H, w(\tilde{y}), r)$ is concave. Finally, for $\hat{y}_p < \hat{y}_s < y$ both are in the domain of gains, the principal offers $w_s^F(y)$ satisfying $(A3)$. There is no kink in that case. ■

When output is high enough to ensure $S(y) \geq r_p + w(y)$, the principal is in the domain of gains and her objective function is identical to that in the problem studied in Section 3. In this case her loss aversion does not affect optimal contracting. As a result, the optimal contracts are exactly those presented in Theorem 1. This part of the solution constitutes Proposition D.1 (i).

For output levels ensuring $S(y) < r_p + w(y)$, the principal's loss aversion can affect optimal contracting. When output is sufficiently low so that $y < \hat{y}_s$ also holds, the loss-averse principal transfers most of the risk to the agent. If $U(e, w_s^*(\tilde{y}), r)$ is S-shaped, transfers are set as low as possible because the agent is risk seeking. Instead, if $U(e, w_s^*(\tilde{y}), r)$ is concave, transfers are non-zero but are low enough as to locate the agent in the domain of losses. This shape of the optimal contracts presented in Proposition D.1 (ii) and (iii) are exactly like those in Theorem 1.

For higher output levels, so that $y \geq \hat{y}_s$ holds but $S(y) < r_p + w(y)$ is also true, the principal needs to offer some insurance to the risk-averse agent. However, the principal's loss aversion implies that more risk will be transferred to the agent as compared to the solution in Theorem 1. This is achieved by offering lower-powered incentives in the segment $y \in (\hat{y}_s, \hat{y}_p)$. This immediately implies the existence of a kink in the incentives scheme around $\hat{y}_p$, the point at which the principal herself transitions from losses to gains. After that point, she offers incentives that are as high-powered as in Theorem 1. This kink around $\hat{y}_p$ and the lower-powered incentives in $y \in (\hat{y}_s, \hat{y}_p)$ constitute a slight modification to Theorem 1. This modification is reflected in Proposition D.1 (i) and (ii).

D.2. Adverse Selection followed by Moral Hazard

The assumption that the principal is fully informed about the agent's risk preferences is typically made in models of moral hazard. However, in the framework considered in this paper, this assumption becomes more demanding as she not only needs to know the agent's utility curvature but also his parameter of loss aversion. This extension relaxes the assumption that the principal exactly knows the agent's risk preferences.

Suppose that the principal is perfectly informed about the agent's utility but that she does not know his degree of loss aversion. For simplicity, assume that she can contract with agents with either high or low degrees of loss aversion. Formally, let $\lambda_i \in \{\lambda_L, \lambda_H\}$ where $\lambda_H > \lambda_L > 1$. Contracting with an agent with λ_H occurs with probability ω, while contracting with an agent with λ_L occurs with the complement probability, $1 - \omega$.

The timing of the interaction between agent and principal is as follows. First, nature moves and determines λ_i, which is private information to the agent. Second, the principal offers a menu of contracts. Third, the agent self-selects into a contract. Fourth, e is chosen by the agent. Finally, y realizes and the agent is paid according to the transfers specified in the contract selected by the agent.

The principal's objective is to design contracts that enable her to screen agents to ensure participation and motivation. The following proposition shows that the optimal menu of contracts consists of two bonus contracts enhanced with informational rents.

Proposition D.2 *Under A1-A4 and that λ_i is unknown to the principal, the optimal menu of contracts is the tuple $\{w_s^*(y)^H, w_s^*(y)^L\}$ such that*

i) $w_s^*(y)^H$ *is the second-best optimal contract from Theorem 1 (i) - (ii),*

ii) $w_s^*(y)^L$ *is the second-best optimal contract from Theorem 1 (i) - (ii) satisfying* $U(e_H, w_s^*(y)^L, r, \lambda_L) = \overline{U} + (\lambda_H - \lambda_L) \int_{\underline{y}}^{\hat{y}_H} v(u(r) - u(w_s^*(y)^H)) f(y|e_H) \mathrm{d}y$.

Where $\hat{y}_H \in (\underline{y}, \overline{y})$ is the critical threshold specified in $w_s^*(y)^H$.

Proof. The agent with λ_H faces the following adverse selection constraint,

$$\int_{\underline{y}}^{\overline{y}} u(w(y)^H) f(y|e_H)\mathrm{d}y + \int_{\hat{y}_H}^{\overline{y}} v(u(w(y)^H) - u(r)) f(y|e_H)\mathrm{d}y - \lambda_H \int_{\underline{y}}^{\hat{y}_H} v(r - w(y)^H) f(y|e_H)\mathrm{d}y - c$$
$$\geq \max_{e\in\{e_L,e_H\}} \left\{ \int_{\underline{y}}^{\overline{y}} u(w(y)^L)\, f(y|e)\mathrm{d}y \right.$$
$$\left. + \int_{\hat{y}_L}^{\overline{y}} v(w(y)^L - r) f(y|e)\mathrm{d}y - \lambda_H \int_{\underline{y}}^{\hat{y}_L} v(r - w(y)^L) f(y|e)\mathrm{d}y - c(e) \right\}, \qquad (D12)$$

moral hazard incentive constraint,

$$\int_{\underline{y}}^{\overline{y}} u(w(y)^H) f(y|e_H)\mathrm{d}y + \int_{\hat{y}_H}^{\overline{y}} v(u(w(y)^H) - u(r)) f(y|e_H)\mathrm{d}y - \lambda_H \int_{\underline{y}}^{\hat{y}_H} u(r - w(y)^H) f(y|e_H)\mathrm{d}y - c$$
$$\geq \int_{\underline{y}}^{\overline{y}} u(w(y)^H) f(y|e_L)\mathrm{d}y$$
$$+ \int_{\hat{y}_H}^{\overline{y}} v\big(u(w(y)^H) - u(r)\big) f(y|e_L)\mathrm{d}y - \lambda_H \int_{\underline{y}}^{\hat{y}_H} v\big(u(r) - u(w(y)^H)\big) f(y|e_L)\mathrm{d}y, \qquad (D13)$$

and participation constraint

$$\int_{\underline{y}}^{\overline{y}} u(w(y)^H) f(y|e_H)\mathrm{d}y$$
$$+ \int_{\hat{y}_H}^{\overline{y}} v(u(w(y)^H) - u(r)) f(y|e_H)\mathrm{d}y - \lambda_H \int_{\underline{y}}^{\hat{y}_H} v\big(u(r) - u(w(y)^H)\big) f(y|e_H)\mathrm{d}y - c$$
$$\geq \overline{U}. \qquad (D14)$$

Similarly, the agent with λ_L faces the following adverse selection constraint,

$$\int_{\underline{y}}^{\overline{y}} u(w(y)^L) f(y|e_H)\mathrm{d}y + \int_{\hat{y}_L}^{\overline{y}} v(u(w(y)^L) - u(r)) f(y|e_H)\mathrm{d}y - \lambda_L \int_{\underline{y}}^{\hat{y}_L} v\big(u(r) - u(w(y)^L)\big) f(y|e_H)\mathrm{d}y - c$$
$$\geq \max_{e\in\{e_L,e_H\}} \left\{ \int_{\underline{y}}^{\overline{y}} u(w(y)^H) f(y|e_H)\mathrm{d}y + \int_{\hat{y}_H}^{\overline{y}} v\big(u(w(y)^H) - u(r)\big) f(y|e)\mathrm{d}y \right.$$
$$\left. - \lambda_L \int_{\underline{y}}^{\hat{y}_H} v\big(u(r) - u(w(y)^H)\big) f(y|e)\mathrm{d}y - c(e) \right\}, \qquad (D15)$$

moral hazard incentive constraint,

$$\int_{\underline{y}}^{\overline{y}} u(w(y)^L)f(y|e_H)\mathrm{d}y + \int_{\hat{y}_L}^{\overline{y}} v(u(w(y)^L) - u(r))f(y|e_H)\mathrm{d}y - \lambda_L \int_{\underline{y}}^{\hat{y}_L} v(u(r) - u(w(y)^L))f(y|e_H)\mathrm{d}y - c$$
$$\geq \int_{\underline{y}}^{\overline{y}} u(w(y)^L)f(y|e_L)\mathrm{d}y + \int_{\hat{y}_L}^{\overline{y}} v\big(u(w(y)^L) - u(r)\big)f(y|e_L)\mathrm{d}y$$
$$- \lambda_L \int_{\underline{y}}^{\hat{y}_L} v\big(u(r) - u(w(y)^L)\big)f(y|e_L)\mathrm{d}y, \qquad (D16)$$

and participation constraint

$$\int_{\underline{y}}^{\overline{y}} u(w(y)^H)f(y|e_H)\mathrm{d}y$$
$$+ \int_{\hat{y}_L}^{\overline{y}} v(u(w(y)^L) - u(r))f(y|e_H)\mathrm{d}y - \lambda_L \int_{\underline{y}}^{\hat{y}_L} v\big(u(r) - u(w(y)^L)\big)f(y|e_H)\mathrm{d}y - c$$
$$\geq \overline{U}. \qquad (D17)$$

The agent with λ_L mimicking the agent with λ_H derives the following utility $U(\hat{e}, w_H, r, \lambda_L)$ for a given effort level $\hat{e}$,

$$U(\hat{e}, w(y)^H, r, \lambda_L)$$
$$= \int_{\underline{y}}^{\overline{y}} u(w(y)^H)f(y|\hat{e})\mathrm{d}y + \int_{\hat{y}_H}^{\overline{y}} v(u(w(y)^H) - u(r))f(y|\hat{e})\mathrm{d}y$$
$$- \lambda_H \int_{\underline{y}}^{\hat{y}_H} v\big(u(r) - u(w(y)^H)\big)f(y|\hat{e})\mathrm{d}y$$
$$+ (\lambda_H - \lambda_L) \int_{\underline{y}}^{\hat{y}_H} v\big(u(r) - u(w(y)^H)\big)f(y|\hat{e})\mathrm{d}y - c(\hat{e})$$
$$= U(\hat{e}, w_H, r, \lambda_H) + (\lambda_H - \lambda_L) \int_{\underline{y}}^{\hat{y}_H} v(u(r) - u(w(y)^H))f(y|\hat{e})\mathrm{d}y. \qquad (D18)$$

Since $\lambda_H > \lambda_L$ and $r > w(y)^H$ in $y \in (\underline{y}, \hat{y}_H)$, the agent derives informational rents. The agent with λ_H mimicking the agent with λ_L derives the following utility for a given effort level $\hat{e}$,

$$U(\hat{e}, w(y)^L, r, \lambda_H)$$
$$= \int_{\hat{y}_L}^{\overline{y}} v(u(w(y)^L) - u(r))f(y|\hat{e})\mathrm{d}y - \lambda_L \int_{\underline{y}}^{\hat{y}_L} v\big(u(r) - u(w(y)^L)\big)f(y|\hat{e})\mathrm{d}y - c(\hat{e})$$
$$- (\lambda_H - \lambda_L) \int_{\underline{y}}^{\hat{y}_L} v\big(u(r) - u(w(y)^L)\big)f(y|\hat{e})\mathrm{d}y - c(\hat{e})$$
$$= U(\hat{e}, w(y)^L, r, \lambda_L) - (\lambda_H - \lambda_L) \int_{\underline{y}}^{\hat{y}_L} v(u(r) - u(w(y)^L))f(y|\hat{e})\mathrm{d}y. \qquad (D19)$$

Eq.$(D19)$ shows that engaging in that strategy is not profitable. Next, use Eqs. $(D18)$ and $(D19)$ to rewrite the adverse selection constraints in Eqs. $(D12)$ and $(D15)$ as follows:

$$\int_{\hat{y}_H}^{\overline{y}} u(w(y)^H - r) f(y|e_H) \mathrm{d}y - \lambda_H \int_{\underline{y}}^{\hat{y}_H} u(r - w(y)^H) f(y|e_H) \mathrm{d}y - c \geq \max_{e \in \{e_L, e_H\}} \left\{ U(e, w(y)^L, r, \lambda_L) - (\lambda_H - \lambda_L) \int_{\underline{y}}^{\hat{y}_L} u(r - w(y)^L) f(y|e) \mathrm{d}y \right\}, \quad (D20)$$

and

$$\int_{\hat{y}_L}^{\overline{y}} u(w(y)^L - r) f(y|e_H) \mathrm{d}y - \lambda_L \int_{\underline{y}}^{\hat{y}_L} u(r - w(y)^L) f(y|e_H) \mathrm{d}y - c \geq \max_{e \in \{e_L, e_H\}} \left\{ U(e, w(y)^H, r, \lambda_H) + (\lambda_H - \lambda_L) \int_{\underline{y}}^{\hat{y}_H} u(r - w(y)^H) f(y|e) \mathrm{d}y \right\}, \quad (D21)$$

respectively.

From the above equations it can be concluded that $(D14)$ and $(D20)$ imply $(D17)$, so it must be that $(D17)$ slacks at the optimum while $(D14)$ binds. Moreover, since $(D18)$ and $(D19)$ show that only the agent with λ_L derives profits when mimicking, then $(D20)$ and is strictly satisfied and $(D21)$ binds at the optimum. Denote by $w_S^*(y)^i$ the contract from Theorem 1. From the proof of that Theorem, it is known that e_H generates high effort. This reduces the number of constraints to two, namely:

$$U(e_H, w_S^*(y)^H, r, \lambda_H) = \overline{U}, \qquad (D22)$$

and

$$U(e_H, w_S^*(y)^L, r, \lambda_L) = U(e_H, w_S^*(y)^H, r, \lambda_H) + (\lambda_H - \lambda_L) \int_{\underline{y}}^{\hat{y}_H} v\left(u(r) - u\left(w_S^*(y)^H\right)\right) f(y|e_H) \mathrm{d}y. \quad (D3)$$

Solving the above equations yields that $w_S^*(y)^H$ must satisfy $U(e_H, w_S^*(y)^H, r, \lambda_H) = \overline{U}$, and $w_S^*(y)^L$ must yield $U(e_H, w_S^*(y)^L, r, \lambda_L) = \overline{U} + (\lambda_H - \lambda_L) \int_{\underline{y}}^{\hat{y}_H} v(u(r) - u(w_S^*(y)^H)) f(y|e_H) \mathrm{d}y$.■

According to Theorem 1, bonus contracts ensure that agents with reference-dependent preferences participate and exert high effort. Furthermore, Corollary 2 shows that the bonus feature of the contract appears regardless of the agent's degree of loss aversion. That is because the location of the bonus will be shifted to account for any level of loss aversion. Therefore, the optimal menu consists of bonus contracts.

Screening is guaranteed by complementing the contract targeting agents with lower loss aversion, $w_S^*(y)^L$, with an informational rent. The goal of the rent is to discourage these agents from mimicking high loss-averse agents. To achieve that, its magnitude is such that they become exactly indifferent between engaging in a strategy of mimicking or not doing so. On the other hand, agents with high loss aversion are not willing to engage in a strategy of mimicking. As suggested by Corollary 2, choosing $w_s^*(y)^L$ instead of $w_s^*(y)^H$ would considerably expose these agents to considerably losses and generate disutility.