Incentive design for reference-dependent preferences

Víctor González-Jiménez[a],

[a]*Department of Economics, University of Vienna*

Abstract

I investigate the optimal design of incentives when the agent's risk preferences exhibit reference dependence. The theoretical framework incorporates the most prominent representations of reference-dependent preferences and reference point rules. A common finding emerges: The optimal incentive scheme must include a bonus that divides the performance space in gains, all performance levels for which the bonus is awarded, and losses, all performance levels that fail to obtain the bonus. The bonus feature of the contract allows the principal to exploit the agents' irrationalities of loss aversion and diminishing sensitivity to either offer optimal insurance or to optimally elicit high effort levels. This paper provides a rationale for incentive schemes including bonuses grounded in preference.

JEL Codes: J41, D90, C91, D81
Keywords: Prospect theory, Contracts, Reference-dependence, Incentives, Bonuses.

1. Introduction

Abundant empirical evidence from decision theory shows that individuals, when making decisions under risk, evaluate outcomes relative to a reference point (Tversky and Kahneman, 1992, Abdellaoui et al., 2007, Abdellaoui et al. 2008, von Gaudecker et al. 2011, Vieider et al. 2015, Baillon et al. 2020). This way of evaluating risky alternatives deviates from expected utility since the reference point can change across alternatives. A property that generates major irrationalities and that is able to explain relevant economic phenomena that expected utility cannot. For example, the regularity that individuals are risk averse for lotteries with small stakes while not being absurdly risk averse for lotteries featuring large stakes (Rabin, 2000).

The solution to the conventional moral hazard problem demonstrates the importance of risk preference for incentive design. According to that solution, the optimal contract results from a tradeoff faced by the principal between providing insurance to the agent and eliciting high effort levels (Mirrlees 1976, Holmström, 1979). As a result, contracts with higher-powered incentives are given as the agent is more risk averse. This optimal tradeoff conveys a powerful, yet relatively unexplored implication: Abandoning expected utility to adopt *descriptively valid* theories of risk to characterize preference generates frameworks with the capacity to design contracts that can more effectively motivate individuals and to predict incentive schemes that are widely used by organizations.

This paper incorporates reference-dependent preferences in the optimal design of incentives. Its main result is that contracts including a bonus, i.e. an upward discontinuity in payments, emerge as optimal solutions regardless of the theory of risk used to model these preferences and irrespective of the rule used to define reference points. I thus provide a foundation to the widespread practice of paying bonuses grounded in preference.

A bonus contract emerges as an optimal solution since it allows the principal to exploit the agent's biases of loss aversion and diminishing to her benefit. Specifically, the optimal contract pays the lowest possible transfer in the contingency that low output levels realize, deliberately locating the agent in the domain of losses. The agent's risk seeking from diminishing sensitivity at losses generates a tolerance to that considerable exposure to risk. Also, the agent's loss aversion motivates him to exert high effort to avoid incurring in these considerable losses. Moreover, a bonus transitions the agent from the domain of losses to the domain gains. In fact, the magnitude of such bonus is designed so that the losses included by the contract are, on expectation, outweighed. In this way the principal ensures that the contract will be accepted by the agent. Hence, a contract paying excessively low amounts at low but likely realizations and paying a sizable bonus at high but unlikely realizations is offered and accepted by the agent.

While I am not the first to incorporate reference-dependence in a principal-agent framework, existing studies differ in the way risk attitude is characterized as well as in the considered reference point rule (de Meza and Webb, 2007, Dittmann et al., 2010, Herweg et al., 2010, Corgnet et al., 2018). These disagreements have led to diverse solutions and interpretations in the literature. For example, two acutely different payment modalities such as stochastic contracts—in which the principal turns a blind-eye to the agent's performance signals with some ex-ante probability—and option-like contracts—in which incentives are performance-insensitive at high performance levels—are optimal under loss aversion (de Meza and Webb, 2007, Herweg et al., 2010).

The theoretical framework of this paper incorporates and unifies the most prominent approaches to modeling reference-dependent preferences, such as cumulative prospect theory (Tversky and Kahneman, 1992), disappointment models with prior (Bell, 1985, Loomes and Sugden, 1986, Gul, 1991) and disappointment models without prior (Delquié and Cilo, 2006, Kőszegi and Rabin, 2006). The model also considers the most prominent reference point rules, such as the status quo, max-min, goals and aspirations as reference points, expectations-based reference point, and the outcomes of the contract as the reference point. To the best of my knowledge, this is the first paper to provide such a general framework. This enables me to obtain novel and generalizable results. For example, that a bonus contract is first-best optimal under all preference representations. Moreover, the generality of the model enables me to reconcile previous findings and to determine how previous approaches arrived to different solutions.

This paper contributes to previous literature in several ways. First, it provides a justification for bonuses that contrasts with the standard explanation of limited liability (Park 1995, Kim 1997, Oyer, 1998). A notorious disadvantage of attributing the optimality of bonus contracts to limited liability is that such result is irreconcilable with the regularity of individuals being predominantly risk averse. This study pursues a completely different approach: It characterizes risk preference using descriptive theories of risk to endow the agent with the well-established biases of loss aversion and diminishing sensitivity. After solving the principal's problem, I find that bonus contracts emerge as optimal as they enable the principal to profitably exploit these irrationalities.

Second, I am the first to provide a *comprehensive* analysis of incentives for reference dependence. This is done by considering both loss aversion and diminishing sensitivity. Moreover, I characterize the first-best contract, which fully insures the agent against risk, to subsequently investigate how the second-best contract deviates from that benchmark to implement punishments and rewards. Such comprehensive analysis is missing in the literature and, due to the novelty of its results, is highly relevant. For instance, I find that the first-best contract already exhibits a bonus. A result that contrasts previous explanations for bonuses based on their motivational influence under moral hazard (Dittman et al. 2010, Herweg et al., 2010). Furthermore, I show that the second-best contracts elicit high effort because its bonus feature generates sizable punishments with respect to the first-best contract.

Third, this study is the first to employ a general specification of reference dependence to study the effect of these preferences in a setting of moral hazard. The specification is taken from Baillon et al. (2020) and is applied to the principal-agent framework. This general framework produces solutions to the principal's problem that incorporate properties that are deemed as desirable by previous studies. For instance, the emergence of discontinuities in the pay schedule (De Meza and Webb, 2007, Dittman et al., 2010, Herweg et al., 2010), having pay insensitive segments at low output levels (Dittman et al., 2010), and exhibiting performance-pay for high performance levels (De Meza and Webb, 2007). Moreover, the model is flexible enough to reproduce other important results in the literature. For example, the emergence of stochastic contracts under severe loss aversion (Herweg et al., 2010, Corgnet et al., 2019).

The generality of my framework, its capacity to incorporate desirable properties, and its flexibility imply that its results can be extrapolated to other economic settings in which incentive design is crucial. For instance, insurance markets, financial markets, and agricultural economies. Besides, this study shows how practitioners can apply this general framework to model of reference-dependence preferences without having to worry about the generality of their findings due to their chosen preference specification and/or reference point rule.

Finally, this paper fills significant gaps in the literature. I am the first to characterize optimal contracts when production goals are taken by the agent as reference points. In a previous paper, Corgnet et al., (2018) characterize second-best contracts in a moral hazard setting but confine themselves to linear contracts. I am also the first to provide a solution to the principal's problem in a setting whereby the agent exhibits expectations-based reference points. A rule that has received considerable attention in the behavioral economics literature (Abeler et al., 2011, Terzi et al. 2015, Sprenger, 2015, Gneezy et al. 2017, Baillon et al. 2020).

2. The general setup

Consider a principal (she) hiring an agent (he) to produce output on a task. Production output is a random variable that may take any value in the interval $y \in [\underline{y}, \overline{y}]$, where $\underline{y} \geq 0$. The agent's action consists on choosing an effort $e \in \{e_L, e_H\}$ on the task. For simplicity, it is assumed that only exerting high effort is costly to the agent.

Assumption 1 (A1). $c(e) = \begin{cases} c \; if \; e_H, \\ 0 \; if \; e_L. \end{cases}$ Where $c > 0$.

Furthermore, it is assumed that both agent and principal know that output is distributed according to the cumulative density function $F(y|e)$, which admits a probability density function $f(y|e)$. Importantly, output and effort relate according to the monotone likelihood ratio property.

Assumption 2 (A2). $\frac{\partial}{\partial y}\left(\frac{f(y|e_L)}{f(y|e_H)}\right) < 0 \; \forall \; y \in [\underline{y}, \overline{y}]$.

The implications of Assumption 2 are well-known: Higher performance levels are more likely to be drawn from a probability density function conditional on high effort rather than from a probability density function conditional on lower effort.

To incentivize high effort, the principal offers the agent a take-it-or-leave-it contract including a transfer $w(y) \geq 0$. The agent can accept the contract and subsequently choose a level of effort e, or, alternatively, reject the contract and obtain his reservation utility, $\overline{U} \geq 0$.[1] When the contract is accepted, payments are made according to the rule $w(y)$ after the realization of y is known to both parties.

I assume that the principal is risk neutral and thus able to pool multiple risks. Specifically, her objective function is:

$$\int_{\underline{y}}^{\overline{y}} \big(S(y) - w(y)\big) f(y|e) dy, \qquad (1)$$

where $S'(y) > 0$ and $S''(y) < 0$.

The agent's preferences, i.e. how the transfer $w(y)$ enters the agent's objective function, is the main investigation of this paper. Standard theory proposes that the agent derives utility from receiving the contract by evaluating its implied transfers in an absolute way. The following assumption formalizes that proposal.

Assumption 3 (A3). *The agent's utility is a C^2 function u: $\mathbb{R}^+ \cup \{+\infty\} \longrightarrow \mathbb{R} \cup \{-\infty, +\infty\}$ with $u'(w(y)) > 0$, $u''(w(y)) < 0$, and differentiable inverse $h\big(w(y)\big) := u^{-1}\big(w(y)\big)$.*

Notice that Assumption 3 also posits that the agent's utility exhibits diminishing returns to transfers. Hence, within the framework of expected utility theory, he is risk averse.

The first-best and second-best solutions to the principal's problem when the agent's risk preferences are characterized by expected utility are attributed to Borch (1962) and Holmström (1979), and are presented in the following proposition.

Proposition 1. *Under assumptions A1-A3:*

i. *The first-best contract, w_{FB}, pays a constant transfer equal to $w_{FB} = h'\left(\frac{1}{\nu}\right)$ where $\nu > 0$ is a constant.*

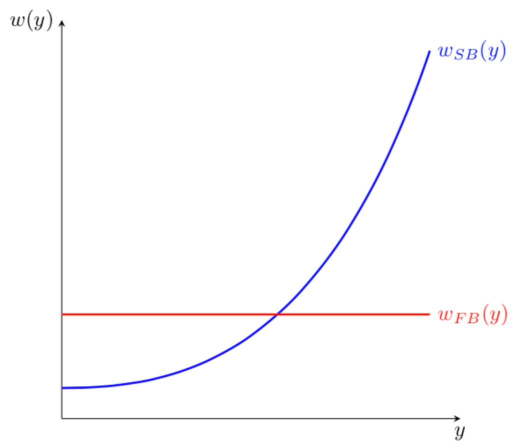
[1] Since $\overline{U} \geq 0$, that $w(y)$ is nonnegative does not necessarily imply an absence of punishments as, according to Assumption 1, $c > 0$. Therefore, by setting a small enough $w(y)$ the principal can punish the agent by making him worse off than his outside option utility level.

ii. *The second-best contract* $w_{SB}(y)$ *is a transfer schedule that is everywhere increasing in* y *and satisfies* $w_{SB}(y) = h'\left(\frac{1}{\nu+\gamma\left(1-\frac{f(y|e_L)}{f(y|e_H)}\right)}\right)$ *where* $\gamma > 0$ *and* $\nu > 0$ *are constants.*

The first part of Proposition 1 shows that when effort is contractible, the optimal contract fully insures the agent. Specifically, the agent is paid a constant amount that not only protects him from the risk inherent in the stochastic nature of output but also that is sizeable enough to make him indifferent between accepting or not the contract.

The second part of the proposition shows that when effort is not contractible, the optimal contract motivates the agent by paying higher transfers in exchange of higher output levels. When facing such scheme, the agent understands that exerting high effort increases the likelihood of obtaining a high compensation. Moreover, to ensure that the contract is accepted the principal ensures that the strength of the incentives, i.e. the degree at which $w_{SB}(y)$ increases in y, implies a degree of risk that is tolerable to the risk averse agent. Figure 1 illustrates the first-best and second-best contracts described in Proposition 1.

Figure 1. Illustration of the contracts from Proposition 1



In the following sections, Assumption 3 will be relaxed to allow for more descriptive theories of risk. Generally speaking, it will be assumed that the agent exhibits reference-dependent preferences. That is, utility is generated by the transfers of the contract implied by the contract *relative* to a reference point, which may be exogenous or endogenous to the principal's offer.

3. Prospect Theory preference

The first approach to model reference dependence posits that that the agent's preferences are characterized by cumulative prospect theory (Tversky and Kahneman, 1992). This theory of risk states that the agent evaluates the outcomes specified by the contract $w(y)$ relative to a reference point, $r > 0$. Outcomes above this point are classified as *gains* while outcomes below this point are classified as *losses*. A relevant departure of prospect theory with respect to expected utility is that the agent exhibits different risk preferences for gains and losses. This property is captured by the value function $v(w,r)$, which is formally described in the next assumption.

Assumption 4 (A4). *The agent's value function* $v\colon \mathbb{R}^+ \cup \{+\infty\} \longrightarrow \mathbb{R} \cup \{-\infty, +\infty\}$ *is the piece-wise function:*

$$v(w,r) = \begin{cases} u(w(y) - r) & \text{if } w(y) \geq r, \\ -\lambda u\big(r - w(y)\big) & \text{if } r < w(y). \end{cases}$$

with properties:

- $u(0) = 0$;
- $u'(\cdot) > 0 \;\forall\, y \in \left[\underline{y}, \overline{y}\right]$;
- $u''(\cdot) < 0$;
- $\lambda > 1$;
- $u(\cdot)$ *has a differentiable inverse* $h(\cdot) := u^{-1}(\cdot)$.

There are two sources of risk attitude emerging from the value function described by Assumption 4. First, the curvature of $v(w,r)$ is concave if $w(y) \geq r$ and convex if $w(y) < r$. These shapes generate risk aversion in the domain of gains and risk seeking in the domain of losses, a property that receives the name of diminishing sensitivity in the literature. In the principal-agent framework diminishing sensitivity implies that the agent is willing to accept contracts that expose him to large degrees of risk in the domain of losses but, at the same time, requires contracts that protect him from risk in the domain of gains.

Second, the assumption $\lambda > 1$ implies that the agent is loss averse because in his utility losses loom larger than equally-sized gains. Loss aversion implies that the agent will reject contracts that expose him to losses and that do not properly compensate him for such exposure.

Analyses of prospect theory typically take the reference point r to be exogenous.[2] This is the approach that will be followed in this section, but that will be relaxed in subsequent sections to gain robustness and generalizability. In the considered setting, that r is exogenous implies that the contract offered by the principal does not alter the reference point, and thus does not modify how the agent evaluates risky alternatives.

All in all, the utility of an agent with cumulative prospect theory preferences can be written as:

$$CPT(e, w(y), r) = \int_{\underline{y}}^{\overline{y}} \theta u(w(y) - r) - \lambda(1 - \theta)\big(r - w(y)\big) f(y|e)\mathrm{d}y - c(e), \quad (2)$$

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

At this point it is relevant to highlight that the chosen characterization of prospect theory depicted in equation (2) abstracts from probability weighting, an important source of risk attitude for prospect theory. The main reason for not considering probability weighting is that the main goal of this paper is to study how reference dependence, on its own, affects the optimal design of incentives. The reader interested in the optimal design of contracts when the agent's preferences exhibit reference-dependence *and* probability weighting is referred to González-Jiménez and Castillo (2021). That paper shows that under reference-dependent preferences, bonus contracts emerge even in the presence of probability weighting.

The principal's program consists on implementing a contract that is both accepted by the agent and that incentivizes him to exert high effort. Formally her program is:

$$\max_{\{\mathrm{w(y)}\}} \int_{\underline{y}}^{\overline{y}} \big(S(y) - w(y)\big) f(y|e_H) dy$$

Subject to

$$PC: \qquad CPT(e_H, w(y), r) \geq \overline{U}, \qquad (3)$$

$$IC: \qquad CPT(e_H, w(y), r) \geq CPT(e_L, w(y), r).$$

The solution to (3) is presented in Proposition 2. The most relevant property of that solution is that both the first-best contract and the second-best contract exhibit a bonus that divides the

[2] In initial formulations of prospect theory, the *status quo* or the current welfare position of the decision-maker when making a choice was assumed to be the reference point (Kahneman and Tversky, 1979).

agent's output space into gains and losses. The proofs of the main results of the paper are relegated to Appendix A.

Proposition 2. *Under assumptions A1, A2, A4, and that the agent's preferences are as described by* (2)*, there exist unique output levels* $\hat{y}_s, \hat{y}_f \in (\underline{y}, \overline{y})$ *such that:*

i) *The first-best contract,* $w_{FB}(y)$*, pays the minimum possible if* $y < \hat{y}_f$*, exhibits a bonus at* $y = \hat{y}_f$*, and pays the constant transfer* $w_{FB}(y), = r + h'\left(\frac{1}{v}\right)$ *if* $y > \hat{y}_f$*.*

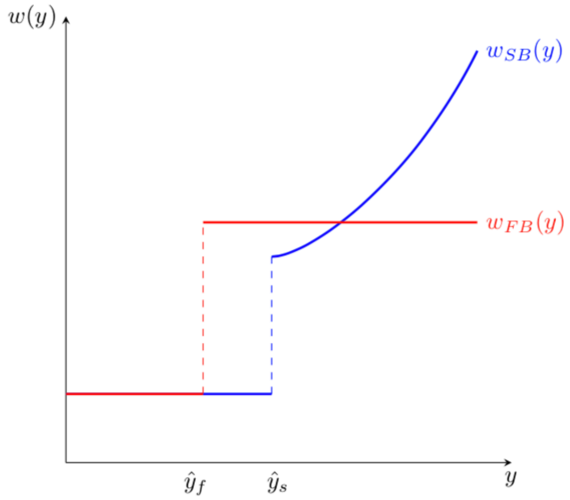
ii) *The second-best contract,* $w_{SB}(y)$*, pays the minimum possible if* $y < \hat{y}_s$*, exhibits a bonus at* $y = \hat{y}_s$*, and increases in* y *according to the schedule* $w_{SB}(y) = r + h'\left(\frac{1}{v+\gamma\left(1-\frac{f(y|e_L)}{f(y|e_H)}\right)}\right)$ *if* $y > \hat{y}_s$*.*

Both first-best and the second-best contract consist of two segments. For low performance levels, i.e. lower than thresholds $\hat{y}_s$ and $\hat{y}_f$, these contracts pay the lowest admissible transfer. Instead, for high performance levels, i.e. higher than thresholds $\hat{y}_s$ and $\hat{y}_f$, the first-best contract offers a constant transfer that exceeds the agent's reference point while the second-best contract consists of an incentive schedule in which transfers strictly increase in the agent's performance. At the performance thresholds $\hat{y}_s$ and $\hat{y}_f$ each contract exhibits a bonus or upward discontinuity in payments that emerges as a result of the combination of these two segments.

Figure 2 presents an example of these incentive schemes. According to standard terminology from personnel economics these schemes can be understood as a lump-sum bonus, in the case of the first-best contract, and an incentive scheme consisting of a salary plus a commission bonus in the case of the second-best contract (See for example Chung et al., (2014)).

The optimal contracts presented in Proposition 2 enable the principal to exploit the agent's biases of loss aversion and diminishing sensitivity. In the case of the first-best contract these biases are used to offer optimal insurance. Specifically, the risk seeking from diminishing sensitivity at losses entails that the agent is willing to tolerate being exposed to large degrees of risk. The principal anticipates this and locates the agent in the domain of losses by giving the lowest possible payment in the contingency that $y < \hat{y}_f$. The bonus feature of the contract, which is given for output levels $y \geq \hat{y}_f$, transitions the agent from the domain of losses to the domain of gains. The existence of this bonus ensures that the agent accepts the contract because its magnitude outweighs the disutility created by the losses included in the contract and generates an expected utility level equal to the agent's outside option.

Figure 2. Illustration of the contracts described by Proposition 3.



In turn, the second-best contract exploits the agent's biases to motivate him. Because of loss aversion, the agent will exert high effort to avoid potential losses. Also, diminishing sensitivity implies that the agent is willing to be more risk seeking, and thus willing to exert high effort, when exposed to losses. As a consequence, the principal locates the agent in the domain of losses by paying the lowest possible in the segment $y < \hat{y}_s$. However, as in the first-best contract, loss aversion implies that a contract that only consists of losses would be rejected. Hence, the contract includes a transition to the domain of gains with a bonus awarded in the segment $y \geq \hat{y}_s$. In addition, in that domain, loss aversion and risk seeking from diminishing sensitivity are absent and the agent is instead risk averse. So, to elicit high effort the principal must offer a high-powered incentive scheme that increases in performance. The bonus thus also reflects a sudden change of incentives at $y = \hat{y}_f$.

Next, I investigate how the second-best contract from Proposition 2 (ii) imparts punishments and rewards with respect to the first-best to motivate the agent in a situation of moral hazard. The following corollary describes these incentives.

Corollary 1. *The second-best contract, $w_{SB}(y)$ from Proposition 2 (ii):*

i) *Exhibits $\hat{y}_s \geq \hat{y}_f$, so punishments are imparted the domain of losses and punishments and rewards are imparted in the domain of gains.*

ii) *Has a unique output level $\tilde{y} \in (\hat{y}_f, \bar{y})$ such that its bonus is weakly smaller than that paid by the first-best from Proposition 2 (i) if $y \geq \tilde{y}$ or larger otherwise.*

The second-best contract elicits high effort both by exploiting the agent's irrationalities and by imparting punishments with respect to the first-best in gains and losses. Figure 2 illustrates these incentives. In $y < \hat{y}_f$, both contracts pay the agent the lowest possible transfer implying an absence of incentives. However, the agent is motivated to exert high effort in that segment by virtue of loss aversion and risk-seeking attitude from diminishing sensitivity. Besides, that the second-best contract exhibits $\hat{y}_s \geq \hat{y}_f$ along with the bonus guarantee the existence of sizeable punishments with respect to the first-best in $y \in (\hat{y}_f, \hat{y}_s)$. Thus, when the agent's irrationalities are, on their own, insufficient to motivate the agent, the contract includes punishments in the domain of losses. Finally, that the second-best contract increases in performance at $y > \hat{y}_s$, generates punishments for low performance and rewards for high performance in the domain of gains. This standard way to motivate the agent, also present in Proposition 1, is a direct consequence of loss aversion and diminishing sensitivity being absent in that segment.

To gain understanding on the influence of risk preference on the implementation of the aforementioned incentives, I present a set of comparative statics. They seek to investigate how the second-best contract and its incentives change when components of the agent's risk preference are changed. First, it is analyzed how changes in the agent's exogenous reference point affects the shape of the second-best contract.

Corollary 2. *A higher reference point r implies that the second-best contract $w_{SB}(y)$ should include a larger monetary bonus that is awarded in the contingency of attaining a higher threshold output $\hat{y}_s$.*

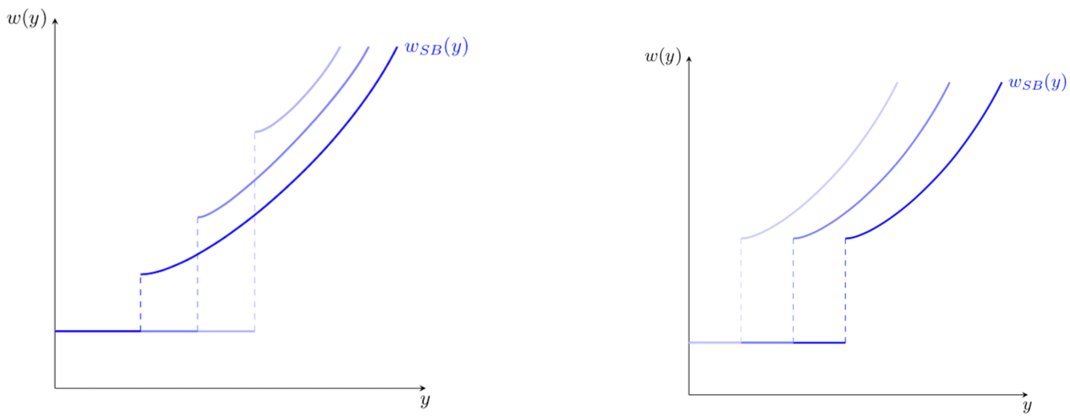
If the worker has a higher exogenous reference point, the size of the bonus must be larger to ensure that the contract indeed transitions the agent to the domain of gains. In turn, the prospect of higher bonuses offered at high output levels enable the principal to extend the output segment at which the agent is exposed to losses. As a consequence, the threshold at which the bonus is awarded, $\hat{y}_s$, is increased.[3] Figure 3a illustrates this comparative static.

The following Corollary shows how changes in loss aversion affects the second-best contract.

Corollary 3. *Higher loss aversion, λ, implies that the second-best contract $w_{SB}(y)$ awards the same monetary bonus in the contingency that a lower threshold output $\hat{y}_s$ is attained.*

Higher loss aversion implies that the agent experiences more disutility when exposed losses. As a consequence, the principal awards the bonus at lower output levels, which in turn reduces the segment in which the agent is exposed to the losses included in the contract. Figure 3 illustrates this comparative static.

Figure 3. Illustration of comparative statics



a) Example of Corollary 2   b) Example of Corollary 3

Notes: In panel a) higher reference point is denoted by lighter color. In panel b) higher loss aversion is denote by lighter color.

The following corollary assumes a CRRA utility to investigate the influence of utility curvature on contract design. It will become evident that this comparative static is less clear-cut than the other two.

Corollary 4. *Let $u(w) = w^{\alpha}$ with $\alpha < 1$. Higher curvature, α, implies that the second-best contract, $w_{SB}(y)$, should include:*

[3] Due to loss aversion the boost in the bonus does not lead to a proportional increment in the threshold $\hat{y}_s$.

i) *a smaller bonus and higher-powered incentives if* $\frac{1}{\nu+\gamma\left(1-\frac{f(y|e_L)}{f(y|e_H)}\right)} > 1$, *or a larger bonus and lower-powered incentives otherwise,*

given in the contingency that

ii) *a smaller threshold* $\hat{y}_s$ *is attained if* $r > 1$ *or a larger threshold* $\hat{y}_s$ *is attained otherwise.*

To explain the intuition of this comparative static, I focus on the cases $r > 1$ and $\frac{1}{\nu+\gamma\left(1-\frac{f(y|e_L)}{f(y|e_H)}\right)} > 1$. When $r > 1$ losses are more sizeable, and a larger utility curvature, α, implies that they will generate more disutility. Therefore, to avoid that the contract is rejected, the principal protects the agent from losses by awarding the bonus at lower output levels. Furthermore, a larger utility curvature, α, implies lower risk seeking in losses and lower risk aversion in gains. Since the bonus reflects a sudden change of incentives in the transition from losses to gains, this change becomes less abrupt.

I conclude this section by relating its results to previous literature. In the closest paper to this one, Dittmann et al. (2010) found a similar second-best contract using prospect theory preferences. This model generalizes their solution by allowing for other reference-point rules, other preference specifications, an arbitrary concave utility function, and an arbitrary probability density. An important takeaway of this paper is that their solution to the principal's problem is more general and powerful than previously thought. In addition, a relevant difference with respect to their study is that Proposition 2 also characterizes the first-best contract, which provides a novel interpretation of bonuses, based on optimal insurance, and enables me to provide a more complete description of the incentives included in the second-best contract (Corollary 1).[4] [5]

---

[4] Another prominent difference with respect to Dittmann et al. (2010) is that I will extend the analysis of this section to account for various reference point rules that are typically used in the literature of decision theory. Among them expectation-based reference points. Such analyses provide robustness and generalize the finding that the optimal contract, first-best or second-best, should include a bonus. Additionally, comparative statics that illustrate the importance of risk preference to incentives that are absent in their paper are included (Corollary 2 and Corollary 4).

[5] The paper by de Meza and Webb (2007) finds, under certain conditions, similar solutions to the principal's problem when the agent exhibits loss aversion and down-side risk loving, i.e. diminishing sensitivity in losses. As in Dittmann et al. (2010) and in stark contrast to my analysis, they solely focus on the second-best contract. Hence, their analysis of incentives and their interpretation for the emergence of bonuses is not as complete. Moreover, their admissible solutions sometimes do not feature bonus, sometimes do not pay the minimum possible wage level, and sometimes do not have a commission payment given for high output levels (See their Figure 2). This

4. Prospect theory with endogenous reference points

4.1 Salience-based reference points

Thus far, the reference point has been assumed to be exogenous to the contract offered by the principal. In this section I analyze how the contracts in Proposition 2 change when endogenous reference point rules that are consistent with prospect theory are assumed. I start by considering salience-based rules. These rules assume that the agent's reference point results from a comparison of the outcomes or probabilities included in the alternatives available to the worker. For the sake of brevity, I focus on the more complete case of moral hazard.

The first rule is the *max-min* (Hershey and Schoemaker, 1985). According to this rule, agents are cautious and take as reference point the maximum value from a set consisting of the minimum possible outcome of each alternative. As an example, suppose that an agent is given two contracts $w_1 = (0.5, 200;\ 0.5, 0)$ and $w_2 = 100$. The set consisting of the minimum possible outcomes of each alternative is $\{0,100\}$, so according to this rule the agent's reference point is 100.

Corollary 5 presents the solution to the principal's problem under the max-min reference point rule. It turns out that the agent's reference point is the first-best contract presented in Proposition 2 (i) and that a second-best contract with a shape similar to that described in Proposition 2 (ii) suffices to motivate the agent.

Corollary 5. *Under assumptions A1, A2, A4, that the agent's preferences are characterized by* (2)*, and the max-min rule, the agent's reference point is* $r = w_{FB}(y)$ *and there exists a unique* $y_{\overline{U}} \in (\underline{y}, \overline{y})$ *such that the second-best contract,* $w_{SB}(y)$*, is constant and pays the minimum possible if* $y < y_{\overline{U}}$*, exhibits a bonus at* $y = y_{\overline{U}}$*, and increases in performance if* $y > y_{\overline{U}}$*.*

The intuition behind this reference point is as follows. The agent makes two choices: whether to accept the contract or not, and whether to exert high effort or not. There exist thus four candidates for reference point, each corresponding to the alternatives implied by the agent's actions. Notice that the expected utility level implied by the first-best contract from Proposition

stark divergence can be partly explained by the fact that they use a different preference representation, which strictly speaking does not belong to cumulative prospect theory nor disappointment models, axiomatized theories of risk that are commonly used to model reference dependence.

2 (i) is the agent's reservation utility, $\overline{U}$, which is, in turn, the welfare level obtained by the agent when the contract is rejected. That utility level is lower than that implied by the minimum amount given to the agent by the second-best contract. Otherwise, that contract would not include punishments for low performance and would not motivate the agent to exert high effort, constituting a contradiction. Therefore, $r = w_{FB}(y)$ under the max-min rule.

Anticipating this reference point, the principal offers a contract that locates the agent in the domain of losses in the contingency of low output levels, that transitions the agent to the domain of gains with a bonus, and that exhibits schedule of transfers that increase in the agent's performance at the high-end of the output space. This contract shape is *similar* to that presented in Proposition (ii) because the agent's risk preferences are characterized by prospect theory under both problems and so that shape guarantees that the principal optimally exploits the agent's irrationalities to her benefit. The only difference between the contracts from Corollary 6 and Proposition 2 and is that $y_{\overline{U}}$, the output level at which the bonus is given under the max-min rule, may not need to coincide with $\hat{y}_s$.

I consider another two salience-based reference points: the *min-max* rule and the $w_{SB}(y)$ *at max* P. The first one implies that individuals take as the reference point the minimum value of a set consisting of the maximum outcome of each alternative. In light of the example given above in which the agent received $w_1 = (0.5, 200;\ 0.5, 0)$ and $w_2 = 100$, this rule implies that his reference point is also 100. The second one states that the output level realizing with the highest probability is adopted as the agent's reference point. In the previously mentioned example, the agent's reference point is 100 under this rule since its associated probability is the highest.

The solutions to the principal's problem when the agent adopts these reference points are described by Corollaries B.1. and B.2. in Appendix B. Not surprisingly, the shapes of these contracts are similar to that described by Proposition 2 (ii). A resemblance explained by agents having risk preferences characterized by prospect theory under all considered rules, and thus requiring a similar exposure to risk to be motivated. As a consequence, bonus contracts remain to be optimal to incentivize the agent despite the reference point being endogenized with salience-based reference point rules.

There is a noteworthy difference among the aforementioned results, which is that the threshold output level that awards the bonus can be different. This is a direct consequence of bonuses, which seek to transition the agent from losses to gains, heavily depending on the specified reference point rule. However, in those solutions both the size of the bonus and the threshold output level irremediably depend on components that remain exogenous to the principal choice: The welfare implied by the agent's outside option and the highest probability implied by the density function. In the next subsection, I investigate the optimal design of incentives when the principal can more directly affect the agent's reference point and thus the implied transition from gains to losses of the contract.

4.2. Goals as reference points

There is abundant evidence showing that individuals incorporate goals as reference points (Heath et al., 1999, Larrick et al., 2009, Allen et al., 2017, Markle et al., 2018). For example, individuals exhibit higher performance in cognitive and/or physical tasks when a high rather than a low goal is set (Heath et al., 1999). The higher performance under high goals is explained by a combination of loss aversion—an aversion to experience psychological losses from missing the goal which motivates higher effort exertion—and diminishing sensitivity—a willingness to exert more effort as the goal is approached as well as a willingness to exert less effort as they move away from the goal. A higher goal thus enlarges the performance levels for which these two phenomena boost motivation. This rationale also explains why consumers save more energy and water when a savings goal is set (Hardig and Hsiaw, 2014, Tiefenbeck et al., 2018) and why college students exhibit better performance when setting a task-based goal (Clark et al., 2020).

In this section, I study the optimal implementation of incentives when the agent incorporates a production goal chosen by the principal as their reference point. Importantly, a production goal can differ from an expectation in that the latter is fully governed by probabilities associated to potential outcomes, while the former can encompass an ambition and hope component. A goal may not need to coincide with the principal's nor the agent's expectations. This important difference makes goals consistent with cumulative prospect theory, a claim that will become evident in the beginning of the next section.

Let $g \in [\underline{y}, \overline{y}]$ be a production goal chosen by the principal. Also, assume that agent's preferences when the goal is taken as the reference point are given by

$$CPT(e, w(y), g) = \int_{g}^{\overline{y}} u(w(y) - w(g)) f(y|e) \mathrm{d}y - \lambda \int_{\underline{y}}^{g} u\big(w(g) - w(y)\big) f(y|e) \mathrm{d}y - c(e). \quad (4)$$

Equation (4) shows that the goal divides the agent's *output space* into gains and losses, and that such division affects the evaluation of the transfers implied in the contract. Specifically, the transfers of the contract are evaluated relative to the payment given when the goal is just met, $w(g)$. Thus, while the agent's reference point is g, such goal enters his objective function in the form of the monetary payment. In this way the unit of the utility domain is kept coherent and consistent through the paper.

In this setting, the principal's problem is dual: choose the production goal, g, and determine how performance around that goal must be incentivized, $w(g)$. Therefore, her proposal consists of the tuple $(w(y), g)$, offered to the agent before he exerts effort on the delegated task.

The following result shows that it is in the best interest of the principal to set higher goals provided that they are accompanied with a contract in which the payment to the agent reflects their magnitude.

Lemma 1. *Under assumptions A1, A2, A4, and that the agent's preferences are characterized by* (4)*, then 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* $y \leq g$.

When goals are accompanied by an extrinsic reward that reflects their difficulty, the principal can motivate the agent by setting challenging goals. However, she must be careful when following this strategy because higher goals generate, under these conditions, larger disutility, which increases the likelihood that the contract is not accepted. In other words, goals are painful motivational devices. A similar result is found by Koch and Nafzinger (2011) in the context of time-inconsistent agents who set non-binding endogenous goals to gain self-regulation.

Knowing that goals are painful motivational devices, the principal needs to design an incentive scheme that is compatible with a strategy of setting challenging goals. Lemma 1 already establishes that such a contract must reflect the magnitude of the goal chosen by the principal, $w'(g) > 0$. However, that lemma is silent about other crucial properties of the contract such the dependence of payments on y, whether $w'(g)$ should hold for all y, and how the agent should be incentivized for levels of output close and away from g.

The following Proposition describes the optimal contract in a setting of moral hazard when the agent's reference point is g, a production goal set by the principal.

Proposition 3. *Under assumptions A1, A2, A4, and that the agent's preferences are characterized by* (4)*, the second-best contract,* $w_{SB}(y)$*, pays the minimum possible in* $y < g$*, increases in performance in* $y > g$*, and exhibits a bonus or discrete jump at* $y = g$ *that increases with the magnitude of* g*.*

The incentive scheme compatible with a strategy of setting challenging goals is similar to that presented in Proposition 2 (ii). However, the bonus included in that contract is now awarded in the contingency that the production goal specified by the principal is achieved or surpassed. Furthermore, a relevant and novel feature of the optimal contract described by Proposition 3 is that the magnitude of the bonus increases with the production goal specified by the principal. This property ensures that the disutility that the agent derives from more challenging goals (Lemma 1) is counteracted with a payment that reflects the magnitude of these goals.[6]

The following corollary fully characterizes the principal's solution for this contracting setting. It does so by describing the optimal goal that accompanies the bonus contract described in Proposition 3. I find that while the optimal goal must be challenging, to motivate high effort, it must be on expectation attainable, so that the bonus included in the contract transitions the agent from losses to gains.

Corollary 6. *Let* $\lim_{x \to 0} u'(x) = +\infty$*. The solution to the principal problem consists of a tuple* $(w_{SB}(y), g^*)$ *where* $w_{SB}(y)$ *is described by Proposition 3 and* g^* *satisfies* $\mathbb{E}(w_{SB}(y)) - w_{SB}(g^*) = \epsilon$ *for arbitrarily small* $\epsilon > 0$*.*

---

[6] This characteristic of the contract is also consistent with the comparative static presented in Corollary 2, which showed that a higher exogenous reference point increased the magnitude of the bonus and the output threshold level after which it is awarded.

Under an additional mild condition on the agent's utility, i.e. the Inada condition, the principal sets a challenging but on expectation attainable production goal and accompanies it with contract that awards a bonus when the production goal is met. This result of the model is able to capture the regularity of firms paying bonuses when an individual or organization goal is achieved. According to Worldatwork (2018) more than 75% of American firms award bonused on such a way when compensating employees.

The theoretical framework considered in this section differs from existing models in the literature of goal setting in economics in several ways. First, I consider a setting of risk in which achieving the goal is uncertain for the decision maker. Most models in the literature assume a deterministic environment in which the goal can be achieved with certainty (Wu et al., 2008, Gomez-Minambres, 2012, Corgnet et al., 2015, Dalton et al., 2016b, Brookins et al., 2017). Second, I model the agent's preferences using cumulative prospect theory. This is in line with early representations of goals as reference points (Heath et al., 1999, Wu et al., 2008), but contrasts most approaches in the literature in which the agent's preference is modeled using disappointment models with the goal as reference point (Koch and Nafziger, 2011, 2016, Gomez-Minambres, 2012, Corgnet et al., 2015, Corgnet et al., 2018).[7] In the next section, it will become evident why such representation of preference along with the considered reference point rule can contradict the traditional justification for disappointment models. Third, in my model the goal divides the output space into gain and losses but enters the agent's utility as the transfer paid by the contract at that goal. This characterization of preference maintains the scale of the utility domain consistent. This approach contrasts that taken by Gomez-Minambres (2012), Corgnet et al. (2015), and Corgnet et al. (2018) in which both the production goal and a wage level enter the utility domain simultaneously.

4.3. Aspirations-based reference points

There is evidence that individuals, when making decisions under risk, are willing to take risks but also want to "earn at least something" be that a daily target income, not falling below a subsistence level, or achieving some earning level (Payne et al., 1980, 1981, Lopes, 1987). These outcomes are referred to as *aspirations* in the literature of decision theory.

[7] Specifically, it is well-known that expectation-based disappointment models and the Koszegi and Rabin (2006)'s choice acclimating equilibria would generate absurd predictions without the consumption utility component. When the goal acts as a reference point, this problem is absent. Thus, from a technical standpoint, it is not clear why most authors have favored this representation.

A way to model aspirations-based decision making is using cumulative prospect theory with the aspiration level acting as a reference point (Dalton et al., 2016a, Bogliacino and Ortoleva 2015). This means of modeling the agent's decision-making implies that outcomes below the aspiration level are classified as losses and that all outcomes above the reference point are classified as gains. If one is also ready to assume that aspirations are either exogenous to the principal's proposal, determined by salience-based judgements, or belong to the realm of goals, then the bonus contracts presented in Proposition 2, Corollary 5, and Proposition 3 suffice to optimally motivate these agents.

An alternative representation of aspirations-based decision making is due to Diecidue and van de Ven (2008). In that model, the decision-maker's utility exhibits a discontinuity around the aspiration level. Such discontinuity generates prominent utility differences between successes, potential outcomes above the aspiration level, and failures, possible outcomes below the aspiration level. The utility differences implied by this discontinuity explain the tendency of individuals to favor outcomes that just meet their aspiration level. A similar approach to model aspirations-based decision making is followed by Genicot and Ray (2017), arguably the most influential paper in economics on aspirations.

Appendix B.2 shows that the optimal incentive scheme when the agent's preferences are characterized by Diecidue and van de Ven (2008)'s preference is identical to that presented in Proposition 1. That is because the discontinuity in utility at the aspiration level does not imply changes in the agent's marginal utility as compared to the expected utility framework. As a consequence, the optimal implementation of incentives is identical under both frameworks. This result substantiates the importance of reference dependence and its implied irrationalities in generating optimal contracts that include bonuses.

5. Disappointment Aversion Models

5.1 The need for Disappointment models

A considerable bulk of the literature in behavioral economics is sympathetic with the idea that the individual's reference point is the expected value of an alternative. Potential outcomes included in that alternative will be thus classified as gains or losses relative to that expectation and, if the value function described in Assumption 4 is assumed, such classification has direct implications on the individual's risk preference.

This intuitive reference-point rule is however not compatible with cumulative prospect theory preferences. To see how, consider an agent with preferences described by (3), with an expectation-based reference point rule, and who is offered a constant contract $w = x$. Since $\mathbb{E}(x) = x$, the agent's utility is $CPT(e, w, \mathbb{E}(x)) = 0$ and such a utility level is attained for any x. An absurd implication.

The merit of the disappointment models is to account for expectations-based reference points without incurring in the aforementioned problem. They do so while maintaining the descriptively relevant phenomena of loss aversion and diminishing sensitivity. The remainder of this section, studies the optimal design of incentives under disappointment aversion.

5.2 Disappointment models with prior

In this section, I characterize the agent's preferences using the disappointment models of Bell (1985), Loomes and Sugden (1986), and Gul (1991). An agent with these preferences forms some prior about the possible transfers included in the contract, and experiences *disappointment* when the contract specifies a payment that is worse than such prior. In contrast, when the payment specified by the contract is better than his prior, the worker experiences *elation*. To avoid confusion and to keep a consistent terminology throughout the paper, I refer to the elation and disappointment outcomes as gains and losses, respectively.

The considered models differ with regard to how the prior is defined. For instance, in the model of Bell (1985) the agent's prior is the expected consumption value of the contract. Alternatively, in the model of Loomes and Sugden (1986) the prior corresponds to the expected consumption utility of the contract. I consider a preference specification in which these two models are equivalent. Specifically, I assume that the disappointment averse agent has preferences that can be characterized by

$$DA\big(e, w(y)\big) = \int_{\underline{y}}^{\overline{y}} w(y) f(y|e) dy + CPT(e, w(y), \widetilde{w}(y)) - c(e). \qquad (5)$$

Where $\widetilde{w} := \int_{\underline{y}}^{\overline{y}} w(y)f(y|e)dy$ and the expression $CPT(e, w(y), \widetilde{w}\,)$ refers to the preferences presented in equation (2) when the reference point is the expected consumption value of the contract, $r = \widetilde{w}$. Such representation of preference is adopted from Baillon et al. (2020).

The preferences described by equation (5) consist of three expressions. The first expression represents the agent's *absolute* evaluation of transfers. I refer to this expression as *consumption utility*. This expression guarantees that the agent does not incur in the absurd evaluation of degenerate lotteries described in section 4.1. The second expression captures loss aversion and diminishing sensitivity and is thus referred as the agent's psychological utility component. The last expression captures the standard cost of effort.

That the models of Bell (1985) and Loomes and Sugden (1986) are equivalent is due to the simplifying assumption that consumption utility is linear. This very same assumption allows me to express these two models as extensions of cumulative prospect theory, as it is evident from equation (5). This assumption is consistent with the observation by Köbberling and Wakker (2005) that consumption utility is the normative component of utility. Moreover, applications of disappointment aversion models impose linearity of the consumption utility (Gill and Prowse 2012, Liu and Shum, 2013).

When the agent's risk preferences can be represented by equation (5), the first-best and second-best contracts are similar to those presented in Proposition 2. Bonus contracts optimally insure and motivate the agent. The following proposition describes these contracts.

Proposition 4. *Under assumptions A1, A2 and A4, and that the agent's preferences are characterized by* (5)*, there exist unique output levels* $y_{m1}, y_{m2} \in \left(\underline{y}, \overline{y}\right)$ *such that*

i) *The first-best contract,* $w_{FB}(y)$*, is constant and pays the minimum possible if* $y < y_{m1}$*, exhibits a bonus at* $y = y_{m1}$*, and pays the constant transfer satisfying* $w_{FB}(y) = \widetilde{w} + h'\left(\frac{1}{v} - 1\right)$ *if* $y > y_{m1}$.

ii) *The second-best contract pays the minimum possible if* $y < y_{m2}$*, exhibits a bonus at* $y = y_{m2}$*, and increases in performance according to* $w_{SB} = \widetilde{w} + h'\left(\frac{1}{v+\gamma\left(1-\frac{f(y|e_L)}{f(y|e_H)}\right)} - 1\right)$ *if* $f\,y \geq y_{m2}$.

The first-best contract is similar to that described by Proposition 2 (i). This is not a surprising result given that the value function from Assumption 4 fully governs the agent's risk attitudes. That is because consumption utility, the component that is added to the preference representation described in (2) to account for disappointment models, is linear and does not affect the agent's evaluation of risky alternatives. Therefore, a contract offering full insurance to an agent with preferences described by (5) must have exhibit a similar shape to one offering insurance under prospect theory preferences. Optimal insurance involves exploiting the agent's loss aversion and diminishing sensitivity with a lump-sum bonus.

A similar rationale explains why the second-best contract resembles that presented in Proposition 2 (ii). The agent's risk preferences are fully determined by the value function of Assumption 4, so he still suffers from loss aversion and diminishing sensitivity. To optimally motivate him, the principal exploits these irrationalities using a contract that includes a bonus.

There are several differences between the contracts presented in Proposition 4 and Proposition 2 that are worth discussion. First, the reference point under disappointment aversion preferences is determined by $\tilde{w}$, the expected value of the contract. As a consequence, contracts with a higher expected value raise the agent's reference point, which, in turn, increase the size of the bonus that needs to be given to transition him to the domain of gains. Formally, the contracts from Proposition 4 offer in the domain of gains $w_{SB}(y) = \tilde{w} + \left(\frac{1}{v+\gamma\left(1-\frac{f(y|e_L)}{f(y|e_H)}\right)} - 1\right)$ and $w_{FB}(y) = \tilde{w} + h'\left(\frac{1}{v} - 1\right)$, both of which increase on $\tilde{w}$. However, that the bonus is raised implies that the principal can afford to extend the output segment in which the agent is exposed to losses without getting his offer rejected. In this way, the higher average transfers of an off-equilibrium contract are counteracted with more exposure to losses until the point in which the agent is just indifferent between accepting and rejecting it the contract. This endogenous adjustment of the bonus and its location due to a higher level of transfers reflects the comparative static presented in Corollary 2.

Second, the reference point rules adopted by prospect theory and disappointment aversion imply that the reference point of an agent can be different depending on the theory of risk used to characterize his preferences. Thus, to transition the agent to the domain of gains, a different bonus might be needed which implies that a threshold output is specified. In other words, y_{m1} and y_{m2} can be different from y_f and y_s.

Third, the inclusion of the consumption utility component in the agent's preference augments the incentives offered by the contracts in Proposition 4. For instance, the fixed transfer given by the first-best contract in the domain of gains surpasses the agent's reference point by an amount $h'\left(\frac{1}{\nu}-1\right)$. Instead, the first-best contract in Proposition 2 surpasses the agent's reference point by $h'\left(\frac{1}{\nu}\right)$. So, the transfer given in excess of the reference point depends positively on the marginal contribution of consumption utility, which is equal to one.

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 assumed to be his certainty equivalent. Importantly, this certainty equivalent includes the agent's psychological utility component. Formally, in equation (5) $\widetilde{w}$ can be replaced by CE, a degenerate lottery that generates $DA(e, w(y))) = CE$.

Appendix B.3 presents the optimal solution to the principal's contracting problem when the agent exhibits Gul (1991)'s preferences. Again, these contracts are similar to those presented in Proposition 2 due to the fact that the value function fully determines risk preference. Moreover, as previously mentioned, that the reference point can be different as compared to previously other disappointment models, implies that that the bonus and as a consequence the transition from gains and losses can be different as compared to Proposition 4. In fact, for a risk (globally) averse agent it must be that $y_{c1} < y_{m1}$ and $y_{c2} < y_{m2}$.[8]

De Meza and Webb (2007) also study the optimal contract under Gul (1991) preferences. My approach to represent preference is acutely different to theirs. The preference representation including in this paper incorporates diminishing sensitivity for both gains and losses, includes a consumption utility component for both gains and losses, and posits that the consumption

[8] Intuitively, a (globally) risk averse agent with risk preferences representable by (6), must exhibit $CE < \widetilde{w}_{SB}(y)$. 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, $CE = \widetilde{w}_{SB}(y)$. Hence, $y_{c1} < y_{m1}$ and $y_{c2} < y_{m2}$.

utility component is linear. These differences lead to considerably different solutions. In particular, their optimal contract does not include a bonus and performance insensitive region located at intermediate output levels, rather than at low output levels. This dissimilarity shows how slight differences in preferences can lead to acutely different incentive contracts.

5.3 The contract as the reference point

In this section, I assume that the agent exhibits risk preferences that can be represented by the disappointment model of Delquié and Cillo (2006) and Kőszegi and Rabin (2006, 2007)'s choice acclimating equilibria. This is arguably the most used model in economics to characterize reference-dependent preferences. This type of disappointment model differs from those analyzed in the previous section as it posits that the agent's reference point is each possible outcome of the contract. Hence, the agent does not need to form a prior in order to define a reference point.

To exemplify this rule, suppose that the worker obtains the contract $w_3 = (0.5,200; 0.5,100)$. In these models, each possible outcome of the contract is taken as the reference point. First, let $r = 200$. Then, obtaining 100 feels like a loss that realizes with 25% probability (0.50*0.50) while obtaining 200 is a neutral outcome which also realizes with 25% probability. Instead when $r = 100$, obtaining 200 feels like a gain realizing with 25% chance and obtaining 100 is a neutral outcome. The contract w_3 is thus reframed as a 25% chance to win 100, a 25% chance to loss 100, and a 50% chance to obtain a neutral outcome.

The reframing of lotteries depicted in the aforementioned example illustrates the need of a standard consumption utility component in the representation of preference that accompanies such reference point rule. Otherwise a problem similar to that explained in Section 5.1 emerges.[9] Hence, the agent's risk preference when each possible outcome of the contract is taken as the reference point is given by

[9] As in section 5.1. let $w = x$ and assume that the agent exhibits prospect theory preferences. When the contract itself is taken as the reference point, then $CPT(e, w, w) = 0$ for any $x > 0$. An absurd implication.

$$KR\big(e, w(y)\big) = \int_{\underline{y}}^{\overline{y}} w(y) f(y|e) dy + \eta \int_{\tilde{y}}^{\overline{y}} \int_{\underline{y}}^{\overline{y}} u\big(w(y) - w(\tilde{y})\big) f(\tilde{y}|e) f(y|e) d\tilde{y} dy$$
$$- \eta\lambda \int_{\underline{y}}^{\tilde{y}} \int_{\underline{y}}^{\overline{y}} u\big(w(\tilde{y}) - w(y)\big) f(y|e)\, f(\tilde{y}|e)\, d\tilde{y} dy - c(e), \qquad (6)$$

where $\eta > 0$ weighs the gain-loss utility component in the agent's overall utility. Equation (6) can be rewritten as

$$KR\big(e, w(y)\big) = \int_{\underline{y}}^{\overline{y}} w(y) f(y|e) dy + \eta \int_{\underline{y}}^{\overline{y}} CPT(e, w(y), w(\tilde{y}))\, f(\tilde{y}|e) d\tilde{y}. \qquad (7)$$

Where $CPT(e, w(y), w(\tilde{y}))$ refers to the preferences described by equation (2) when the contract's outcome $w(\tilde{y})$ for a given output level $\tilde{y}$ is taken as the agent's reference point.

That the models of Delquié and Cillo (2006)'s and Kőszegi and Rabin (2006, 2007) choice acclimating equilibria can be treated as extensions of cumulative prospect theory, as in equation (8), is due to the simplifying assumption that consumption utility is linear. This is the approach followed by Baillon et al. (2020) in their general representation of reference dependence. Also, this assumption is typically made in applications of this type of preferences (Heidhues and Kőszegi 2008, Abeler et al. 2011, Crawford and Meng 2011, Eil and Lien 2014).

The following proposition shows that the solution to the principal's problem when the agent's preferences are characterized by (8) consists of a stochastic contract. This property makes this solution acutely different from those presented in Propositions 2-4.

Proposition 5. *Under assumptions A1, A2, A4, and that the agent's preferences are characterized by* (7)*, there exist unique output levels* $y_{k1}, y_{k2} \in [\underline{y}, \overline{y}]$ *such that:*

i) *The first-best contract, consists of a stochastic contract* $L^* = (w_{FB}(\, y_{k1})\,, p_1^*; 0)$ *with* $p_1^* = \frac{1}{2} + \frac{w_{FB}(\, y_{k1})}{2\eta\left(u(w_{FB}(\, y_{k1})\,) - \lambda u(-w_{FB}(\, y_{k1}))\right)}$ *if* $y < y_{k1}$ *and a constant payment,* $w_{FB}(y)$*, satisfying* $\left(1 + \eta \int_{\underline{y}}^{\overline{y}} u'\big((w_{FB}(y) - w_{FB}(y_{k1}))\big) f(\tilde{y}|e) d\tilde{y}\right)^{-1} = v$ *if* $y \geq y_{k1}$.

ii) The second-best contract, $w_{SB}(y)$*, consists of a stochastic contract* $L^* = (w_{SB}(y_{k2}), p_2^*; 0)$ *with* $p_2^* = \frac{1}{2} + \frac{w_{SB}(y_{k2})}{2\eta\left(-\lambda u(-w_{SB}(y_{k2})) + u(w_{SB}(y_{k2}))\right)}$ *if* $y < y_{k2}$ *and a payment scheme that increases in performance if* $y \geq y_{k2}$.

The principal provides optimal insurance with a contract composed of two segments. For low output realizations, i.e. lower than a threshold y_{k1}, the agent is exposed to risk with a lottery-like compensation. This way of providing insurance substantially differs to the first-best contracts presented in Propositions 1-4 and, as it will be explained below, is a direct consequence of considered reference point rule. Instead, for high output levels, i.e. higher than a threshold y_{k1}, the contract is constant in performance, resembling more standard ways of insurance.

To understand how the assumed preference representation generates the stochastic component included in the contract, suppose that $w = 0$, the minimum admissible transfer, is paid when $y < y_{k1}$. This strategy, followed by the principal in the contracts described in Propositions 2-3, leads to greater disutility as compared to the hypothetical case in which the agent's reference point was either exogenous or, say, the mean value of the contract. That is because this low transfer is eventually taken as the reference point, so for any w_{FB} and r such that $r > w_{FB} > 0$, then $-\lambda\, u(-w_{FB}) > -\lambda\, u(r - w_{FB})$. To avoid exposing the agent to such sizeable losses that could lead the agent to reject the contract, the principal offers the lottery L^* including a non-zero probability to locate the agent in the domain of gains.

To further elucidate the above intuition, suppose that $u'' = 0$ and $\eta = 1$. Under these conditions, the probability governing the lottery-like payment becomes $p_1^* = \frac{1}{2} + \frac{1}{2(\lambda+1)}$. Note that for any $\lambda > 1$ the contract is stochastic inasmuch as $p_1^* \in \left(\frac{1}{2}, \frac{3}{4}\right)$. More importantly, note that the higher the agent's loss aversion, as reflected by λ, the higher becomes p_1^* and the agent is less exposed to potential losses included by the contract.

The second-best contract described by Proposition 5 also consists of a stochastic component that applies at low output levels, i.e. lower than a threshold y_{k2}. However, in contrast to the first-best contract, that lottery-like payment is complemented with an incentive scheme in which transfers increase in the agent's performance. This more traditional shape of the incentive scheme applies at high output levels, i.e. higher than a threshold $\tilde{y}_{k2}$.

With this two-segment contract the principal motivates the agent in two different ways. First, the deterministic monetary incentives that increase in performance motivate the agent in the classical way. More interestingly, the lottery-like payment motivates the agent to exert high effort by exploiting the agent's irrationalities. The intuition of this result is similar to that given in Propositions 2-4. Specifically, the loss averse agent will be incentivized to choose high effort in order to avoid the loss implied by the lowest outcome of L^*. Similarly, diminishing sensitivity implies that the agent will be risk seeking and thus willing to exert high effort when facing the potential loss implied by the lowest outcome of L^*

In a well-known paper, Herweg et al. (2010) also find that, under the considered preference representation, the second-best contract is stochastic. However, in stark contrast to Proposition 5 they find that this result only holds for high levels of loss aversion, i.e. $\lambda > 2$. Instead, for low loss aversion levels, i.e. $\lambda < 2$, their contract becomes a lump-sum bonus. This result can be captured by Proposition 5 but, however, under a different utility specification.

Let $x := w(y) - w(g)$. Replace the global concavity in utility from Assumption 4, i.e. $u''(x) < 0$ for all x, for $u''(x) < 0$ if $x \geq 0$ and $u''(x) > 0$ if $x < 0$. An assumption made in Kőszegi and Rabin (2006, 2007).[10] This assumption is naturally accompanied with losses entering positively in the agent's utility, so instead of having $-u(-x)$ for $x < 0$ as assumed throughout the paper, such a loss enters the utility as $u(-x)$.

After making such an assumptions substitution, receiving a lottery $L = (w_{SB}(y_{k2}), 1-p; 0, p)$ generates utility:

$$KR(L) = (1-p)w_{SB}(\tilde{y}) + p(1-p)\eta\left(u(w_{SB}(y_{k2})) + \lambda u(-w_{SB}(y_{k2}))\right). \quad (8)$$

The first-order condition of (9) with respect to p is

$$-w_{FB}(\tilde{y}) + (1-2p)\eta\left(u(w_{SB}(y_{k2})) + \lambda u(-w_{SB}(y_{k2}))\right) = 0, \quad (9)$$

[10] See their A3 in both papers.

which after some rearranging leads to $p_2^* = \frac{1}{2} - \frac{w_{SB}(\tilde{y})}{2\eta\left(-\lambda u(-w_{SB}(\tilde{y}))+u(w_{SB}(\tilde{y}))\right)}$.[11] Assuming, as Herweg et al. (2010) do, that $\eta = 1$ and $u'' = 0$, the optimal probability collapses to $p_2^* = \frac{1}{2} + \frac{1}{2(\lambda-1)}$. Hence, if $\lambda \in (1,2]$ the contract indeed becomes a lump-sum bonus since $p_2^* = 1$ and the lowest admissible pay is given if $y < \tilde{y}_{k2}$ while a larger fixed payment is given if $y \geq \tilde{y}_{k2}$. Instead, if $\lambda > 2$, the optimal contract is stochastic. This solution thus emerges as a consequence of losses entering positively in the agent's utility and linearity of the utility function.

All in all, the model of Delquié and Cillo (2006) and Kőszegi and Rabin (2006, 2007) generates a solution to the principal's problem that includes a stochastic component. This solution starkly contrasts the contracts of Proposition 2-4. Furthermore, it was shown that under some changes in the agent's utility, that reflect the model of Kőszegi and Rabin (2006, 2007), bonus contracts similar to those presented in Proposition 2 and 4 emerge as the first-best and second-best solutions if one is ready to assume a mild degree of loss aversion and no diminishing sensitivity.

6. Extensions

Thus, far it has been shown that bonus contracts similar to those presented in Proposition 2 emerge as optimal solutions under different preference representations and reference point rules. This section extends the results presented in Section 2 to further gain generalizability and highlight the significance of Proposition 2.

6.1 CPT principal

Throughout the paper it was assumed that the principal was an expected value maximizer. A natural extension is to consider a setting in which she, just like the agent, exhibits reference-dependent preferences. If that was the case, the principal's preferences are characterized by a value function with properties similar to those presented in Assumption 4. Formally, let the principal's objective function be

---

[11] The second-order condition is $-2\eta\left(u(w_{SB}(y_{k2})) + \lambda u(-w_{SB}(y_{k2}))\right)$, which is negative since $u'(\cdot) > 0$. Therefore, the probability given by p_2^* is optimal.

$$\Pi\big(S(y), r_p, w\big) = \begin{cases} P\big(S(y) - r_p\big) - w(y) & if\ S(y) \geq r_p\ , \\ -\lambda_p P\left(r_p - S(y)\right) - w(y) & if\ S(y) < r_p. \end{cases} \quad (10)$$

Where $r_p \geq 0, \lambda_p > 1, P'(\cdot) > 0$, and $P''(\cdot) < 0$.

This assumption together with the assumption that the agent's preference can be represented by equation (2) imply that the contracts in Proposition 2 remain to be optimal. This is because the principal's loss aversion and diminishing sensitivity do not apply to her cost component, $w(y)$, but to her benefit. This assumption is consistent with the approach taken throughout the paper to model reference dependence for the agent.

Instead, assume that the principal's irrationalities apply to both her benefit and cost. This assumption is consistent with the idea that these biases apply to monetary outcomes, in her case both benefits and costs. Specifically, let the principal's preferences be given by

$$\Pi\big(S(y), r_p, w\big) = \begin{cases} S(y) - r_p - w(y) & if\ S(y) \geq r_p + w(y), \\ -\lambda_p \left(r_p + w(y) - S(y)\right) & if\ S(y) < r_p + w(y). \end{cases} \quad (11)$$

Hence, unlike the agent, the principal, only suffers from loss aversion. This assumption is imposed for simplicity but can be justified on the grounds of the principal being able to pool multiple risks and as a result not exhibiting utility curvature. When focusing on the more general case in which effort is not observable nor contractable, the solution to the principal's program when she exhibits preferences as in (11) is presented in the following proposition. The proofs of the most important results of this section are relegated to Appendix C.

Proposition 6. *Let $\hat{y}_s \in (\underline{y}, \overline{y})$ be the unique output level from Proposition 2. Under assumptions A1, A2, A4, that the agent's preferences are as described by (2), and that the principal's preferences are given by (11), there exists a unique output level $\hat{y}_p \in \left[\underline{y}, \overline{y}\right]$ such that the second-best contract:*

i) Is identical to the contract presented in Proposition 2 (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 $\hat{y}_p \geq \hat{y}_s$.

When the principal is in the domain of gains, $S(y) \geq r_p + w(y)$, the principal's marginal cost, associated to $w(y)$, is identical to that faced in Section 2. As a result, the solution to the principal's problem is exactly that presented in Proposition 2 (ii).

Instead, in the domain of losses, $S(y) < r_p + w(y)$, the principal is more risk averse than in the domain of gains due to her loss aversion. If the agent is also the domain of losses, the principal can transfer all risk to him by offering the lowest possible payment at low output levels. Recall that the agent would accept that proposal due to the risk seeking from diminishing sensitivity. This excessively low payment component is also included in the second-best contract of Proposition 2 (ii). Instead, if the agent is in the domain of gains, the principal needs to insure the agent from risk. However, due to her loss aversion the principal provides a less generous insurance as compared to Proposition 2 (ii). In particular, incentives will be less higher-powered because the principal is willing to expose the agent to more risk. This different ways of insuring the agent when he is in the domain of gains leads to a kink in the incentives scheme at $\hat{y}_p$, as shown by the second part of the proposition.

6.2 Adverse selection followed by moral hazard

The assumption that the principal is informed about the agent's risk preferences is typically made in models of moral hazard. However, this assumption becomes more prominent and stringent in the framework of this paper. This extension considers a setting in which the principal does not exactly know the agent's risk preferences.

Suppose that the principal is informed about the utility shape of the agent, as it is typically done in the literature, but that the degree of loss aversion is unknown to her. Specifically, assume that she can contract with agents having either high or low degrees of loss aversion, $\lambda_i \in \{\lambda_L, \lambda_H\}$ where $\lambda_H > \lambda_L$ and $\lambda_L > 1$. Also, the principal knows that contracting with an agent with λ_H takes place with probability ω, while contracting with an agent with λ_L happens with the complement probability, $1 - \omega$.

The exact timing of the interaction between agent and principal is described next for clarity. First, nature moves and determines λ_i which is private information to the agent. Second, the principal offers a menu of contracts $w(y)^i$ with $i = \{L, H\}$ that seeks to screen agents according to their level of loss aversion. Third, the agent selects a contract. Fourth, e is chosen by the agent. Finally, y realizes and the agent is paid according to the transfers specified in the selected contract. In this framework, the principal's objective is to offer a contract that both motivates the agent to exert high effort and enables her to screen agents according to their degree of loss aversion.

The following proposition shows that the principal offers a menu of contracts $w(y)^i$ in which the contract of Proposition 2 (ii) is complemented with informational rents that enable the agent to perform a screening of agents according to their levels of loss aversion.

Proposition 7. *Under assumptions A1, A2, A4, that the agents' preferences are described by* (2)*, and that* λ_i *is unknown to the principal, the optimal menu consists of:*

i) $w(y)^H$ *which exhibits the shape of the contract described in Proposition 2 (ii) and satisfies* $CPT(e_H, w(y)^H, r, \lambda_H) = \overline{U}$*, and*

ii) $w_{SB}(y)^L$ *which exhibits the shape of the contract described in Proposition 2 (ii) and satisfies* $CPT(e_H, w(y)_L, r, \lambda_L) = \overline{U} + (\lambda_H - \lambda_L) \int_{\underline{y}}^{\hat{y}_H} u(r - w(y)^H) f(y|e_H)\, \mathrm{d}y$.

Where $\hat{y}_H, \hat{y}_L \in \left(\underline{y}, \overline{y}\right)$ *are the points at which* $w(y)^H$*and* $w(y)^L$*, respectively, transition the agent from losses to gains.*

The solution to the principal's problem consists of a menu of contracts with the shape of the second-best contract in Proposition 2 (ii), i.e. bonus contracts. Such a shape guarantees that both types exert high effort at a minimal cost for the principal.

Moreover, to screen among agents with different levels of loss aversion, the principal offers informational rents to the more *efficient* types, that is agents for who the irrationality of loss aversion is less severe. Specifically, to deter the agent with low loss aversion from mimicking the agent with higher loss aversion the principal includes in his contract a rent $(\lambda_H - \lambda_L) \int_{\underline{y}}^{\hat{y}_H} u(r - w(y)^H) f(y|e_H)\, \mathrm{d}y$. This rent increases in $\lambda_H - \lambda_L$, the difference at which loss aversion differs across the two agents. In this way, that agent is indifferent between incurring in a strategy of mimicking or not. Also, note that the agent with high loss aversion is not willing to mimic the agent with lower loss aversion, since, as suggested by Corollary 3, such a contract would expose him to sizable losses that generate large degrees of disutility.

6.3 Ambiguity

The assumption that principal and agent know the probabilities associated to different output levels can be relaxed. In this setting, both agent and principal know that performance may take any value in the set $[\underline{y}, \overline{y}]$ but they do not know the exact distribution that underlies the realizations of the random variable y.

A first approach is to assume probability sophistication (Machina and Schmeidler, 1992). According to this assumption, the agent quantifies uncertainties as probabilities. Thus, there exists a probability measure P on $[\underline{y}, \overline{y}]$ such that $w(y)$ is evaluated by the agent as a probability-contingent pay schedule. The objective function of the agent with prospect theory preference and probabilistic sophistication becomes

$$CPT(e, w(y), r) := \int_{\underline{y}}^{\overline{y}} \theta u(w(y) - r) - \lambda(1-\theta)\big(r - w(y)\big)\mathrm{d}P(y) - c(e), \quad (12)$$

Under (13), the solution to the maximization problem of the agent is similar to those presented in Proposition 2 with the difference that probabilities assigned to output are now subjective.

The condition of probabilistic sophistication is stringent and has been invalidated by robust empirical phenomena such as the ``home bias" or the Ellsberg paradox. A less stringent condition made in non-EU models is that probabilistic sophistication holds within sources of uncertainty, but not necessarily between sources of uncertainty (Chew and Sagi, 2008, Abdellaoui et al., 2011, Baillon et al. 2018). In the present setting, this assumption amounts to posit a uniform degree of ambiguity for the source that determines performance. If one is ready to assume that the level of the agent's effort does not alter the uniformity of the ambiguity from which performance originates, probabilistic sophistication holds within the source and the optimal contracts are again as described in Proposition 2.

7. Conclusion

This paper provided a preference foundation for the conventional compensation practice of offering bonuses. A contract with a bonus exploits the agent’s loss aversion and diminishing sensitivity in a way that allows the principal to offer insurance and generate motivation in a cost effective-way. I also demonstrated that regardless of the theory of risk chosen to characterize reference-dependent preferences or the rule chosen to define a reference point, the bonus feature of the contract emerges as optimal solution to the principal’s problem.

Recent and robust empirical evidence on the type of preferences that individuals exhibit in settings of risk has emerged. For instance, according to Baillon et al. (2020) most subjects' preferences can be characterized by prospect theory with a status-quo or max-min rule as reference points. The results of this paper provide a sharp prediction in light of these findings: Contracts including bonuses should be widely used by organizations in practice. That prediction is corroborated by compensation practices around the world. For instance, according to Worldatwork (2018) 90% of American enterprises use incentive schemes that include a bonus.

References

Abdellaoui, M., Baillon, A., Placido, L., & Wakker, P. P. (2011). The rich domain of uncertainty: Source functions and their experimental implementation. *American Economic Review*, *101*(2), 695-723.

Abdellaoui, M., Bleichrodt, H., & l'Haridon, O. (2008). A tractable method to measure utility and loss aversion under prospect theory. *Journal of Risk and uncertainty*, *36*(3), 245-266.

Abdellaoui, M., Bleichrodt, H., & Paraschiv, C. (2007). Loss aversion under prospect theory: A parameter-free measurement. *Management Science*, *53*(10), 1659-1674.

Abeler, J., Falk, A., Goette, L., & Huffman, D. (2011). Reference points and effort provision. *American Economic Review*, *101*(2), 470-92.

Allen, E. J., Dechow, P. M., Pope, D. G., & Wu, G. (2017). Reference-dependent preferences: Evidence from marathon runners. *Management Science*, *63*(6), 1657-1672.

Baillon, A., Bleichrodt, H., & Spinu, V. (2020). Searching for the reference point. *Management Science*, *66*(1), 93-112.

Baillon, A., Huang, Z., Selim, A., & Wakker, P. P. (2018). Measuring ambiguity attitudes for all (natural) events. *Econometrica*, *86*(5), 1839-1858.

Bell, D. E. (1985). Disappointment in decision making under uncertainty. *Operations research*, *33*(1), 1-27.

Bogliacino, F., & Ortoleva, P. (2015). The behavior of other as a reference point. *Available at SSRN 2653343*.

Borch K.H.. (1962). Equilibreium in a Reinsurance Market. *Econometrica, 30*, 424-444.

Brookins, P., Goerg, S., & Kube, S. (2017). Self-chosen goals, incentives, and effort. *Unpublished manuscript*.

Crawford, V. P., & Meng, J. (2011). New york city cab drivers' labor supply revisited: Reference-dependent preferences with rational-expectations targets for hours and income. *American Economic Review*, *101*(5), 1912-32.

Chew, S. H., & Sagi, J. S. (2008). Small worlds: Modeling attitudes toward sources of uncertainty. *Journal of Economic Theory*, *139*(1), 1-24.

Chung, D. J., Steenburgh, T., & Sudhir, K. (2014). Do bonuses enhance sales productivity? A dynamic structural analysis of bonus-based compensation plans. *Marketing Science*, *33*(2), 165-187.

Clark, D., Gill, D., Prowse, V., & Rush, M. (2020). Using goals to motivate college students: Theory and evidence from field experiments. *Review of Economics and Statistics*, *102*(4), 648-663.

Corgnet, B., Gómez-Miñambres, J., & Hernán-Gonzalez, R. (2015). Goal setting and monetary incentives: When large stakes are not enough. *Management Science*, *61*(12), 2926-2944

Corgnet, B., Gómez-Miñambres, J., & Hernán-Gonzalez, R. (2018). Goal setting in the principal–agent model: Weak incentives for strong performance. *Games and Economic Behavior*, *109*, 311-326.

Corgnet, B., & Hernán-González, R. (2019). Revisiting the trade-off between risk and incentives: The shocking effect of random shocks?. *Management Science*, *65*(3), 1096-1114.

Dalton, P. S., Ghosal, S., & Mani, A. (2016)a. Poverty and aspirations failure. *The Economic Journal*, *126*(590), 165-188.

Dalton, P. S., Gonzalez, V., & Noussair, C. N. (2016)b. Self-chosen goals: Incentives and gender differences.

De Meza, D., & Webb, D. C. (2007). Incentive design under loss aversion. *Journal of the European Economic Association*, *5*(1), 66-92.

Delquié, P., & Cillo, A. (2006). Disappointment without prior expectation: a unifying perspective on decision under risk. *Journal of Risk and Uncertainty*, *33*(3), 197-215.

Diecidue, E., & Van De Ven, J. (2008). Aspiration level, probability of success and failure, and expected utility. *International Economic Review*, *49*(2), 683-700.

Dittmann, I., Maug, E., & Spalt, O. (2010). Sticks or carrots? Optimal CEO compensation when managers are loss averse. *The Journal of Finance*, *65*(6), 2015-2050.

Eil, D., & Lien, J. W. (2014). Staying ahead and getting even: Risk attitudes of experienced poker players. *Games and Economic Behavior*, *87*, 50-69.

Genicot, G., & Ray, D. (2017). Aspirations and inequality. *Econometrica*, *85*(2), 489-519.

Gill, D., & Prowse, V. (2012). A structural analysis of disappointment aversion in a real effort competition. *American Economic Review*, *102*(1), 469-503.

Gneezy, U., Goette, L., Sprenger, C., & Zimmermann, F. (2017). The limits of expectations-based reference dependence. *Journal of the European Economic Association*, *15*(4), 861-876.

González-Jiménez, V, Castillo G. (2021) Optimal incentives without expected utility. Unpublished.

Gómez-Miñambres, J. (2012). Motivation through goal setting. *Journal of Economic Psychology*, *33*(6), 1223-1239.

Gul, F. (1991). A theory of disappointment aversion. *Econometrica: Journal of the Econometric Society*, 667-686.

Harding, M., & Hsiaw, A. (2014). Goal setting and energy conservation. *Journal of Economic Behavior & Organization*, *107*, 209-227.

Heath, C., Larrick, R. P., & Wu, G. (1999). Goals as reference points. *Cognitive psychology*, *38*(1), 79-109.

Hershey, J. C., & Schoemaker, P. J. (1985). Probability versus certainty equivalence methods in utility measurement: Are they equivalent?. *Management Science*, *31*(10), 1213-1231

Herweg, F., Müller, D., & Weinschenk, P. (2010). Binary payment schemes: Moral hazard and loss aversion. *American Economic Review*, *100*(5), 2451-77.

Heidhues, P., & Kőszegi, B. (2008). Competition and price variation when consumers are loss averse. *American Economic Review*, *98*(4), 1245-68.

Holmström, B. (1979). Moral hazard and observability. *The Bell journal of economics*, 74-91.

Hsiaw, A. (2013). Goal-setting and self-control. *Journal of Economic Theory*, *148*(2), 601-626.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263–291.

Kim, S. K. (1997). Limited liability and bonus contracts. *Journal of Economics & Management Strategy*, *6*(4), 899-913.

Koch, A. K., & Nafziger, J. (2011). Self-regulation through goal setting. *Scandinavian Journal of Economics*, *113*(1), 212-227.

Koch, A. K., & Nafziger, J. (2016). Goals and bracketing under mental accounting. *Journal of Economic Theory*, *162*, 305-351.

Köbberling, V., & Wakker, P. P. (2005). An index of loss aversion. *Journal of Economic Theory*, *122*(1), 119-131.

Kőszegi, B., & Rabin, M. (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics*, *121*(4), 1133-1165.

Kőszegi, B., & Rabin, M. (2007). Reference-dependent risk attitudes. *American Economic Review*, *97*(4), 1047-1073.

Larrick, R. P., Heath, C., & Wu, G. (2009). Goal-induced risk taking in negotiation and decision making. *Social Cognition*, *27*(3), 342-364.

Levy, H., & Levy, M. (2009). The safety-first expected utility model: Experimental evidence and economic implications. *Journal of Banking & Finance*, *33*(8), 1494-1506.

Liu, Q., & Shum, S. (2013). Pricing and capacity rationing with customer disappointment aversion. *Production and Operations Management*, *22*(5), 1269-1286.

Loomes, G., & Sugden, R. (1986). Disappointment and dynamic consistency in choice under uncertainty. *The Review of Economic Studies*, *53*(2), 271-282.

Lopes, L. L. (1987). Between hope and fear: The psychology of risk. In *Advances in experimental social psychology* (Vol. 20, pp. 255-295). Academic Press.

Machina, M. J., & Schmeidler, D. (1992). A more robust definition of subjective probability. *Econometrica: Journal of the Econometric Society*, 745-780.

Markle, A., Wu, G., White, R., & Sackett, A. (2018). Goals as reference points in marathon running: A novel test of reference dependence. *Journal of Risk and Uncertainty*, *56*(1), 19-50.

Mirrlees, J. A. (1976). The optimal structure of incentives and authority within an organization. *The Bell Journal of Economics*, 105-131.

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

Park, E. S. (1995). Incentive contracting under limited liability. *Journal of Economics & Management Strategy*, *4*(3), 477-490.

Payne, J. W., Laughhunn, D. J., & Crum, R. (1980). Translation of gambles and aspiration level effects in risky choice behavior. *Management Science*, *26*(10), 1039-1060.

Payne, J. W., Laughhunn, D. J., & Crum, R. (1981). Note—further tests of aspiration level effects in risky choice behavior. *Management Science*, *27*(8), 953-958.

Rabin, M. (2000). Risk aversion and expected-utility theory: A calibration theorem. *Econometrica*, *68*(5), 1281-1281.

Sprenger, C. (2015). An endowment effect for risk: Experimental tests of stochastic reference points. *Journal of Political Economy*, *123*(6), 1456-1499.

Terzi, A., Koedijk, K., Noussair, C. N., & Pownall, R. (2016). Reference point heterogeneity. *Frontiers in psychology*, *7*, 1347.

Tiefenbeck, V., Goette, L., Degen, K., Tasic, V., Fleisch, E., Lalive, R., & Staake, T. (2018). Overcoming salience bias: How real-time feedback fosters resource conservation. *Management science*, *64*(3), 1458-1476.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, *5*(4), 297-323.

Vieider, F. M., Lefebvre, M., Bouchouicha, R., Chmura, T., Hakimov, R., Krawczyk, M., & Martinsson, P. (2015). Common components of risk and uncertainty attitudes across contexts and domains: Evidence from 30 countries. *Journal of the European Economic Association*, *13*(3), 421-452.

Von Gaudecker, H. M., Van Soest, A., & Wengstrom, E. (2011). Heterogeneity in risky choice behavior in a broad population. *American Economic Review*, *101*(2), 664-94.

Worldatwork. (2018). *2018 Incentive Pay Practices for Publicly-Traded Companies.*

Wu, G., Heath, C., & Larrick, R. (2008). A prospect theory model of goal behavior. *Unpublished manuscript*.

Appendix A. Proofs

Proposition 1

See Holmström (1979) and Borch (1962)

Proposition 2.

i) Lagrangian and first-order conditions.

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

$$\mathcal{L} = \big(S(y) - w(y)\big)f(y|e_H)$$
$$+\nu\big(\theta u(w(y) - r)f(y|e_H) - \lambda(1-\theta)u\big(r - w(y)\big)f(y|e_H) - c - \overline{U}\big) \qquad (A1)$$
$$+\gamma\big(\theta u(w(y) - r)\big(f(y|e_H) - f(y|e_L)\big) - \lambda(1-\theta)u\big(r - w(y)\big)\big(f(y|e_H) - f(y|e_L)\big) - c\big).$$

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

$$-f(y|e_H) + \nu\left(\theta u'(w(y) - r)f(y|e_H) + \lambda(1-\theta)u'\left(r - w(y)\right)f(y|e_H)\right) \qquad (A2)$$
$$+\gamma\left(\theta u'(w(y) - r)\big(f(y|e_H) - f(y|e_L)\big) + \lambda(1-\theta)u'\big(r - w(y)\big)\big(f(y|e_H) - f(y|e_L)\big)\right) = 0.$$

Denoting by $w_{SB}^{fo}(y)$ the transfer satisfying $(A2)$, the following expressions are obtained after some manipulations:

$$\frac{1}{u'\big(w_{SB}^{fo}(y) - r\big)} = \nu + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right), \qquad (A3)$$

if $\theta = 1$, and

$$\frac{1}{\lambda u'\left(r - w_{SB}^{fo}(y)\right)} = \nu + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right), \qquad (A4)$$

if $\theta = 0$. Using $h(\cdot)$ from A4 rewrite $(A3)$ as:

$$w_{SB}^{fo}(y) = r + h'\left(\frac{1}{\nu + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right)}\right), \qquad (A5)$$

and $(A4)$ as:

$$w_{SB}^{fo}(y) \; = r - h'\left(\frac{1}{\lambda\left(v+\gamma\left(1-\frac{f(y|e_L)}{f(y|e_H)}\right)\right)}\right). \qquad (A6)$$

ii) Suboptimality of solution from the first-order approach in losses.

If $w_{SB}^{fo}(y)$ satisfying $(A6)$ exhibits $0 < w_{SB}(y) < r$, the principal is better off offering a lottery payment $L = (r, p; 0, 1-p)$ with $p \in [0,1]$. Since the agent's utility in losses, $-\lambda u\left(r - w_{SB}^{fo}(y)\right)$, increases in $w_{SB}^{fo}(y)$, there must exist a number $p \in (0,1)$ such that:

$$-\lambda u\left(r - w_{SB}^{fo}(y)\right) = -(1-p)\lambda u(r). \quad (A7)$$

Therefore, replacing $w_{SB}^{fo}(y)$ from $(A6)$ by a lottery-like contract $L = (r, p; 0, 1-p)$ leaves the agent's participation constraint and the incentive compatibility constraint unchanged. Moreover, the convexity of $-u(\cdot)$ in the domain of losses (Assumption 4), implies:

$$-\lambda u\left(r - w_{SB}^{fo}(y)\right) < -\lambda u\big((1-p)r\big). \quad (A8)$$

Since $u(\cdot)$ increases everywhere (Assumption 4), Equation $(A8)$ implies

$$w_{SB}(y) > pr. \qquad (A9)$$

Hence, the lottery $L = (r, p; 0, 1-p)$ is more cost-effective for the principal than the candidate solution given by $(A6)$.

I turn to analyze the marginal incentives of offering $L = (r, p; 0, 1-p)$. Denote by $\overline{L} = pr$ its expected value. Substituting $\overline{L}$ into the agent's expected utility in losses yields:

$$\mathbb{E}\big(CPT(L)\big) = -\lambda\left(1 - \frac{\overline{L}}{r}\right)\big(u(r)\big) - c, \qquad (A10)$$

which is linear in $\overline{L}$. As a result, changes in $\overline{L}$, through adjustments of p, do not alter the agent's marginal expected utility.

Let $\hat{y}_s \in [\underline{y}, \overline{y}]$ be the performance level satisfying:

$$\frac{1}{\frac{\lambda u(r)}{r}} = v + \gamma\left(1 - \frac{f(\hat{y}_s|e_L)}{f(\hat{y}_s|e_H)}\right), \qquad (A11)$$

The uniqueness of $\hat{y}_s$ is guaranteed by the fact that the right-hand side of $(A11)$ is strictly increasing in performance due to Assumption 2, while the left-hand side of that equation is constant. Hence, if these expressions cross at any $\hat{y}_s \in (\underline{y}, \overline{y})$, that crossing must be unique.

Instead, failing to cross at any interior output level yields $\hat{y}_s = \underline{y}$ if $\frac{1}{\frac{\lambda u(r)}{r}} > \nu + \gamma\left(1 - \frac{f(\hat{y}_s|e_L)}{f(\hat{y}_s|e_H)}\right)$, or $\hat{y}_s = \overline{y}$, otherwise.

For performance levels $y < \hat{y}_s$ the expected value of L should be reduced. As established by equation $(A10)$, such a change does not alter the agent's marginal utility so $\overline{L}$ can be reduced all the way to zero by setting $p = 0$. Instead, for any $y > \hat{y}_s$ the expected value of L should be increased. Again, since $\mathbb{E}(CPT(L))$ is linear in $\overline{L}$, the expected value of L should be increased all the way to r by setting $p = 1$, bringing the agent to the domain of gains. Therefore, in $y \leq \hat{y}_s$ the agent is located in the domain of losses with the following incentives scheme:

$$w_{SB}(y) = \begin{cases} 0 \; if \; y < \hat{y}_s, \\ L \; if \; y = \hat{y}_s. \end{cases} \qquad (A12)$$

Instead, for $y > \hat{y}_s$ the optimal incentives scheme satisfies $(A5)$ so $w_{SB}(y) = w_{SB}^{fo}$ in that segment.

iii) Shape of the second-best contract

By $(A12)$ it is clear that the optimal contract is performance insensitive if $y < \hat{y}_s$. That equation also shows that at $y = \hat{y}$ the agent is indifferent between getting $w_{SB} = 0, w_{SB} = r$, or any convex combination between those two performance-insensitive points. Alternatively, for $y > \hat{y}$, the optimal solution is given by $(A5)$ which is evidently larger than r. Therefore, the solution to the principal's problem exhibits a discrete jump at $y = \hat{y}_s$ since $\lim_{y \to \hat{y}_s^+} w_{SB} > r$ and $\lim_{y \to \hat{y}_s^-} w_{SB} = 0$. To further investigate the shape of $w_{SB}(y)$ for $y > \hat{y}_s$, compute the derivative of $(A3)$ with respect to y to obtain:

$$\frac{-u''\,(w_{SB}(y) - r)w'_{SB}(y)}{\left(u'(w_{SB}(y) - r)\right)^2} = -\frac{\partial}{\partial y}\left(\frac{f(y|e_L)}{f(y|e_H)}\right). \qquad (A13)$$

Equation $(A13)$ shows that $w'_{SB}(y) > 0$ in the domain of gains due to Assumption 2, $u''(\cdot) < 0$, and $u'(\cdot) > 0$ from Assumption 4. Hence, the optimal contract increases in performance in $y > \hat{y}$.

I conclude this part of the proof by studying the properties of $\hat{y}_s$ First, I show that $\hat{y}_s \in (\underline{y}, \overline{y})$. Suppose instead that $\hat{y}_s \geq \overline{y}$. In that case, $w_{SB}(y)$ is fully described by $(A12)$ and the agent experiences losses for all $y \in [\underline{y}, \overline{y}]$. Such a contract is rejected by the agent as it generates sizeable disutility, due to $\lambda \geq 1$, and because the outside option of the agent is assumed to be non-negative, $\overline{U} \geq 0$.

Now suppose that $\hat{y}_s \leq \underline{y}$. In that case the solution is fully described by $(A5)$ which is increasing everywhere. The principal can profitably deviate from that equilibrium by paying $w_{SB}(y) = 0$, the lowest amount at the lowest end of the output interval. Such a deviation does not necessarily lead the agent to reject the contract, since such low payment locates the agent in the domain of losses where he is risk seeking (Assumption 4) and thus willing to be exposed to such amount of risk. Hence it must be that $\hat{y}_s \in (\underline{y}, \overline{y})$.

iv) First-best (incentive compatible constraint not considered)

Let now $\gamma = 0$. Denote by $w_{FB}^{fo}(y)$ the candidate solution from the first-order approach under that restriction. Equation $(A3)$ collapses to

$$\frac{1}{u'\left(w_{FB}^{fo}(y) - r\right)} = \nu. \qquad (A14)$$

Which can be rewritten as

$$w_{FB}(y) = r + h'\left(\frac{1}{\nu}\right). \qquad (A15)$$

As in the derivation of the second-best contract, it can be shown that the principal is better off offering a lottery $L = (r, p; 0,1-p)$ rather than the solution resulting from the first-order approach, i.e. $w_{FB}^{fo}(y)$ satisfying $(A4)$ under $\gamma = 0$. To see how, note that $L = (r, p; 0,1-p)$ can be offered to the agent with a p satisfying

$$-\lambda u\left(r - w_{FB}^{fo}(y)\right) = -(1-p)\lambda u(r). \quad (A16)$$

As a consequence, replacing $w_{FB}^{fo}(y)$ for L leaves the agent's participation constraint unchanged. Moreover, because of the convexity of $u(\cdot)$ in the domain of losses (Assumption 4), it can be stated that

$$-\lambda u\left(r - w_{FB}(y)\right) < -\lambda u((1-p)r). \quad (A17)$$

Since $u(\cdot)$ is increasing everywhere (Assumption 4), $(A17)$ implies

$$w_{FB}(y) > pr. \qquad (A18)$$

Hence, offering L is more cost-effective to the principal.

From $(A11)$ with $\gamma = 0$, it can be established that when principal offers L, the Lagrangian multiplier, $\nu > 0$ can be sufficiently large to ensure

$$\frac{1}{\frac{\lambda u(r)}{r}} = \nu. \qquad (A19)$$

However, if ν is smaller, so that $(A19)$ does not hold, the expected value of L can be reduced by means of adjustments in p. Equation $(A10)$ shows that such reduction of the expected value does not change marginal utility, since the expected value of the lottery, $\overline{L}$, enters the agent's utility linearly. Therefore, the expected value can be reduced to $\overline{L} = 0$ by setting $p = 0$.

Instead, if v is too large, then $\overline{L}$, should be increased. Again, since $\mathbb{E}(CPT(L))$ is linear in $\overline{L}$, the contract's probability is set to $p = 1$, and $\overline{L} = r$.

Next, I study the shape of the candidate solutions. It was established above that in the domain of losses $w_{FB} = 0$ should be paid, which is performance insensitive. Moreover, the solution given by $(A15)$, paid in the domain of gains, is also performance insensitive. To see how, take the derivative of equation $(A14)$ with respect to y to obtain:

$$\frac{-u''\,(w_{FB}(y) - r)w'_{FB}(y)}{\left(u'(w_{FB}(y) - r)\right)^2} = 0. \qquad (A20)$$

Equation $(A20)$ implies that $w'_{FB}(y) = 0$; the contract pays a fixed amount in the domain of gains.

An implication of $w_{FB}(y) = 0$ and $w_{FB}(y)$ satisfying $(A5)$ being constant in performance is that, when implemented on their own, they either locate the agent in the domain of gains, $w_{FB}(y) \geq r$, or in the domain of losses, $w_{FB}(y) < r$. Since $\overline{U} \geq 0$ and $\lambda > 1, w_{FB} = 0$ cannot be implemented on its own as it induces considerable disutility and the agent would be better off rejecting the contract. Suppose instead that the principal's offer consists on the solution $w_{FB}(y)$ satisfying $(A15)$ everywhere. The agent will accept such a proposal as it fully protects him from risk and generates utility gains, $w_{FB}(y) > r$. The principal however, can profitably deviate from that solution by paying $w_{FB}(y) = 0$ for the lowest output levels. Due to the convexity of the agent's utility in losses, such a deviation will not necessarily lead the agent to reject the contract. As a consequence, the optimal contract consists of a combination of the schedules $w_{FB}(y) = 0$ and $w_{FB}(y)$ satisfying $(A15)$.

To conclude the proof, I define the point at which the transition from losses to gains takes place. To keep the agent indifferent between accepting and rejecting the contract, this transition must be given by an output level $y_f \in (\underline{y}, \overline{y})$ satisfying

$$\int_{y_f}^{\overline{y}} u\left(w_{FB}^{fo}(y) - r\right) f(y|e_H)\, \mathrm{d}y - \lambda \int_{\underline{y}}^{y_f} u(r) f(y|e_H) \mathrm{d}y - c \; = \overline{U}. \qquad (A21)$$

The existence of y_f is guaranteed by the facts that $w_{FB}^{fo}(y)$ satisfying $(A15)$ and L make the participation constraint bind for gains and losses, respectively.

As a result, the optimal incentive scheme locates the agent in the domain of losses with the following payment

$$w_{FB} = \begin{cases} 0 \; if \; y < \hat{y}_f, \\ L \; if \; y = \hat{y}_f. \end{cases} \qquad (A22)$$

Instead, when $y > \hat{y}_f$, $w_{SB}(y)$ satisfying $(A15)$ is proposed to the agent, locating him in the domain of gains. ■

Corollary 1

Suppose instead that $\hat{y}_f > \hat{y}_s$. That ordering implies a non-monotonic implementation of rewards. To see how, denote by y^* be the output level in $y > \hat{y}_f$ that satisfies $w_{SB}(y^*) = w_{FB}$. The existence of that output level is guaranteed by Assumption 2, since the second-best contract increases in $y > \hat{y}_s$ according to $w_{SB}(y) = h'\left(\frac{1}{v+\gamma\left(1-\frac{f(y|e_L)}{f(y|e_H)}\right)}\right)$ and the first best is constant in $y > \hat{y}_f$ and equal to $w_{SB}(y) = h'\left(\frac{1}{v}\right)$. For output levels $y \in (\hat{y}_s, \hat{y}_f)$ the second-best contract implements rewards with respect to the first-best, followed by punishments in $y \in (\hat{y}_f, y^*)$, to subsequently exhibit rewards in $y^* > y$. Such non-monotonic implementation of incentives does not motivate high effort in $y \in (\hat{y}_s, y^*)$. A contradiction since at the optimum the incentive compatibility constraint binds. Following a similar rationale, it can be shown that for $\hat{y}_s \geq \hat{y}_f$ incentives are monotonic. Hence, it must be that $\hat{y}_s \geq \hat{y}_f$. This proves part (i) of the corollary.

To prove the part (ii) of the corollary, suppose that $\hat{y}_s = \hat{y}_f$. Notice that in $y < \hat{y}_s$ the second-best contract does not implement punishments as both contracts exhibit $w_{SB} = w_{FB} = 0$. Hence, the inclusion of punishments in the domain of gains implies that, at $y = \hat{y}_s$, $1 < \frac{f(y|e_L)}{f(y|e_H)}$, which in turn implies, due to the fact that $h'(\cdot)$ is a decreasing function, that $h'\left(\frac{1}{v+\gamma\left(1-\frac{f(\hat{y}_s|e_L)}{f(\hat{y}_s|e_H)}\right)}\right) < h'\left(\frac{1}{v}\right)$. Such inequality together with $(A5)$ and $(A15)$ entail that the bonus of the first-best contract is larger at $y = \hat{y}_s$.

Let now $\hat{y}_s > \hat{y}_f$. The second best-contract includes punishments regardless of whether $h'\left(\frac{1}{v+\gamma\left(1-\frac{f(\hat{y}_s|e_L)}{f(\hat{y}_s|e_H)}\right)}\right) > h'\left(\frac{1}{v}\right)$ or $h'\left(\frac{1}{v+\gamma\left(1-\frac{f(\hat{y}_s|e_L)}{f(\hat{y}_s|e_H)}\right)}\right) < h'\left(\frac{1}{v}\right)$. However, note that if $\hat{y}_s = \overline{y}$ and $h'\left(\frac{1}{v+\gamma\left(1-\frac{f(\hat{y}_s|e_L)}{f(\hat{y}_s|e_H)}\right)}\right) < h'\left(\frac{1}{v}\right)$, the second-best contract does not include rewards for exerting high effort. Hence, it must be that $h'\left(\frac{1}{v+\gamma\left(1-\frac{f(\hat{y}_s|e_L)}{f(\hat{y}_s|e_H)}\right)}\right) > h'\left(\frac{1}{v}\right)$ at $\hat{y}_s = \overline{y}$. As a consequence, the second-best contract should exhibit a bonus larger than that of the first-best as y approaches $\overline{y}$.

Corollary 2

From $(A5)$ is evident that $\frac{\partial w_{SB}(y)}{\partial r} > 0$ in the domain of gains. Also, since $\lim_{y \to \hat{y}^-} w_{SB}(y) = 0$ and $\lim_{y \to \hat{y}^+} w_{SB}(y) > r$, that the optimal contract exhibits $\frac{\partial w_{SB}(y)}{\partial r} > 0$ in the domain of gains implies that the bonus becomes larger as r increases.

To investigate the influence of changes in r on the location of the bonus, compute the derivative of r with respect to the left-hand side of $(A11)$, $\frac{\partial}{\partial r}\left(\frac{r}{\lambda u(r)}\right) = \frac{\lambda\left(u(r)-ru'(r)\right)}{\left(\lambda u(r)\right)^2}$. Using the Taylor theorem around zero gives $u(0) = 0 = u(r) - u'(r)r + \frac{u''(r)r^2}{2}$. Hence, $\frac{\partial}{\partial r}\left(\frac{r}{\lambda u(r)}\right) > 0$ since $u''(r) < 0$. Therefore, a higher r implies that the equality in $(A11)$ is maintained for a higher output level, the bonus is given at a higher threshold $\hat{y}_s$.■

Corollary 3.

The left-hand side of equation $(A11)$ decreases as λ increases. To maintain that equality, and due to Assumption 2, $\hat{y}_s$ must decrease. Hence, for a larger value of λ, the second-best contract $w_{SB}(y)$ exhibits a smaller segment for which $(A22)$ is the solution. Finally, notice that λ does not enter $(A5)$, so changes in that parameter do not influence the shape of the second-best contract in the domain of gains nor the magnitude of the bonus. ■

Corollary 4

Under the assumed parametric utility function, equation $(A5)$ becomes

$$w_{SB}(y) = r + \left(\frac{1}{v + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right)}\right)^{\frac{1}{\alpha-1}} \qquad (A23)$$

Therefore, the derivative of $w_{SB}(y)$ with respect to α is equal to

$$\frac{\partial w_{SB}(y)}{\partial \alpha} = -\left(\frac{1}{v + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right)}\right)^{\frac{1}{\alpha-1}} \ln\left(\frac{1}{v + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right)}\right)\frac{1}{(\alpha-1)^2}. \qquad (A24)$$

From $(A24)$ it can be established that $\frac{\partial w_{SB}(y)}{\partial \alpha} < 0$ if $1 > v + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right)$

Instead, $\frac{\partial w_{SB}(y)}{\partial \alpha} > 0$ if and $1 < v + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right)$.

Moreover, under the assumed parametric utility, equation $(A11)$ becomes

$$\frac{r}{\lambda r^{\alpha}} = v + \gamma\left(1 - \frac{f(\hat{y}_s|e_L)}{f(\hat{y}_s|e_H)}\right). \qquad (A25)$$

The derivative of the left-hand side of the above equation is $\frac{\partial}{\partial \alpha}\left(\frac{r}{\lambda r^{\alpha}}\right) = -\frac{r^{1-\alpha}}{\lambda}\ln r$. which is negative if $r > 1$. In that case, equation $(A25)$ holds for a smaller $\hat{y}_s$, so a larger value of α leads to a lower threshold output at which the bonus is given. Instead, if $r < 1$ equation $(A25)$ holds for a larger $\hat{y}_s$. The proof is concluded by noting that the solution given by $(A22)$ does not depend on α.

Corollary 5.

Proof. In the present framework the agent makes, at most, two choices: accepting the contract or not, and choosing an amount of effort. There are thus three candidates for reference points: rejecting the contract and obtaining $\overline{U} \geq 0$, obtaining the minimum payment from accepting the contract after choosing e_H, and obtaining the minimum payment from accepting the contract after choosing e_L.

Denote the optimal contract for an agent with preferences characterized by (2) and the max-min reference point rule by $w_{SB}(y)$. Notice that $\min\{w_{SB}(y)\} = 0$. Moreover, an optimal contract generating utility $\overline{U} \geq 0$, is $w_{FB}(y)$ from Proposition 2 (i). Due to the fact that $c > 0$ and $\overline{U} \geq 0$, it must be that $\mathbb{E}\big(w_{FB}(y)\big) > 0$. Thus, under the max-min rule the agent's reference point is $w_{FB}(y)$.

Proposition 2 considered an exogenous but arbitrary reference point. Since the agent's preferences described by prospect theory, this solution remains optimal once the reference point $r = w_{FB}(y)$ is accounted for. Let $y_{\overline{U}}$ be the level of output satisfying:

$$\frac{1}{\frac{\lambda u\left(w_{FB}(y)\right)}{w_{FB}(y)}} = \nu + \gamma\left(1 - \frac{f\left(y_{\overline{U}}|e_L\right)}{f\left(y_{\overline{U}}|e_H\right)}\right), \qquad (A26)$$

Under the max-min rule the second-best contract consists of the solution given by the first-order condition:

$$\frac{1}{u'\left(w_{SB}(y) - w_{FB}(y)\right)} = \nu + \gamma\left(1 - \frac{f(y|e_H)}{f(y|e_L)}\right), \qquad (A27)$$

if $y > y_{\overline{U}}$, and the performance insensitive segment

$$w_{SB}(y) = \begin{cases} 0 \ if \ y < y_{\overline{U}} \\ L \ if \ y = y_{\overline{U}.} \end{cases} \qquad (A28)$$

Using the same rationale as in Proposition 2 (ii) it can be shown that $y_{\overline{U}} \in \left(\underline{y}, \overline{y}\right)$. An incentive scheme consisting of $(A28)$ will be rejected by the agent so $y_{\overline{U}} < \overline{y}$. Moreover, an incentive scheme paying $(A27)$ everywhere is dominated by a contract paying $w_{SB}(y) = 0$ at the low end of the output space. ■

Lemma 1.

Proof. To show that higher goals generate disutility, compute the derivative of (4) with respect to g to obtain

$$\frac{\partial CPT}{\partial g} = -\int_{g}^{\overline{y}} w'(g)u'\big(w(y) - w(g)\big)f(y|e)dy$$
$$- \lambda \int_{\underline{y}}^{g} w'(g)u'\big(w(g) - w(y)\big)f(y|e)\mathrm{d}y. \quad (A29)$$

Due to the properties $u'(\cdot) > 0$ and $\lambda > 1$ from Assumption 4, equation $(A30)$ is negative if $w'(g) > 0$. In that case, higher goals, g, induce disutility. This proves part (i) of the lemma.

Next, it is shown that higher goals incentivize high effort. To do so, integration by parts applied to equation (4) gives:

$$CPT = u\big(w(\overline{y}) - w(g)\big) - \int_{g}^{\overline{y}} w'(y)u'\big(w(y) - w(g)\big)F(y|e)\mathrm{d}y$$
$$- \lambda \int_{\underline{y}}^{g} w'(y)u'\big(w(g) - w(y)\big)\, F(y|e)dy. \qquad (A30)$$

The incentive compatible constraint for a given g can be rewritten using $(A30)$ as follows

$$IC\text{:} \; -\int_{g}^{\overline{y}} w'(y)u'\big(w(y) - w(g)\big)\big(F(y|e_H) - F(y|e_L)\big)\mathrm{d}y$$
$$- \lambda \int_{\underline{y}}^{g} w'(y)u'\big(w(g) - w(y)\big)\, \big(F(y|e_H) - F(y|e_L)\big)\mathrm{d}y \geq c. \quad (A31)$$

To investigate whether higher goals incentivize higher effort, derive $(A31)$ with respect to g to obtain:

$$-(\lambda - 1)w'(g)u'\big(w(g) - w(g)\big)\big(F(g|e_H) - F(g|e_L)\big)$$
$$+ \int_{g}^{\overline{y}} w'(y)w'(g)u''\big(w(y) - w(g)\big)\big(F(y|e_H) - F(y|e_L)\big)\mathrm{d}y$$
$$- \lambda \int_{\underline{y}}^{g} w'(y)w'(g)u''\big(w(g) - w(y)\big)\, \big(F(y|e_H) - F(y|e_L)\big)\mathrm{d}y. \qquad (A32)$$

Recall that $u'(\cdot) > 0$ and $u''(\cdot) < 0$ from Assumption 4, and that $F(y|e_L) \geq F(y|e_H)$ from Assumption 2. Hence, equation $(A32)$ is positive if $w'(g) > 0$ and if the incentive scheme exhibits $w'(y) > 0$ for $y > g$ and $w'(y) = 0$ for $y < g$. ■

Proposition 3.

Proof. Denote by $\nu > 0$ and $\gamma > 0$ the Lagrangian multipliers of the agent's participation and incentive compatibility constraints, respectively. The Lagrangian of the principal's maximization program can be written as:

$$\mathcal{L} = \big(S(y) - w(y)\big)f(y|e_H) +$$

$$\nu\big(\theta u(w(y) - w(g))f(y|e_H) - \lambda(1 - \theta)u\big(w(g) - w(y)\big)f(y|e_H) - c\big) \; +$$

$\gamma\big(\theta u(w(y) - w(g))\big(f(y|e_H) - f(y|e_L)\big) - \lambda(1-\theta)u\big(w(g) - w(y)\big)\big(f(y|e_H) - f(y|e_L)\big) - c\big).$ $\quad (A33)$

Where θ is an indicator function taking value $\theta = 1$ if $w(y) \geq w(g)$ and $\theta = 0$ otherwise.

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

$$-f(y|e_H) + \nu\left(\theta u'\big(w(y) - w(g)\big)f(y|e) + \lambda(1-\theta)u'\left(w(g) - w(y)\right)f(y|e)\right) + \gamma\left(\theta u'\big(w(y) - w(g)\big)\big(f(y|e_H) - f(y|e_L)\big) + \lambda(1-\theta)u'\big(w(g) - w(y)\big)\big(f(y|e_H) - f(y|e_L)\big)\right) = 0. \quad (A34)$$

Denoting by $w_{SB}^{fo}(y)$ the transfer that satisfies $(A34)$, the following equations are obtained after some re-arranging:

$$\frac{1}{u'\big(w_{SB}^{fo}(y) - w_{SB}^{fo}(g)\ \big)} = \nu + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right), \quad (A35)$$

if $\theta = 1$, and

$$\frac{1}{\lambda u'\left(w_{SB}^{fo}(g)\ - w_{SB}^{fo}(y)\right)} = \nu + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right), \quad (A36)$$

if $\theta = 0$. Using the inverse function $h(\cdot)$ from Assumption 4, rewrite equation $(A35)$ as

$$w_{SB}^{fo}(y) = w_{SB}^{fo}(g) + h'\left(\frac{1}{\nu + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right)}\right), \quad (A37)$$

and rewrite $(A36)$ as,

$$w_{SB}^{fo}(y) = w_{SB}^{fo}(g) - h'\left(\frac{1}{\lambda\left(\nu + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right)\right)}\right). \quad (A38)$$

From $(A37)$ and $(A38)$ it is evident that implementing the condition $w'(g) > 0$ from Lemma 1, implies $\frac{\partial w_{SB}^{fo}(y)}{\partial g} > 0$. The aim of the proof is to find a contract consistent with the strategy of setting high goals, so $\frac{\partial w_{SB}^{fo}(y)}{\partial g} > 0$ is imposed. Corollary 6 demonstrates that such strategy is optimal for the principal.

As in Proposition 2, I show that if $w_{SB}^{fo}(y)$ satisfying $(A36)$ exhibits, $0 < w_{SB}^{fo}(y)\ < w_{SB}^{fo}(g)$, then that candidate solution is not optimal. To see how, note that $-\lambda u\big(w_{SB}^{fo}(g) - w_{SB}^{fo}(y)\ \big)$ increases in $w_{SB}^{fo}(y)$, so there must exist a number $p \in (0,1)$ such that:

$$-\lambda u\left(w_{SB}^{fo}(g) - w_{SB}^{fo}(y)\right) = -(1-p)\lambda u\left(w_{SB}^{fo}(g)\right). \quad (A39)$$

Therefore, replacing $w_{SB}^{fo}(y)$ satisfying $(A36)$ by $L = (w_{SB}^{fo}(g), p; 0,1-p)$ leaves the agent's participation and incentive compatibility constraints unchanged. Moreover, the convexity of $-u(\cdot)$ in the domain of losses (Assumption 4), implies:

$$-\lambda u\big(w_{SB}^{fo}(g) - w_{SB}^{fo}(y)\big) < -\lambda u\Big((1-p)w_{SB}^{fo}(g)\Big). \qquad (A40)$$

Since $u(\cdot)$ increases everywhere (Assumption 4), Equation $(A40)$ implies

$$w_{SB}^{fo}(y) > p w_{SB}^{fo}(g). \qquad (A41)$$

Hence, L is more cost-effective for the principal than the candidate solution given by $(A36)$.

I turn to analyze the marginal incentives of $L = \big(w_{SB}^{fo}(g), p; 0,1-p\big)$. Substituting the expected value of that lottery, $\overline{L} = p w_{SB}^{fo}(g)$, into the agent's expected utility yields:

$$\mathbb{E}\big(CPT(L)\big) = -\lambda\left(1 - \frac{\overline{L}}{w_{SB}^{fo}(g)}\right)\Big(u\big(w_{SB}(g)\big)\Big) - c, \qquad (A42)$$

which is linear in $\overline{L}$. Implying that changes in the expected value of the lottery through adjustments of p, do not alter the agent's expected marginal utility.

Denote by $\hat{y}_g$ the performance level satisfying:

$$\frac{1}{\frac{\lambda u\left(w_{SB}^{fo}(g)\right)}{g}} = \nu + \gamma\left(1 - \frac{f(\hat{y}_g|e_L)}{f(\hat{y}_g|e_H)}\right). \qquad (A43)$$

That output level $\hat{y}_g$ is unique since the right-hand side of $(A43)$ strictly increases with performance due to Assumption 2 while the left-hand side of $(A43)$ is constant in performance. If these expressions cross, they do it once in $y \in (\underline{y}, \overline{y})$. Instead, if $\hat{y}_g = \underline{y}$ if the left-hand side of $(A44)$ is larger than the right-hand side for all y and otherwise, $\hat{y}_g = \overline{y}$.

Equation $(A42)$ shows that the agent's expected marginal utility does not change with changes in the expected value of lottery L. Thus, if $y \leq \hat{y}_g$, $\overline{L} = 0$ by setting $p = 0$. Instead, for $y > \hat{y}_g$, $\overline{L} = r$, the agent is brought to the domain of gains with $p = 1$. Therefore, the optimal incentive scheme in the domain of losses is:

$$w_{SB} = \begin{cases} 0 \; if \; y < \hat{y}_g, \\ L \; if \; y = \hat{y}_g. \end{cases} \qquad (A44)$$

The optimal contract thus consists of $(A44)$ and $(A37)$. A solution that exhibits a discrete jump at $y = \hat{y}_g$ since $\lim_{y \to \hat{y}_g^+} w_{SB}(y) = w_{SB}(g)$ and $\lim_{y \to \hat{y}_g^-} w_{SB}(y) = 0$. A direct consequence of $w'_{SB}(g) > 0$ is thus that the bonus at $\hat{y}_g$ becomes more sizeable the larger g is.

From equation $(A44)$ it can be established that in the domain of losses $w'_{SB}(y) = 0$. Furthermore, for performance levels $y > \hat{y}_g$, the agent is given the solution given by $(A37)$. To investigate the shape of that solution derive $(A35)$ with respect to y to obtain:

$$\frac{-u''\left(w_{SB}(y) - w_{SB}(g)\right)w'_{SB}(y)}{\left(u'\left(w_{SB}(y) - w_{SB}(y)\right)\right)^2} = -\frac{\partial}{\partial y}\left(\frac{f(y|e_L)}{f(y|e_H)}\right). \qquad (A45)$$

Equation $(A45)$ shows that in the domain of gains $w'_{SB}(y) > 0$ due to Assumption 2 and $u''(\cdot) < 0$ and $u'(\cdot) > 0$ from Assumption 4.

Next, I show that $\hat{y}_g = g$. Suppose instead that $\hat{y}_g < g$, in that case the principal is overinsuring the agent from risk in $y \in [\hat{y}_g, g]$, a segment located in the domain of losses where he is risk seeking due to Assumption 4. This is inefficient as the principal could increase profits by paying $w_{SB} = 0$ and exposing the agent to more risk. Due to the convexity of the agent's utility in segment function he would be willing to accept such contract.

Moreover, suppose that $\hat{y}_g > g$ The agent is overexposed to risk in $y \in [g, \hat{y}_g]$ a segment located in the domain of gains where he is risk averse. Such risk exposure that leads the agent to reject the contract. To ensure that the contract is accepted, the principal needs to insure the agent by offering w_{SB} given by $(A38)$ for any $y \geq g$. Then it must be that $\hat{y}_g = g$.

That $\hat{y}_g = g$ has several important implications. First, since $\lim_{y \to g^+} w_{SB}(y) = w_{SB}(g)$ and $\lim_{y \to g^-} w_{SB}(y) = 0$, a bonus of size $w_{SB}(g)$ is given at $y = g$. Second, to fulfill the condition $w'_{SB}(g)$ from Lemma 1, that bonus increases with the size of g. Finally, the conditions on the dependence of $w_{SB}(y)$ on y are met since $w'_{SB}(y) > 0$ if $y \geq g$ and $w'_{SB}(y) = 0$ if $y < g$. Thus, the incentives scheme given by $(A44)$ and $(A37)$ ensures that goals are motivational devices.■

Corollary 6.

Proof. Optimization of $(A33)$ with respect to g gives

$$\nu\left(-w'(g)\theta u'\left(w(y) - w(g)\right)f(y|e) - \lambda(1-\theta)w'(g)u'\left(w(g) - w(y)\right)f(y|e)\right)$$
$$+\gamma\left(-w'(g)\theta u'\left(w(y) - w(g)\right)\left(f(y|e_H) - f(y|e_L)\right)\right.$$
$$\left.-\lambda w'(g)(1-\theta)u'\left(w(g) - w(y)\right)\left(f(y|e_H) - f(y|e_L)\right)\right) = 0. \qquad (A46)$$

Denoting by g^* the goal level that satisfies $(A46)$, the following equations are obtained after some re-arranging

$$-w'(g)\left(\nu u'\left(w(y) - w(g^*)\right)f(y|e) + \gamma u'\left(w(y) - w(g^*)\right)\left(f(y|e_H) - f(y|e_L)\right)\right) = 0, \quad (A47)$$

if $\theta = 1$, and

$$-\lambda w'(g)\left(\nu u'\big(w(g^*)-w(y)\big)f(y|e)+\gamma u'\big(w(g^*)-w(y)\big)\big(f(y|e_H)-f(y|e_L)\big)\right)=0. \quad (A48)$$

if $\theta = 0$. Equations $(A47)$ and $(A48)$ both imply:

$$\nu+\gamma\left(1-\frac{f(y|e_L)}{f(y|e_H)}\right)=0. \qquad (A49)$$

Use $(A35)$ and $(A49)$ to obtain

$$u'(w_{SB}(y)-w_{SB}(g^*))=+\infty. \qquad (A50)$$

Since $\lim_{y\to g^+} w_{SB}(y)=w_{SB}(g)$ then $\lim_{y\to g^+} u'\big(w_{SB}(y)-w_{SB}(g^*)\big)=+\infty$. However, given that y is a random variable, the principal cannot set a goal that is just met. Instead, she can set a goal such that $(A50)$ holds on expectation. Using the fact that $u'(\cdot)$ is a decreasing and concave function, then

$$\mathbb{E}\big(u'(w_{SB}(y)-w_{SB}(g))\big)\geq u'\left(\mathbb{E}\big(w_{SB}(y)\big)-w_{SB}(g)\right). \quad (A51)$$

So g^* is set such that $\mathbb{E}\big(w_{SB}(y)\big)-w_{SB}(g^*)=\epsilon$ for arbitrarily small $\epsilon>0$ generating $u'\left(\mathbb{E}\big(w_{SB}(y)\big)-w_{SB}(g^*)\right)=+\infty$. As a result, $g^*>0$ is on expectation attainable.

Proposition 4.

i) Lagrangian and first-order conditions

Denote by $\nu\geq 0$ 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

$$\mathcal{L}=\big(S(y)-w(y)\big)f(y|e_H)$$

$$+\nu\big(w(y)f(y|e_H)+\theta u(w(y)-\widetilde{w})f(y|e_H)-\lambda(1-\theta)u\big(\widetilde{w}-w(y)\big)f(y|e_H)-c\big)$$

$$+\gamma\big(w(y)\big(f(y|e_H)-f(y|e_L)\big)+\theta u(w(y)-\widetilde{w})\big(f(y|e_H)-f(y|e_L)\big)-\lambda(1-\theta)u\big(\widetilde{w}-w(y)\big)\big(f(y|e_H)-f(y|e_L)\big)-c\big). \quad (A52)$$

Where θ is an indicator function taking value $\theta=1$ if $w(y)\geq\widetilde{w}$ and $\theta=0$ otherwise. Optimizing pointwise with respect to $w(y)$ gives

$$-f(y|e_H)+\nu\left(f(y|e_H)+\theta u'(w(y)-\widetilde{w})f(y|e_H)+\lambda(1-\theta)u'\left(\widetilde{w}-w(y)\right)f(y|e_H)\right)$$
$$+\gamma\Big(\big(f(y|e_H)-f(y|e_L)\big)+\theta u'(w(y)-\widetilde{w})\big(f(y|e_H)-f(y|e_L)\big)$$
$$+\lambda(1-\theta)u'\big(\widetilde{w}-w(y)\big)\big(f(y|e_H)-f(y|e_L)\big)\Big)=0. \qquad (A53)$$

Denote by $w_{SB}^{fo}(y)$ the transfer schedule satisfying $(A53)$ and by $\widetilde{w}_{SB}^{fo} := \int_{\underline{y}}^{\bar{y}} w_{SB}^{fo}(y)\, f(y|e)dy$. Rearranging the above equation gives:

$$\frac{1}{1+u'\left(w_{SB}^{fo}(y)-\widetilde{w}_{SB}^{fo}\right)} = \nu + \gamma\left(1-\frac{f(y|e_L)}{f(y|e_H)}\right), \qquad (A54)$$

if $\theta = 1$, and

$$\frac{1}{1+\lambda u'\left(\widetilde{w}_{SB}^{fo}(y)-w_{SB}^{fo}(y)\right)} = \nu + \gamma\left(1-\frac{f(y|e_L)}{f(y|e_H)}\right), \qquad (A55)$$

if $\theta = 0$. Using the inverse function $h(\cdot)$ from Assumption 4, equation $(A54)$ can be rewritten as follows

$$w_{SB}^{fo}(y) = \widetilde{w}_{SB}^{fo} + h'\left(\frac{1}{\nu+\gamma\left(1-\frac{f(y|e_L)}{f(y|e_H)}\right)} - 1\right), \qquad (A56)$$

and equation $(A55)$ can be rewritten as,

$$w_{SB}^{fo}(y) = \widetilde{w}_{SB}^{fo} - h'\left(\frac{1}{\lambda\left(\nu+\gamma\left(1-\frac{f(y|e_L)}{f(y|e_H)}\right)\right)} - 1\right). \qquad (A57)$$

ii) Suboptimality of first-order condition in the domain of losses

Next, I demonstrate that if $w_{SB}^{fo}(y)$ satisfying $(A57)$ is such that $0 < w_{SB}^{fo}(y) < \widetilde{w}_{SB}^{fo}$, then the principal is better off paying the lottery $L = \left(\widetilde{w}_{SB}^{fo}, p; 0, 1-p\right)$ with expected value $\overline{L} = p\widetilde{w}_{SB}^{fo}$. To see how, note that $\widetilde{w}_{SB}^{fo} - \lambda u\left(\widetilde{w}_{SB}^{fo} - w_{SB}^{fo}(y)\right)$ increases in $w_{SB}(y)$. Hence, there must exist a number $p \in (0,1)$ such that:

$$\widetilde{w}_{SB}^{fo} - \lambda u\left(\widetilde{w}_{SB}^{fo} - w_{SB}^{fo}(y)\right) = p\widetilde{w}_{SB}^{fo}(y) - (1-p)\lambda u\left(\widetilde{w}_{SB}^{fo}\right). \quad (A58)$$

So, replacing $w_{SB}^{fo}(y)$ from $(A57)$ by L leaves the agent's participation and the incentive compatibility constraints unchanged.

Using the fact that $-u(\cdot)$ is convex at $w_{SB}(y) < \widetilde{w}_{SB}$, as implied by Assumption 4, and equation $(A58)$, yields

$$\begin{aligned}(1-p)w_{SB}^{fo}(y) &= \lambda u\left(\widetilde{w}_{SB}^{fo} - w_{SB}^{fo}(y)\right) - (1-p)\lambda u\left(\widetilde{w}_{SB}^{fo}\right) \\ &< \lambda u\left(\widetilde{w}_{SB}^{fo} - w_{SB}^{fo}(y)\right) - \lambda u\left((1-p)\widetilde{w}_{SB}^{fo}\right). \quad (A59)\end{aligned}$$

Since $u(\cdot)$ is increasing, from the last inequality in the above equation it can be established that

$$w_{SB}^{fo}(y) > p\widetilde{w}_{SB}^{fo}. \qquad (A60)$$

Therefore, the lottery payment L improves upon the candidate solution from $(A57)$ since it provides the same incentives at lower costs for the principal.

I turn to analyze the marginal incentives of offering L. To that end, substitute the expected value of that lottery, $\overline{L}$, into the utility, $DA(L) = p\widetilde{w}_{SB} - (1-p)\big(\lambda u(\widetilde{w}_{SB})\big)$, to obtain:

$$\mathbb{E}\big(DA(L)\big) = \overline{L} + \left(1 - \frac{\overline{L}}{\widetilde{w}_{SB}^{fo}}\right)\left(\lambda u\big(\widetilde{w}_{SB}^{fo}\big)\right), \qquad (A61)$$

an expression that is linear in $\overline{L}$. Thus, changes in $\overline{L}$, through adjustments of p, do not alter the agent's expected marginal utility.

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

$$\frac{1}{1 + \frac{\lambda u\left(\widetilde{w}_{SB}^{fo}\right)}{\widetilde{w}_{SB}^{fo}}} = \nu + \gamma\left(1 - \frac{f(\hat{y}, |e_L)}{f(\hat{y}, |e_H)}\right). \qquad (A62)$$

For performance levels $y < \hat{y}$ the expected value of the lottery, $\overline{L}$, must be reduced by decreasing p. As established above, that reduction does not change marginal utility, so $\overline{L}$ should be reduced all the way to zero by setting $p = 0$. Instead, for performance levels $y > \hat{y}$ the expected value of the lottery $\overline{L}$ should be increased. Again, this change does not affect marginal utility, so $\overline{L}$ should be increased all the way to $\widetilde{w}_{SB}^{fo}$ by setting $p = 1$, bringing the agent to the domain of gains.

When the agent is in the domain of losses the optimal incentive scheme is:

$$w_{SB}(y) = \begin{cases} 0 \ if \ y < \hat{y}, \\ L \ if \ y = \hat{y}. \end{cases} \qquad (A63)$$

Alternatively, when the agent is in the domain of gains the optimal contract is given by $(A56)$, so for this domain $w_{SB}(y) = w_{SB}^{fo}(y)$. Notice that this solution exhibits a discrete jump at $y = \hat{y}$ since $\lim_{y \to \hat{y}^+} w_{SB}(y) > r$, as shown by $(A51)$ and $\lim_{y \to \hat{y}^-} w_{SB}(y) = 0$ as shown by $(A58)$.

iii) Shape of the second-best contract

Next, the shape of the incentives scheme is analyzed. By $(A63)$ is evident that for the domain of losses, $w_{SB}(y)$ is performance insensitive. To investigate the shape of $w_{SB}(y)$ for the domain of gains, take the derivative of $(A54)$ with respect to y to obtain:

$$\frac{-u''\left(w_{SB}(y)-\widetilde{w}_{SB}\right)w'_{SB}(y)}{\left(u'\left(w_{SB}(y)-\widetilde{w}_{SB}\right)\right)^2}=-\frac{\partial}{\partial y}\left(\frac{f(y|e_L)}{f(y|e_H)}\right). \qquad (A64)$$

where $\widetilde{w}_{SB} := \int_{\underline{y}}^{\bar{y}} w_{SB}(y)\, f(y|e)dy$. Equation $(A64)$ shows that in the domain of gains $w_{SB}{}'(y) > 0$ due to Assumption 2, and $u'' < 0$ and $u' > 0$ from Assumption 4. The contract thus increases in performance in the domain of gains

Finally, I show that $\hat{y} = y_{m2}$ where y_{m2} is the performance level satisfying $w_{SB}(y_{m2}) = \widetilde{w}_{SB}(y)$. Suppose instead that $\hat{y} < y_{m2}$, then the principal is overinsuring the agent in the segment $y \in [\hat{y}, y_{m2}]$, where the agent is risk seeking due to Assumption 4. This is inefficient as the princpal could increase profits by setting $w_{SB} = 0$ for all $y < y_{m2}$. Moreover, suppose that $\hat{y} > y_{m2}$. In that case, the agent is underinsured in the segment $y \in [y_{m2}, \hat{y}]$ which belongs to the domain of gains where the agent is risk averse (Assumption 4). This overexposure to risk leads the agent to reject the contract. To ensure that the contract is accepted, the principal needs to insure the agent by offering the transfer $w_{SB}(y)$ given by $(A56)$ for any $y \geq y_{m2}$. Hence, $\hat{y} = y_{m2}$.

iv) First-best contract

Let $\gamma = 0$. Denote by $w_{FB}^{fo}(y)$ the solution from the first-order approach to the program described in $(A53)$ under that restriction. Also, let $\widetilde{w}_{FB}^{fo}$ be the expected value of that solution. Equation $(A56)$ becomes

$$w_{FB}^{fo}(y) = \widetilde{w}_{FB}^{fo} + h'\left(\frac{1}{\nu} - 1\right), \qquad (A65)$$

and $(A57)$ becomes,

$$w_{FB}^{fo}(y) = \widetilde{w}_{FB}^{fo} - h'\left(\frac{1}{\lambda}\left(\frac{1}{\nu} - 1\right)\right). \qquad (A66)$$

If $w_{FB}^{fo}(y)$ satisfying $(A66)$ is such that $0 < w_{FB}^{fo}(y) < \widetilde{w}_{FB}^{fo}$ then it cannot be optimal. Since $\widetilde{w}_{FB}^{fo} - \lambda u\left(\widetilde{w}_{FB}^{fo} - w_{FB}^{fo}(y)\right)$ increases in $w_{FB}^{fo}(y)$, then there must exist a number $p \in (0,1)$ such that:

$$\widetilde{w}_{FB}^{fo} - \lambda u\left(\widetilde{w}_{FB}^{fo} - w_{FB}^{fo}(y)\right) = p\widetilde{w}_{FB}^{fo} - (1-p)\lambda u\left(\widetilde{w}_{FB}^{fo}\right). \quad (A67)$$

So, replacing $w_{FB}^{fo}(y)$ from $(A66)$ by $L = (\widetilde{w}_{FB}^{fo}, p; 0, 1-p)$ leaves the participation constraint and the incentive compatibility constraint unchanged.

Using $(A67)$ and the convexity of $u(\cdot)$ in the domain of losses gives

$$\begin{aligned}(1-p)\widetilde{w}_{FB}^{fo} &= \lambda u\left(\widetilde{w}_{FB}^{fo} - w_{FB}^{fo}(y)\right) - (1-p)\lambda u\left(\widetilde{w}_{FB}^{fo}\right) \\ &< \lambda u\left(\widetilde{w}_{FB}^{fo} - w_{FB}^{fo}(y)\right) - \lambda u\left((1-p)\widetilde{w}_{FB}^{fo}\right). \qquad (A68)\end{aligned}$$

Since $u(\cdot)$ increases everywhere (Assumption 4), equation $(A68)$ implies that

$$w_{FB}^{fo}(y) > p\widetilde{w}_{FB}^{fo}. \qquad (A69)$$

Therefore, the lottery payment $L = \left(\widetilde{w}_{FB}^{fo}, p; 0,1-p\right)$ improves upon $w_{FB}^{fo}(y)$, the candidate solution given by $(A66)$, as it provides the same incentives at lower costs for the principal.

Next, I investigate the incentives generated from offering L. To that end, substitute the expected value of the lottery, $\overline{L} = p\widetilde{w}_{FB}^{fo}$, into the agent's expected utility, $DA(L) = p\widetilde{w}_{FB}^{fo} - (1-p)\left(\lambda u\left(\widetilde{w}_{FB}^{fo}\right)\right)$, to obtain:

$$\mathbb{E}\big(DA(L)\big) = \overline{L} + \left(1 - \frac{\overline{L}}{\widetilde{w}_{FB}}\right)\left(\lambda u(\widetilde{w}_{FB})\right). \qquad (A70)$$

Notice that $(A70)$ is linear in $\overline{L}$. Implying that changes in $\overline{L}$, through adjustments in p, do not alter the agent's marginal utility.

When the principal offers L, the Lagrangian multiplier, ν can be large enough to ensure:

$$\frac{1}{1 - \frac{\lambda u\left(\widetilde{w}_{FB}^{fo}\right)}{\widetilde{w}_{FB}^{fo}}} = \nu. \qquad (A71)$$

However, if ν is too small, so that the left-hand side of $(A71)$ is larger than the right-hand side, $\overline{L}$ should be reduced by decreasing p. As established above, a reduction of the expected value of the lottery does not change marginal utility. Thus, $\overline{L}$ can be set to zero by $p = 0$. This brings the agent in the domain of losses. In contrast, if ν is too large, such that the left-hand side of $(A67)$ is smaller than the right-hand side of that equation, then $\overline{L}$, should be increased. Again, since $\mathbb{E}\big(DA(L)\big)$ is linear in $\overline{L}$, the expected value of the lottery should be increased to $\widetilde{w}_{FB}$ by setting $p = 1$, bringing the agent to the domain of gains.

I turn to analyze the shape of the first-best contract. For the domain of losses that the optimal contract pays $w_{FB}(y) = 0$ which is performance insensitive. The solution for the domain of gains, the contract $w_{FB}(y)$ satisfying $(A65)$ is also constant in performance. To see how, derive $(A54)$ under $\gamma = 0$ with respect to $w_{FB}(y)$ to obtain:

$$\frac{-u''\left(w_{FB}(y) - \widetilde{w}_{FB}\right)\left(w'_{FB}(y)\right))}{\left(1 + u'\left(w_{FB}(y) - \widetilde{w}_{FB}\right)\right)^2} = 0 \qquad (A72)$$

Equation $(A67)$ shows that $w'_{FB}(y) = 0$ and the first-best contract is performance insensitive in the domain of gains.

The first-best contract consists of two components: one that pays $w_{FB}(y) = 0$ and brings the agent to the domain of losses, and another one paying $w_{FB}(y)$ satisfying $(A65)$. These components cannot be implemented on their own. Suppose instead that the first-best contract pays $w_{FB}(y)$ satisfying $(A65)$ everywhere. This contract is constant and necessitates $w_{FB}(y) > \widetilde{w}_{FB}$, two requirements that together imply that its transfers are unbounded. A contradiction since the first-best contract maximizes the principal's objective function. Instead, a contract paying $w_{FB}(y) = 0$ everywhere will be rejected by the agent, since it generates disutility and does not satisfy the participation constraint. A contradiction since at the optimum the agent's participation constraint binds.

Hence, it must be that the second-best pays $w_{FB}(y) = 0$ at the lower end of the output space, while $w_{FB}(y)$ satisfying $(A65)$ is paid at the higher end of the output space. The optimal contract consists thus of $w_{FB}(y) = 0$ if $y < y_{m_1}$ and $w_{FB}(y)$ satisfying $(A61)$ if $y \geq y_{m_1}$, where $y_{m1} \in (\underline{y}, \overline{y})$ satisfies $w_{FB}(y_{m1}) = \widetilde{w}_{FB}$. ■

Proposition 5.

i) Lagrangian and first-order conditions

Let $\nu \geq 0$ and $\gamma \geq 0$ be the Lagrangian multipliers of the agent's participation and incentive compatibility constraints, respectively. The Lagrangian of the principal's maximization problem is

$$\begin{aligned}\mathcal{L} = &\big(S(y) - w(y)\big)f(y|e_H) \\ &+ \nu\Bigg(w(y)f(y|e_H) + \eta\theta\int_{\underline{y}}^{\overline{y}} u(w(y) - w(\tilde{y})f(y|e_H)f(\tilde{y}|e)\mathrm{d}\tilde{y} \\ &- \eta\lambda(1-\theta)\int_{\underline{y}}^{\overline{y}} u\big(w(\tilde{y}) - w(y)\big)f(y|e_H)f(\tilde{y}|e)\mathrm{d}\tilde{y} - c\Bigg) \\ &+ \gamma\Bigg(w(y)\big(f(y|e_H) - f(y|e_L)\big) \\ &+ \theta\eta\int_{\underline{y}}^{\overline{y}} u(w(y) - w(\tilde{y})\big(f(y|e_H) - f(y|e_L)\big)f(\tilde{y}|e)\mathrm{d}\tilde{y} \\ &- \lambda\eta(1-\theta)\int_{\underline{y}}^{\overline{y}} u\big(w(\tilde{y}) - w(y)\big)\big(f(y|e_H) - f(y|e_L)\big)f(\tilde{y}|e)\mathrm{d}\tilde{y} - c\Bigg). \quad (A73)\end{aligned}$$

Where θ is an indicator function taking value $\theta = 1$ if $w(y) \geq w(\tilde{y})$ and $\theta = 0$ otherwise. Pointwise maximization with respect to $w(y)$ gives

$$
\begin{aligned}
-f(y|e_H) + \nu\Bigg(& f(y|e_H) + \theta\eta \int_{\underline{y}}^{\overline{y}} u'\big(w_{SB}(y) - w_{SB}(\tilde{y})\big) f(y|e_H) f(\tilde{y}|e)\mathrm{d}\tilde{y} \\
& + \eta\lambda(1-\theta) \int_{\underline{y}}^{\overline{y}} u'\big(w_{SB}(\tilde{y}) - w_{SB}(y)\big) f(y|e_H) f(\tilde{y}|e)\mathrm{d}\tilde{y}\Bigg) \\
& + \gamma\Bigg(\big(f(y|e_H) - f(y|e_L)\big) \\
& + \eta\theta \int_{\underline{y}}^{\overline{y}} u'\big(w_{SB}(y) - w_{SB}(\tilde{y})\big)\big(f(y|e_H) - f(y|e_L)\big) f(\tilde{y}|e)\mathrm{d}\tilde{y} \\
& + \lambda(1-\theta)\eta \int_{\underline{y}}^{\overline{y}} u'\big(w_{SB}(\tilde{y}) - w_{SB}(y)\big)\big(f(y|e_H) - f(y|e_L)\big) f(\tilde{y}|e)\mathrm{d}\tilde{y} - c\Bigg) \\
& = 0. \qquad (A74)
\end{aligned}
$$

Denoting by $w_{SB}^{fo}(y)$ the transfer that satisfies $(A74)$, the following equations are obtained after some rearranging:

$$
\frac{1}{1 + \eta \int_{\underline{y}}^{\overline{y}} u'\big(w_{SB}^{fo}(y) - w_{SB}^{fo}(\tilde{y})\big) f(\tilde{y}|e_H)\mathrm{d}\tilde{y}} = \nu + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right), \qquad (A75)
$$

if $\theta = 1$, and

$$
\frac{1}{1 + \lambda\eta \int_{\underline{y}}^{\overline{y}} u'\big(w_{SB}^{fo}(\tilde{y}) - w_{SB}^{fo}(y)\big) f(\tilde{y}|e_H)\mathrm{d}\tilde{y}} = \nu + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right), \qquad (A76)
$$

if $\theta = 0$.

ii) Suboptimality of first-order condition in losses

I first show that if the candidate solution satisfying equation $(A76)$ is interior, $w_{SB}^{fo}(y) < w_{SB}^{fo}(\tilde{y})$, then it cannot be optimal. To see why, notice that

$$
\mathbb{E}_{\tilde{y}}\left(w_{SB}^{fo}(\tilde{y})\right) - \eta\lambda \int_{\underline{y}}^{\tilde{y}} \int_{\underline{y}}^{\overline{y}} u\big(w_{SB}^{fo}(\tilde{y}) - w_{SB}^{fo}(y)\big)\, f(y|e)\, f(\tilde{y}|e)\, \mathrm{d}\tilde{y}\mathrm{d}y, \qquad (A77)
$$

increases in $w_{SB}^{fo}(y)$. Hence, there must exist a number $p \in [0,1]$ such that:

$$
\begin{aligned}
\mathbb{E}_{\tilde{y}}\left(w_{SB}^{fo}(\tilde{y})\right) - \eta\lambda \int_{\underline{y}}^{\tilde{y}} \int_{\underline{y}}^{\overline{y}} u\big(w_{SB}^{fo}(\tilde{y}) - w_{SB}^{fo}(y)\big)\, f(y|e)\, f(\tilde{y}|e)\, \mathrm{d}\tilde{y}\mathrm{d}y \\
= p\, \mathbb{E}_{\tilde{y}}\left(w_{SB}^{fo}(\tilde{y})\right) \\
- p(1-p)\eta\lambda\left(\int_{\underline{y}}^{\tilde{y}} \int_{\underline{y}}^{\overline{y}} u\left(w_{SB}^{fo}(\tilde{y})\right) f(y|e)\, f(\tilde{y}|e)\, \mathrm{d}\tilde{y}\mathrm{d}y\right). \quad (A78)
\end{aligned}
$$

So, replacing $w_{SB}^{fo}(y)$ from $(A76)$ by the lottery-like payment $L = \left(w_{SB}^{fo}(\tilde{y}), p; 0,1-p\right)$ leaves the participation constraint and the incentive compatibility constraint unchanged in the domain of losses.

Due to the convexity of $u(\cdot)$ in the domain of losses (Assumption 4), equation $(A78)$ implies

$$\begin{aligned}
&\mathbb{E}_{\tilde{y}}\left(w_{SB}^{fo}(\tilde{y})\right) - p\,\mathbb{E}_{\tilde{y}}\left(w_{SB}^{fo}(\tilde{y})\right) \\
&\quad = \eta\lambda \int_{\underline{y}}^{\tilde{y}} \int_{\underline{y}}^{\overline{y}} u\left(w_{SB}^{fo}(\tilde{y}) - w_{SB}^{fo}(y)\right) f(y|e)\, f(\tilde{y}|e)\, \mathrm{d}\tilde{y}\mathrm{d}y \\
&\quad - p(1-p)\eta\lambda \left(\int_{\underline{y}}^{\tilde{y}} \int_{\underline{y}}^{\overline{y}} u\left(w_{SB}^{fo}(\tilde{y})\right) f(y|e)\, f(\tilde{y}|e)\, \mathrm{d}\tilde{y}\mathrm{d}y\right) \\
&\quad < \eta\lambda \int_{\underline{y}}^{\tilde{y}} \int_{\underline{y}}^{\overline{y}} u\left(w_{SB}^{fo}(\tilde{y}) - w_{SB}^{fo}(y)\right) f(y|e)\, f(\tilde{y}|e)\, \mathrm{d}\tilde{y}\mathrm{d}y \\
&\quad - p\eta\lambda \left(\int_{\underline{y}}^{\tilde{y}} \int_{\underline{y}}^{\overline{y}} u\left((1-p)w_{SB}^{fo}(\tilde{y})\right) f(y|e)\, f(\tilde{y}|e)\, \mathrm{d}\tilde{y}\mathrm{d}y\right). \qquad (A79)
\end{aligned}$$

Since $u'(\cdot) > 0$, it can be established from $(A79)$ that:

$$w_{SB}^{fo}(y) > p w_{SB}^{fo}(\tilde{y}). \qquad (A80)$$

Therefore, the payment L improves upon the candidate solution from $(A76)$ since it provides the same incentives at lower costs for the principal.

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

$$KR(L) = \overline{L} + \frac{\overline{L}}{w_{SB}^{fo}(\tilde{y})}\left(1 - \frac{\overline{L}}{w_{SB}^{fo}(\tilde{y})}\right)\left(u\left(w_{SB}^{fo}(\tilde{y})\right) - \lambda u\left(-w_{SB}^{fo}(\tilde{y})\right)\right), \qquad (A81)$$

an expression that is not linear in $\overline{L}$. Hence, changes in $\overline{L}$, via changes in p, affect the agent's marginal utility. So unlike Propositions 2,3, and 4 there is a $p^* \in [0,1]$ that maximizes the agent's utility that can be away from the boundaries. To find that optimal value, express $(A81)$ in terms of p:

$$KR(L) = p w_{SB}^{fo}(\tilde{y}) + p(1-p)\eta\left(u\left(w_{SB}^{fo}(\tilde{y})\right) - \lambda u\left(-w_{SB}^{fo}(\tilde{y})\right)\right). \qquad (A82)$$

The first-order condition of $(A82)$ with respect to p is:

$$w_{SB}^{fo}(\tilde{y}) + (1-2p)\eta\left(u\left(w_{SB}^{fo}(\tilde{y})\right) - \lambda u\left(-w_{SB}^{fo}(\tilde{y})\right)\right) = 0. \qquad (A83)$$

The second-order condition of $(A82)$ with respect to p is:

$$-2\eta\left(u\left(w_{SB}^{fo}(\tilde{y})\right) - \lambda u\left(-w_{SB}^{fo}(\tilde{y})\right)\right). \qquad (A84)$$

Equation $(A84)$ is negative as long as $u\left(w_{SB}^{fo}(\tilde{y})\right) - \lambda u\left(-w_{SB}^{fo}(\tilde{y})\right) > 0$ which is implied by $u'(\cdot) > 0$ from Assumption 4. Therefore, the optimal probability satisfies $(A83)$ and has the closed-form $p_2^* = \frac{1}{2} + \frac{w_{SB}^{fo}(\tilde{y})}{2\eta\left(u\left(w_{SB}^{fo}(\tilde{y})\right) - \lambda u\left(-w_{SB}^{fo}(\tilde{y})\right)\right)}$. Hence, for outcomes in the domain of losses lottery $L^* = (w_{SB}(\tilde{y}), 1 - p_2^*; 0, p_2^*)$ is offered. Notice that $p_2^* \in \left[\frac{1}{2}, 1\right]$ since $0 < \frac{w_{SB}^{fo}(\tilde{y})}{u\left(w_{SB}^{fo}(\tilde{y})\right) - \lambda u\left(-w_{SB}^{fo}(\tilde{y})\right)} < 1$.

iii) The shape of the second-best contract

The shape of the incentives scheme is analyzed. For the domain of gains the shape of the optimal schemes is given by the derivative of $(A74)$ with respect to y, which is computed next:

$$\frac{-\eta \int_{\underline{y}}^{\overline{y}} u''\left(w_{SB}^{fo}(y) - w_{SB}^{fo}(\tilde{y})\right) f(\tilde{y}|e_H)\mathrm{d}\tilde{y}\, w'^{fo}_{SB}(y)}{\left(1 + \eta \int_{\underline{y}}^{\overline{y}} u'\left(w_{SB}^{fo}(y) - w_{SB}^{fo}(\tilde{y})\right) f(\tilde{y}|e_H) d\tilde{y}\right)^2} = -\frac{\partial}{\partial y}\left(\frac{f(y|e_L)}{f(y|e_H)}\right). \qquad (A85)$$

Equation $(A85)$ shows that in the domain of gains $w_{SB}^{fo\prime}(y) > 0$ due to Assumption 2 and $u''(\cdot) < 0$ and $u'(\cdot) > 0$ from Assumption 4. Instead, the lottery L^* depends on constants of the model and on $w_{SB}^{fo}(\tilde{y})$ for given $\tilde{y}$.

Next, I show that the second-best contract unavoidably consists of two segments, one paying $w_{SB}(y) = L^*$ and another one paying $w_{SB}(y)$ satisfying $(A75)$. Suppose instead that the optimal contract only pays $w_{SB}^{fo}(y)$ satisfying $(A75)$. Since $\mathbb{E}(L^*) < w_{SB}^{fo}(y)$, the principal can profitably deviate from that solution by paying L^* at the lower end of the output space. The payment L^* will be evaluated as a loss when $w_{SB}^{fo}(y)$ satisfying $(A75)$ is taken as reference point. Due to the convexity of $u(\cdot)$ in the domain of losses (Assumption 4), the agent will be willing to accept that contract. Suppose now that the optimal contract pays L^* for all y. Both outcomes of this lottery are performance insensitive and thus not providing any incentives to exert high effort. This is a contradiction since at the optimum the agent's incentive compatibility constraint binds. Hence, it must be that $w_{SB}^{fo}(y)$ satisfying $(A75)$ is included in the contract.

Finally, I show that there exists a unique threshold output below which L^* is paid and above which $w_{SB}(y)$ satisfying $(A74)$ is paid. The optimal incentives of paying L^* are given by

$$\frac{1}{1 + \left(1 - 2\overline{L}\right)\frac{\left(u(w_{SB}(\tilde{y})) - \lambda u(-w_{SB}(\tilde{y}))\right)}{w_{SB}(\tilde{y})}} = \nu + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right), \qquad (A86)$$

Note that $\nu + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right) > 0$. Therefore, equation $(A86)$ entails that $\overline{L} = w_{FB}(\tilde{y})p_2^* < \frac{1}{2}$ which in turn implies $w_{FB}(\tilde{y}) < 1$ since $p_2^* \in \left[\frac{1}{2}, 1\right]$. Rewrite $(A86)$ as

$$\frac{1}{1+\left(\frac{2\eta\left(u(w_{SB}(\tilde{y}))-\lambda u(-w_{SB}(\tilde{y}))\right)}{w_{SB}(\tilde{y})}-\eta\left(u(w_{SB}(\tilde{y}))-\lambda u(-w_{SB}(\tilde{y}))\right)-w_{SB}(\tilde{y})\right)} = \nu + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right), \quad (A87)$$

where the left-hand side of $(A87)$ strictly decreases in $w_{SB}(\tilde{y})$ since $\frac{\partial}{\partial w_{SB}(\tilde{y})}\left(\frac{w_{SB}(\tilde{y})}{2\eta\left(u(w_{SB}(\tilde{y}))-\lambda u(-w_{SB}(\tilde{y}))\right)}\right) < 0$ and $w_{SB}(\tilde{y}) < 1$. Since the right-hand side of $(A86)$ increases in y (Assumption 2), there exists an output level $\tilde{y}_{k2} \in (\underline{y}, \overline{y})$ such that $(A86)$ holds. For all output levels $y < \tilde{y}_{k2}$ then $L^* = (w_{SB}(\tilde{y}_{k2}), 1 - p_2^*; 0, p_2^*)$ is given, otherwise $w_{SB}(y)$ satisfying $(A75)$ is paid.

iv) *First-best contract*

Let $\gamma = 0$. Denote by $w_{FB}^{fo}(y)$ the candidate solution from the first-order approach. This solution is now given by,

$$\frac{1}{1 + \eta\mathbb{E}_{\tilde{y}}\left(u'\left(w_{FB}^{fo}(y) - w_{FB}^{fo}(\tilde{y})\right)\right)} = \nu, \qquad (A88)$$

in the domain of gains, and

$$\frac{1}{1 + \eta\lambda\mathbb{E}_{\tilde{y}}\left(u'\left(w_{FB}^{fo}(\tilde{y}) - w_{FB}^{fo}(y)\right)\right)} = \nu, \qquad (A89)$$

In the domain of losses. From $(A88)$ and $(A89)$ it can be established that $\nu > 0$ due to $u'(\cdot) > 0$, $\eta > 0$, and $\lambda > 1$ from Assumption 4.

I first show that if $w_{FB}(y)$ satisfying $(A89)$ exhibits $0 < w_{FB}^{fo}(y) < w_{FB}^{fo}(\tilde{y})$ then it cannot be optimal. Recall that in the domain of losses $w_{FB}^{fo}(\tilde{y}) > w_{FB}^{fo}(y)$ and notice that the expression

$$\mathbb{E}_{\tilde{y}}(w_{FB}(\tilde{y})) - \eta\lambda\int_{\underline{y}}^{\tilde{y}}\int_{\underline{y}}^{\overline{y}} u\left(w_{FB}(\tilde{y}) - w_{FB}(y)\right) f(\tilde{y}|e)\, d\tilde{y} f(y|e)\, dy \qquad (A90)$$

increases in $w_{FB}^{fo}(y)$. Hence, there must exist a number $p \in [0,1]$ such that:

$$\begin{aligned}\mathbb{E}_{\tilde{y}}\left(w_{FB}^{fo}(\tilde{y})\right) - \eta\lambda\int_{\underline{y}}^{\tilde{y}}\int_{\underline{y}}^{\overline{y}} u\left(w_{FB}^{fo}(\tilde{y}) - w_{FB}^{fo}(y)\right) f(\tilde{y}|e)\, d\tilde{y} f(y|e)\, dy \\ = p\ \mathbb{E}_{y}\left(w_{FB}(\tilde{y})\right) \\ - p(1-p)\eta\lambda\left(\int_{\underline{y}}^{\tilde{y}}\int_{\underline{y}}^{\overline{y}} u\left(w_{FB}^{fo}(\tilde{y})\right) f(\tilde{y}|e)\, d\tilde{y} f(y|e) dy\right) \quad (A91)\end{aligned}$$

So, replacing $w_{FB}(y)$ from $(A89)$ by the lottery-like transfer $L = (w_{FB}(\tilde{y}), p; 0,1-p)$ leaves the participation and the incentive compatibility constraints unchanged.

Equation $(A91)$ together with the convexity of $u(\cdot)$ in $w_{FB}(y) < w_{FB}^{fo}(\tilde{y})$(Assumption 4), imply

$$\begin{aligned}
\mathbb{E}_{\tilde{y}}\left(w_{FB}^{fo}(\tilde{y})\right) - p\,\mathbb{E}_{\tilde{y}}\left(w_{FB}^{fo}(\tilde{y})\right) \\
&= \eta\lambda \int_{\underline{y}}^{\tilde{y}} \int_{\underline{y}}^{\overline{y}} u\left(w_{FB}^{fo}(\tilde{y}) - w_{FB}^{fo}(y)\right) f(y|e)\, f(\tilde{y}|e)\, d\tilde{y}dy \\
&\quad - p(1-p)\eta\lambda \left(\int_{\underline{y}}^{\tilde{y}} \int_{\underline{y}}^{\overline{y}} u\left(w_{FB}^{fo}(\tilde{y})\right) f(y|e)\, f(\tilde{y}|e)\, d\tilde{y}dy \right) \\
&< \eta\lambda \int_{\underline{y}}^{\tilde{y}} \int_{\underline{y}}^{\overline{y}} u\left(w_{FB}^{fo}(\tilde{y}) - w_{FB}^{fo}(y)\right) f(y|e)\, f(\tilde{y}|e)\, d\tilde{y}dy \\
&\quad - p\eta\lambda \left(\int_{\underline{y}}^{\tilde{y}} \int_{\underline{y}}^{\overline{y}} u\left((1-p)w_{FB}^{fo}(\tilde{y})\right) f(y|e)\, f(\tilde{y}|e)\, d\tilde{y}dy \right). \qquad (A93)
\end{aligned}$$

Since $u(\cdot)$ increases everywhere, then from equation $(A93)$ it can be established that:

$$w_{FB}^{fo}(y) > p w_{FB}^{fo}(\tilde{y}). \qquad (A94)$$

Therefore, $L = (w_{FB}(\tilde{y}), p; 0,1-p)$ improves upon the candidate solution from $(A89)$ since it provides the same incentives at lower costs for the principal and the optimal contract cannot exhibit $0 < w_{FB}^{fo}(y) < w_{FB}^{fo}(\tilde{y})$.

I turn to analyze the marginal incentives of offering the contract $L = \left(w_{FB}^{fo}(\tilde{y}), p; 0,1-p\right)$. Denote its expected value by $\overline{L} = p w_{FB}^{fo}(\tilde{y})$ and replace it from the agent's expected utility to obtain

$$\begin{aligned}
KR(L) = \overline{L} \\
&+ \frac{\overline{L}}{w_{FB}^{fo}(\tilde{y})}\left(1 - \frac{\overline{L}}{w_{FB}^{fo}(\tilde{y})}\right)\left(u\left(w_{FB}^{fo}(\tilde{y})\right)\right. \\
&\left. - \lambda u\left(-w_{FB}^{fo}(\tilde{y})\right)\right). \qquad (A95)
\end{aligned}$$

Equation $(A95)$ shows that the agent's utility is not linear in $\overline{L}$ and changes in p affect the agent's marginal utility. As a result, there must be an optimal probability $p^* \in [0,1]$ that maximizes the agent's marginal utility in the loss domain. To find that optimal probability, express $(A93)$ in terms of p to obtain,

$$KR(L) = p\, w_{FB}^{fo}(\tilde{y}) + p(1-p)\eta\left(u\left(w_{FB}^{fo}(\tilde{y})\right) - \lambda u\left(-w_{FB}^{fo}(\tilde{y})\right)\right). \qquad (A96)$$

The first-order condition of $(A96)$ with respect to p is:

$$w_{FB}^{fo}(\tilde{y}) + (1-2p)\eta\left(u\left(w_{FB}^{fo}(\tilde{y})\right) - \lambda u\left(-w_{FB}^{fo}(\tilde{y})\right)\right) = 0. \qquad (A97)$$

The second-order condition of $(A96)$ with respect to p is:

$$-2\eta\left(u\left(w_{FB}^{fo}(\tilde{y})\right) - \lambda u\left(-w_{FB}^{fo}(\tilde{y})\right)\right). \qquad (A98)$$

Equation $(A98)$ is negative as long as $u\left(w_{FB}(\tilde{y})\right) > \lambda u\left(-w_{FB}(\tilde{y})\right)$ which is implied by $u'(\cdot) > 0$ from Assumption 4. Hence, the lottery-like payment is offered with interior probability p_1^*, that satisfies $(A97)$ and has closed-form $p_1^* = \frac{1}{2} + \frac{w_{FB}^{fo}(\tilde{y})}{2\eta\left(u\left(w_{FB}^{fo}(\tilde{y})\right) - \lambda u\left(-w_{FB}^{fo}(\tilde{y})\right)\right)}$. The first-best contract in the loss domain is $L^* = (w_{FB}(\tilde{y}), p_1^*; 0, 1-p_1^*)$.

I turn to analyze the shape of the two possible solutions L^* and $w_{FB}^{fo}(y)$ satisfying $(A88)$. The solution to $(A88)$, which is the contract that is given in the domain of gains, is constant in performance. To see why compute the derivative of $(A88)$ with respect to y to obtain:

$$\frac{-\eta\mathbb{E}_{\tilde{y}}\left(u''\left(w_{FB}^{fo}(y) - w_{FB}^{fo}(\tilde{y})\right)\right)w'^{fo}_{FB}(\tilde{y})}{\left(1+\eta\mathbb{E}_{\tilde{y}}\left(u'\left(w_{FB}^{fo}(y) - w_{FB}^{fo}(\tilde{y})\right)\right)\right)^2} = 0 \qquad (A99)$$

Equation $(A99)$ shows that $w_{FB}^{fo\prime}(\tilde{y}) = 0$. Furthermore, in the domain of losses the stochastic contract L^* is given to the agent, that contract depends only on parameters of the model and on the constant payment $w_{FB}^{fo}(\tilde{y})$ for given $\tilde{y}$.

Next, I show that the first-best contract necessarily consists of two segments, one paying $w_{FB}(y) = L^*$ and another one paying $w_{SB}(y)$ satisfying $(A88)$. Suppose that the optimal contract pays $w_{FB}^{fo}(y)$ satisfying $(A88)$. Since $\mathbb{E}(L^*) < w_{FB}^{fo}(y)$, the principal can profitably deviate from that solution by paying L^* at the lower end of the output space. L^* will be evaluated as a loss when $w_{FB}^{fo}(y)$ satisfying $(A88)$ is taken as reference point. Due to the convexity of $u(\cdot)$ in the domain of losses (Assumption 4), the agent will be willing to accept the contract.

Suppose instead that the optimal contract pays L^* for all y. When compensated in such way, the agent evaluates the outcome $w_{FB}^{fo}(\tilde{y})$ for a given $\tilde{y}$ as a gain when $w_{FB}(y) = 0$ is taken as the reference point. However, the fact that this payment only applies with some probability entails that is not providing properly insuring the agent. This is a contradiction since at the optimum the agent's participation compatibility constraint binds.

To conclude the proof, I show that there exists a unique threshold output below which L^* is paid and above which $w_{FB}(y)$ satisfying $(A88)$ is paid. The optimal incentives of paying L^*, subject to the participation constraint is given by

$$\frac{1}{1+\left(1-2\overline{L}\right)\frac{\left(u(w_{FB}(\tilde{y}))-\lambda u(-w_{FB}(\tilde{y}))\right)}{w_{FB}(\tilde{y})}}=\nu, \qquad (A100)$$

From $(A100)$ note that for $\nu > 0$, $\overline{L} = w_{FB}(\tilde{y})p_1^* < \frac{1}{2}$ which in turn implies $w_{FB}(\tilde{y}) < 1$ since $p_1^* \in \left[\frac{1}{2}, 1\right]$. Rewrite $(A100)$ using $(A97)$ as

$$\frac{1}{1+\left(\frac{2\eta\left(u(w_{FB}(\tilde{y}))-\lambda u(-w_{FB}(\tilde{y}))\right)}{w_{FB}(\tilde{y})}-\eta\left(u(w_{FB}(\tilde{y}))-\lambda u(-w_{FB}(\tilde{y}))\right)-w_{FB}(\tilde{y})\right)}=\nu, \qquad (A101)$$

where the left-hand side of $(A101)$ strictly decreases in $w_{FB}(\tilde{y})$ since $\frac{\partial}{\partial w_{FB}(\tilde{y})}\left(\frac{w_{FB}(\tilde{y})}{2\eta\left(u(w_{FB}(\tilde{y}))-\lambda u(-w_{FB}(\tilde{y}))\right)}\right) < 0$ and $w_{FB}(\tilde{y}) < 1$. Since the right-hand side of that equation is constant, there exists an output level $\tilde{y}_{k1} \in (\underline{y}, \overline{y})$ such that $(A101)$ holds. For all output levels $y < \tilde{y}_{k1}$ then $L^* = (w_{FB}(\tilde{y}_k), 1 - p_1^*; 0, p_1^*)$ is given, otherwise $w_{FB}(y)$ satisfying $(A88)$ is paid. ■

Appendix B.

B.1 Other salience-based reference points.

The *min-max* rule implies that individuals are bold: They take as the reference point the minimum value of a set consisting of the maximum outcome of each alternative. In light of the example given above in which the agent received w_1 and w_2, this rule implies that his reference point is also 100.

As in the previous corollary, the reference point results from considering the agent's choices and the implied alternatives. It turns out that the resulting reference point is identical to that resulting from the min-max rule, and as a consequence the same optimal contract motivates the agent. Corollary B.1. presents this result.

Corollary B.1. *Under the assumptions A1, A2, A4, that the agent's preferences are characterized by* (2)*, and the max-min rule, the agent's reference point and the second-best contract are identical to those presented in Corollary 5.*

Proof. In the present framework the agent makes, at most, two choices: deciding whether to accept the contract or not, and choosing amount of effort if the contract is accepted. These choices generate three candidates for reference points under the max-min rule: rejecting the contract and obtaining $\overline{U} \geq 0$, obtaining the maximum payment resulting from accepting the contract after choosing e_H, and obtaining the maximum payment resulting from accepting the contract after choosing e_L.

Since the agent's preferences are characterized by (2) a contract similar to that described in Proposition 2 remains to be optimal although with a different transition from losses to gains. Denote that contract by $w_{SB}(y)$. From equation $(A5)$ it can be established that the maximum amount paid by the second-best contract is $\max\{w_{SB}(y)\} = r + \max\left\{h'\left(\frac{1}{v+\gamma\left(1-\frac{f(y|e_L)}{f(y|e_H)}\right)}\right)\right\}$. Assumption 2 states that $\frac{\partial}{\partial y}\left(\frac{f(y|e_L)}{f(y|e_H)}\right) < 0$ and Assumption 4 that $h'(\cdot)$ is a decreasing function. Hence, the expression $\max\left\{h'\left(\frac{1}{v+\gamma\left(1-\frac{f(y|e_L)}{f(y|e_H)}\right)}\right)\right\}$ is larger when e_H is chosen than when e_L is chosen, implying that the second-best contract from Proposition 2 yields higher payments under high effort. Moreover, an optimal payment generating utility $\overline{U} \geq 0$, is the first-best contract, $w_{FB}(y)$ from Proposition 2 (i).

Next, I show that $w_{FB}(y) \leq \max\{w_{SB}(y)\}$ for any e. Suppose instead that $w_{FB}(y) > \max\{w_{SB}(y)\}$. In that case the agent would not accept the contract since $CPT(e_H, w_{FB}(y), r) = \overline{U} > CPT(e_H, \max\{w_{SB}(y)\}, r)$. This is a contradiction since at the optimum the participation constraint binds. Hence, it must be $w_{FB}(y) \leq \max\{w_{SB}(y)\}$. Moreover, since the incentive compatibility constraint binds at the optimum it must be that $w_{FB}(y) < \max\{w_{SB}(y)\}$ for e_H, otherwise the second-best contract would not provide rewards to exert high effort. As a result, the min-max rule thus implies that $r = w_{FB}(y)$.

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

The welfare level implied by the first-best contract, $\overline{U}$, is lower than that implied by the maximum amount given to the agent by the second-best contract. Otherwise, the second-best contract would not include rewards for good performance and would not be effective in incentivizing the agent to exert high effort. Hence, under the max-min rule $r = w_{FB}(y)$ a contract that implies $\overline{U}$.

An important property of the contract presented in Corollaries 5 and B.1 is that the rewards and punishments that it conveys locate the agent in the domain of losses and in the domain of gains, respectively. That is because the reference point under these two rules, $r = w_{FB}(y)$, is such that rewards with respect to the first-best contract necessarily entail $w_{SB}(y) \geq r$, while punishments necessarily entail $w_{SB}(y) < r$. This feature contrasts the finding, presented in Corollary 1, that the second-best contract under an exogenous reference point elicits high effort by implementing punishments for low performance in the domain of gains. Thus, the different reference point rule changed the way in which the second-best contract imparts incentives.

I consider also consider a salience-based rule that states that the output level realizing with the highest probability is adopted as the agent's reference point. This rule is called $w_{SB}(y)$ *at max* P. The following corollary describes the agent's reference point and the optimal contract when the agent uses $w_{SB}(y)$ at max P.

Corollary B.2 *Under assumptions A1, A2, A4, that the agent's preferences are characterized by* (2), *and the* $w_{SB}(y)$ *at max P rule, the agent's reference point is* $r = w_{SB}(y_p)$ *where* y_p *satisfies* $f(y_p|e_H) = \max\{f(y|e_H)\}$ *and the second-best contract,* $w_{SB}(y)$, *is constant and pays the minimum possible in* $y < \hat{y}_p$, *exhibits a bonus at* $y = y_p$, *and increases in performance in* $y > \hat{y}_p$.

Proof. Let $y_p \in [\underline{y}, \overline{y}]$ be a performance level satisfying $f(y_p|e) = \max\{f(y|e)\}$. If $f(y_p|e)$ is multimodal, let y_p be the smallest output level satisfying $f(y_p|e) = \max\{f(y|e)\}$. The continuity of the support of $f(y|e)$, the set $[\underline{y}, \overline{y}]$, implies that y_p is the output realization that attains the highest probability. Since the agent's preferences are characterized by prospect theory a contract similar to that described in Proposition 2 remains to be optimal although with a different transition from losses to gains. Denote that contract by $w_{SB}(y)$. The $w_{SB}(y)$ *at max* P rule entails that $r = w_{SB}(y_p)$.

The solution presented in Proposition 2 can be adjusted to account for this reference point as follows. Let $\hat{y}_p \in [\underline{y}, \overline{y}]$ satisfy:

$$\frac{1}{\frac{\lambda u(w_{SB}(\hat{y}_p))}{w_{SB}(\hat{y}_p)}} = \nu + \gamma\left(1 - \frac{f(\hat{y}_p|e_L)}{f(\hat{y}_p|e_H)}\right), \qquad (B1)$$

Under the max-min rule, the second-best contract consists of $w_{SB}(y)$ satisfying the following first-order condition

$$\frac{1}{u'\big(w_{SB}(y)-w_{SB}(y_p)\big)}=\nu+\gamma\left(1-\frac{f(y|e_H)}{f(y|e_L)}\right), \qquad (B2)$$

if $y > \hat{y}_p$, and

$$w_{SB}=\begin{cases}0 \; if \; y<\hat{y}_p,\\ L \; if \; y=\hat{y}_p,\end{cases} \qquad (B3)$$

if $y \leq \hat{y}_p$.

Finally, I demonstrate that $\hat{y}_p = y_p$. Suppose instead that $\hat{y}_p < y_p$. In that case, the principal is over-insuring the agent from risk in $y \in [\hat{y}_p, y_p]$ by offering $w_{SB}(y)$ satisfying $(B2)$ in a segment where he is risk seeking due to diminishing sensitivity (Assumption 4). The principal could increase profits by exposing the agent to large amounts of risk by setting $w_{SB}(y) = 0$ for all $y < y_p$, including $y \in [\hat{y}_p, y_p]$, and the agent would accept such contract due to his risk seeking attitudes emerging from diminishing sensitivity. Notice that such a change in the contract's transfers does not change the agent's reference point, since y_p does not depend on $w_{SB}(y)$ but on the probability distribution $f(y|e)$, which is exogenous to the principal's choice.

Next, suppose that $\hat{y}_p > y_p$. In that case the agent is being exposed to large amounts of risk in $y \in [y_p, \hat{y}_p]$ by obtaining $w_{SB}(y) = 0$ in 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 $w_{SB}(y)$ satisfying $(B2)$ for $y \geq y_p$. Hence, it must be that $\hat{y}_p = y_p$. ■

The reference point found under the $w_{SB}(y)$ at max P rule is determined by the density function $f(y|e)$, which is exogenous to the principal's choice. Specifically, the agent's reference point corresponds to the payment given by the second-best contract at the output level realizing with the highest probability, whatever that payment is. Anticipating that, the principal offers a contract in which transfers below that point are set as small as possible. The rationale behind the optimality of that solution is similar to the one given in intuition of Proposition 2 (ii): Loss aversion and diminishing sensitivity generate powerful incentives to motivate the agent to exert high effort despite the agent being paid the lowest amount possible. In the absence of these biases, i.e. in the domain of gains, the incentive scheme has to be high-powered and, due to the agent's risk aversion it must increase. The change in incentives at $y = y_p$ generates a bonus.

B.2 Aspirations-based utility

This appendix investigates the optimal incentive scheme under aspirations-based decision-making a la Diecidue and van de Ven (2008). Specifically, the agent's preference is given by

$$\mathbb{E}\big(U(e,w(y),a)\big) = \int_a^{\overline{y}}\Big(u\big(w(y)\big)+\tau\Big)f(y|e)\mathrm{d}y+\int_{\underline{y}}^{a}\Big(u\big(w(y)\big)-\delta\Big)f(y|e)\mathrm{d}y - c(e). \qquad (B4)$$

Where $a \in [\underline{y}, \overline{y}]$ is the agent' aspiration level, $\tau > 0$ weighs the relevance probabilities of success, and $\delta > 0$ weighs the relevance of probabilities of failure. Notice that the relationship between τ and δ determines whether success or failure is more prominent in the decision maker's utility. I impose $\tau > \delta$, but, as it will be evident later on, the results presented in this section do not depend on the relationship between τ and γ.

When facing an agent with an aspirations-based utility a la Diecidue and van de Ven (2008), the principal implements an incentive scheme that is identical to the traditional optimal solution attributed to Holmström (1979). The following Proposition presents this result.

Proposition B.1 *Under assumptions, A1, A2, A3, and that the agent's preferences are characterized by* $(B4)$ *the optimal contract is identical to that in Proposition 1.*

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

$$\mathcal{L} = \big(S(y) - w(y)\big)f(y|e_H)$$

$$+\nu\Big(\theta\big(u\big(w(y)\big) + \tau\big)f(y|e_H) + (1-\theta)u\big(w(y)\big) - \delta\big)f(y|e_H) - c - \overline{U}\Big) \qquad (B5)$$

$$+\gamma\Big(\theta\big(u\big(w(y)\big) + \tau\big)\big(f(y|e_H) - f(y|e_L)\big) + (1-\theta)u\big(w(y)\big) - \delta\big)\big(f(y|e_H) - f(y|e_L)\big) - c\Big).$$

Where θ is an indicator function taking value $\theta = 1$ if $y \geq a$ and $\theta = 0$ otherwise. Pointwise maximization with respect to $w(y)$ gives

$$-f(y|e_H) + \nu\left(\theta u'^{(w(y))}f(y|e_H) + \lambda(1-\theta)u'\left(w(y)\right)f(y|e_H)\right) \qquad (B6)$$
$$+\gamma\left(\theta u'\big(w(y)\big)\big(f(y|e_H) - f(y|e_L)\big) + \lambda(1-\theta)u'\big(w(y)\big)\big(f(y|e_H) - f(y|e_L)\big)\right) = 0.$$

Denoting by $w_{SB}^{fo}(y)$ the transfer satisfying $(B6)$, and after some manipulations then the following expression is obtained

$$\frac{1}{u'\big(w_{SB}^{fo}(y)\big)} = \nu + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right), \qquad (B7)$$

if $y < a$, and

$$\frac{1}{\lambda u'\left(w_{SB}^{fo}(y)\right)} = \nu + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right), \qquad (B8)$$

if $y \geq a$. Using the inverse function $h(\cdot)$ rewrite $(B7)$ as:

$$w_{SB}^{fo}(y) = h'\left(\frac{1}{\nu + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right)}\right), \qquad (B9)$$

and $(B8)$ as:

$$w_{SB}^{fo}(y) = \left(\frac{1}{v + \gamma \left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right)} \right). \qquad (B10)$$

Notice that $(B9)$ and $(B10)$ are identical so $w_{SB}^{fo}(y)$ satisfying $(B9)$ suffices for all y. Also, equation $(B9)$ is identical to the second-best contract characterized in Proposition 1. ■

The significance of Proposition B.1 is to show that despite the fact that the worker's preferences exhibit a discontinuity at the aspiration level, such feature does not change in a fundamental way the optimal design of incentives. The intuition is that the discontinuity in utility used in the model of Diecidue and van de Ven (2008) does not imply changes in marginal utility with respect to the standard expected utility framework. Hence, marginal benefits and costs from exerting effort remain similar, and the standard solution from Holmström (1979) presented in Proposition 1 (ii) suffices to optimally incentivize the agent.

An extension of the model of Diecidue and van de Ven (2008) that includes reference dependence, as is for example done in Levy and Levy (2009), would lead to contracting modalities that deviate from standard solutions. In that paper, the individual's aspiration level is also assumed to be a reference point, so that the discontinuity in utility is enhanced with the properties diminishing sensitivity and loss aversion. These two properties fundamentally change marginal utility, and thus the tradeoffs faced by the agent. In such case, the optimal contracts in Propositions 2 and 3 suffice to incentivize the agent.

B.3. 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 assumed to be his certainty equivalent. Importantly, this certainty equivalent includes the agent's psychological utility component.

As with the previous disappointment models, this model can be treated as an extension of prospect theory. Formally, the agent's preferences can be represented as,

$$DA\big(e, w(y)\big) = \int_{\underline{y}}^{\overline{y}} w(y) f(y|e) dy + CPT(e, w(y), CE)) - c(e), \qquad (B11)$$

where $CE$is a fixed monetary quantity inducing $DA(e_H, w(y))) = CE$ for given $w(y)$. Again, the key assumption behind this preference representation is that consumption utility is linear. Notice that, due to the assumption that the value function is strictly increasing (Assumption 4) and is thus invertible, the agent's certainty equivalent exists.

The first-best and second-best contracts when the agent's risk preferences are captured by $(B7)$ are presented in the next corollary.

Corollary B.2 *Under assumptions A1, A2, A4 and that the agent's preferences are characterized by* $(B11)$, *there exist unique output levels* $y_{c1}, y_{c2} \in [\underline{y}, \overline{y}]$ *such that:*

i) *The first-best contract, $w_{FB}(y)$, is similar to that of Proposition 4 (i) with y_{m1} replaced by y_{c1}.*

ii) *The second-best contract, w_{SB}, is similar to that of Proposition 4 (ii) with y_{m2} replaced by y_{c2}.*

Proof. Since the agent's preferences are characterized by $(B11)$ a contract similar to that described in Proposition 4 remains to be optimal although with a potentially different transition from losses to gains since $r = CE$ and $CE = \widetilde{w}_{SB}$ might not hold. From Proposition 4 it follows that the second-best is the contract is

$$w_{SB} = CE + h'\left(\frac{1}{v + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right)} - 1\right), \qquad (B12)$$

for the domain of gains and

$$w_{SB} = \begin{cases} 0 \; if \; y < y_{c2} \\ L \; if \; y = y_{c2}. \end{cases} \qquad (B13)$$

for the domain of losses. Where $y_{c2} \in \left[\underline{y}, \overline{y}\right]$ is an output level satisfying $w_{SB}(y_{c2}) = CE, L$ is the lottery-like payment $L = (CE, p; 0, 1-p)$, and $p \in (0,1)$ an interior probability satisfying $CE - \lambda u\left(CE - w_{SB}^{fo}(y)\right) = pCE - (1-p)\lambda u(CE)$ for $w_{SB}^{fo}(y) = CE - h'\left(\frac{1}{\lambda\left(v+\gamma\left(1-\frac{f(y|e_L)}{f(y|e_H)}\right)\right)} - 1\right)$.

Instead, the first-best contract is given by

$$w_{SB} = CE + h'\left(\frac{1}{v} - 1\right), \qquad (B14)$$

In the domain of gains and,

$$w_{SB} = \begin{cases} 0 \; if \; y < y_{c1}, \\ L \; if \; y = y_{c1}. \end{cases} \qquad (B15)$$

For the domain of losses. Where $y_{c1} \in \left[\underline{y}, \overline{y}\right]$ is an output level such that $w_{FB}(y_{c1}) = CE$ and $p \in (0,1)$ an interior probability satisfying $CE - \lambda u\left(CE - w_{FB}^{fo}(y)\right) = pCE - (1-p)\lambda u(CE)$ for $w_{SB}^{fo}(y) = CE - h'\left(\frac{1}{\lambda(v)} - 1\right)$.

The proof of the uniqueness of y_{c1} and y_{c2} and the fact that they are interior points of the output set exhibits a similar reasoning as that followed in parts iii) and iv) in Proposition 4. It is cost-effective to expose the agent to losses, but a contract consisting of only losses will be rejected. ■

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 Proposition 4. This similarity is, again, due to the fact that his risk preferences are fully determined by the value function. 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 included in the contracts described in Corollary B.2. Intuitively, a (globally) risk averse agent with risk preferences representable by (6), must exhibit $CE < \widetilde{w}_{SB}(y)$. 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, $CE = \widetilde{w}_{SB}(y)$. Hence, $y_{c1} < y_{m1}$ and $y_{c2} < y_{m2}$. 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.[12]

[12] For the sake of completeness, note that this comparison is different from that given in Corollary 4. There different degrees of utility curvature are considered. Here, these comparisons emerge when similar levels of utility curvature emerge but different reference points are considered.

Appendix C. Proofs of the extensions presented in Section 5.

Proposition 6.

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

$$\mathcal{L} = \left(-\lambda_p\left(r_p + w(y) - S(y)\right)\right) f(y|e_H)$$
$$+\nu\big(\theta u(w(y) - r) f(y|e_H) - \lambda(1-\theta)u\big(r - w(y)\big) f(y|e_H) - c - \overline{U}\big) \qquad (C1)$$
$$+\gamma\big(\theta u(w(y) - r)\big(f(y|e_H) - f(y|e_L)\big) - \lambda(1-\theta)u\big(r - w(y)\big)\big(f(y|e_H) - f(y|e_L)\big) - c\big).$$

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

$$-f(y|e_H)\lambda_p + \nu\left(\theta u'(w(y) - r) f(y|e_H) + \lambda(1-\theta)u'\left(r - w(y)\right) f(y|e_H)\right) \qquad (C2)$$
$$+\gamma\left(\theta u'(w(y) - r)\big(f(y|e_H) - f(y|e_L)\big) + \lambda(1-\theta)u'\big(r - w(y)\big)\big(f(y|e_H) - f(y|e_L)\big)\right) = 0.$$

Denoting by $w_{SB}^{fo}(y)$ the transfer satisfying $(c2)$ the following expression is obtained and after some manipulations

$$\frac{\lambda_p}{u'\left(w_{SB}^{fo}(y) - r\right)} = \nu + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right), \qquad (C3)$$

if $\theta = 1$, and

$$\frac{\lambda_p}{\lambda u'\left(r - w_{SB}^{fo}(y)\right)} = \nu + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right), \qquad (C4)$$

if $\theta = 0$. Using the inverse function $h(\cdot)$ rewrite $(C3)$ as:

$$w_{SB}^{fo}(y) = r + h'\left(\frac{\lambda_p}{\nu + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right)}\right), \qquad (C5)$$

and $(C4)$ as:

$$w_{SB}^{fo}(y) = r - h'\left(\frac{\lambda_p}{\lambda\left(\nu + \gamma\left(1 - \frac{f(y|e_L)}{f(y|e_H)}\right)\right)}\right). \qquad (C6)$$

As in proposition 2 it can be shown that the solution given by equation $(C6)$ can be replaced by a lottery paying $L = (r, p, 0, 1-p)$. Where $p \in (0,1)$ satisfies

$$-\lambda u\left(r - w_{SB}^{fo}(y)\right) = -(1-p)\lambda u(r), \qquad (C7)$$

and $w_{SB}^{fo}(y)$ satisfies $(C4)$. That lottery does not change participation and incentive compatibility constraint and is more cost effective for the principal since from $(C7)$

$$-\lambda u\left(r - w_{SB}^{fo}(y)\right) < -\lambda u\big((1-p)r\big), \qquad (C8)$$

Implying $w_{SB}^{fo}(y) > pr$.

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

$$\frac{\lambda_p}{\frac{\lambda u(r)}{r}} = \nu + \gamma\left(1 - \frac{f(\hat{y}_s|e_L)}{f(\hat{y}_s|e_H)}\right). \qquad (C9)$$

When $\hat{y}_s > y$, the scheme pays the minimum admissible. That is because when offered L the agent's utility can be expressed as

$$\mathbb{E}\big(CPT(L)\big) = -\lambda\left(1 - \frac{\overline{L}}{r}\right)\big(u(r)\big) - c, \qquad (C10)$$

which is linear in $\overline{L}$, the expected value of the contract. So, changes in $\overline{L}$ do not affect the agent's marginal utility and the principal can afford to pay the minimum possible without changing the agent's incentive. Instead if $\hat{y}_s < y$, the agent's payment can be large enough to bring him to the domain of gains without changing the agent's incentives. In that domain, the principal should be paid $w_{SB}^{fo}(y)$ satisfying $(C3)$. The solution to the principal's problem is

$$w_{SB}(y) = \begin{cases} 0 \ if \ y < \hat{y}_s, \\ L \ if \ y = \hat{y}_s, \end{cases} \quad (C11)$$

and $w_{SB} = w_{SB}^{fo}(y)$ from $(C3)$ if $y > \hat{y}_s$.

Let $S(y) \geq r_p + w(y)$. Since principal's and agent's objective functions are identical to the case studied in Proposition 2, that solution remains optimal.

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

There are two relevant cases. If $\hat{y}_s < \hat{y}_p$ then both agent and principal are in the domain of losses in $\hat{y}_s < y$ and $(C11)$ is offered. In $\hat{y}_s < y < \hat{y}_p$, the principal is in the domain of losses, while the agent is in the domain of gains. As a consequence, the principal offers insurance to the agent by paying $w_{SB}^{fo}(y)$ satisfying $(C3)$. Finally, for $y > \hat{y}_p$ the principal offers $w_{SB}^{fo}(y)$ satisfying $(A5)$. Since λ_p is absent in $(A5)$ but present in $(C3)$, there exists a kink in $y > \hat{y}_p$.

Let $\hat{y}_s > \hat{y}_p$. Both agent and principal are in the domain of losses in $\hat{y}_p < y$ and $(C11)$ is offered. In $\hat{y}_p < y < \hat{y}_s$, the agent is in the domain of losses, while the principal is in the domain of gains. As a consequence, the principal offers insurance to the agent by paying $(A12)$, which is identical to $(C11)$ except for the fact that the transition to gains can differ. Finally, for $y > \hat{y}_s$ the principal offers $w_{SB}^{fo}(y)$ satisfying $(A5)$. There is no kink in this case. ■

Proposition 7.

The agent with λ_H faces the following adverse selection constraint,

$$\int_{\hat{y}_H}^{\overline{y}} u(w(y)^H - r) f(y|e_H) \mathrm{d}y - \lambda_H \int_{\underline{y}}^{\hat{y}_H} u(r - w(y)^H) f(y|e_H) \mathrm{d}y - c$$
$$\geq \max_{e \in \{e_L, e_H\}} \left\{ \int_{\hat{y}_L}^{\overline{y}} u(w(y)^L - r) f(y|e) \mathrm{d}y - \lambda_H \int_{\underline{y}}^{\hat{y}_L} u(r - w(y)^L) f(y|e) \mathrm{d}y - c(e) \right\}, \quad (C12)$$

the moral hazard incentive constraint,

$$\int_{\hat{y}_H}^{\overline{y}} u(w(y)^H - r) f(y|e_H) \mathrm{d}y - \lambda_H \int_{\underline{y}}^{\hat{y}_H} u(r - w(y)^H) f(y|e_H) \mathrm{d}y - c$$
$$\geq \int_{\hat{y}_H}^{\overline{y}} u(w(y)^H - r) f(y|e_L) \mathrm{d}y - \lambda_H \int_{\underline{y}}^{\hat{y}_H} u(r - w(y)^H) f(y|e_L) \mathrm{d}y, \quad (C13)$$

and the participation constraint

$$\int_{\hat{y}_H}^{\overline{y}} u(w(y)^H - r) f(y|e_H) \mathrm{d}y - \lambda_H \int_{\underline{y}}^{\hat{y}_H} u(r - w(y)^H) f(y|e_H) \mathrm{d}y - c \geq \overline{U}. \quad (C14)$$

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

$$\int_{\hat{y}_L}^{\overline{y}} u(w(y)^L - r) f(y|e_H) \mathrm{d}y - \lambda_L \int_{\underline{y}}^{\hat{y}_L} u(r - w(y)^L) f(y|e_H) \mathrm{d}y - c$$
$$\geq \max_{e \in \{e_L, e_H\}} \left\{ \int_{\hat{y}_H}^{\overline{y}} u(w(y)^H - r) f(y|e) \mathrm{d}y - \lambda_L \int_{\underline{y}}^{\hat{y}_H} u(r - w(y)^H) f(y|e) \mathrm{d}y - c(e) \right\}, \quad (C15)$$

the moral hazard incentive constraint,

$$\int_{\hat{y}_L}^{\overline{y}} u(w(y)^L - r) f(y|e_H) \mathrm{d}y - \lambda_L \int_{\underline{y}}^{\hat{y}_L} u(r - w(y)^L) f(y|e_H) \mathrm{d}y - c$$
$$\geq \int_{\hat{y}_L}^{\overline{y}} u(w(y)^L - r) f(y|e_L) \mathrm{d}y - \lambda_L \int_{\underline{y}}^{\hat{y}_L} u(r - w(y)^L) f(y|e_L) \mathrm{d}y, \quad (C16)$$

and the participation constraint

$$\int_{\hat{y}_L}^{\overline{y}} u(w(y)^L - r) f(y|e_H) \mathrm{d}y - \lambda_L \int_{\underline{y}}^{\hat{y}_L} u(r - w(y)^L) f(y|e_H) \mathrm{d}y - c \geq \overline{U}. \quad (C17)$$

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

$$\begin{aligned} CPT(\hat{e}, w(y)^H, r, \lambda_L) \\ &= \int_{\hat{y}_H}^{\overline{y}} u(w(y)^H - r) f(y|\hat{e}) \mathrm{d}y + (\lambda_H - \lambda_L) \int_{\underline{y}}^{\hat{y}_H} u(r - w(y)^H) f(y|\hat{e}) \mathrm{d}y - c(\hat{e}) \\ &\geq CPT(\hat{e}, w_H, r, \lambda_H) + (\lambda_H - \lambda_L) \int_{\underline{y}}^{\hat{y}_H} u(r - w(y)^H) f(y|\hat{e}) \mathrm{d}y. \qquad (C18) \end{aligned}$$

Since $\lambda_H > \lambda_L$ and $r > w(y)^H$ in the domain of losses, 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 $\hat{e}$,

$$\begin{aligned} CPT(\hat{e}, w(y)^L, r, \lambda_H) &= \int_{\hat{y}_L}^{\overline{y}} u(w(y)^L - r) f(y|\hat{e}) \mathrm{d}y - (\lambda_H - \lambda_L) \int_{\underline{y}}^{\hat{y}_L} u(r - w(y)^L) f(y|\hat{e}) \mathrm{d}y - c(\hat{e}) \\ &\geq CPT(\hat{e}, w(y)^L, r, \lambda_L) - (\lambda_H - \lambda_L) \int_{\underline{y}}^{\hat{y}_L} u(r - w(y)^L) f(y|\hat{e}) \mathrm{d}y. \qquad (C19) \end{aligned}$$

From $(C19)$ it is evident that engaging in that strategy is not profitable. Express the adverse selection constraints $(C12)$ and $(C15)$ in terms of informational rents in

Use $(C18)$ and $(C19)$ to rewrite the adverse selection constraints in $(C12)$ and $(C15)$ as follows:

$$\begin{aligned} \int_{\hat{y}_H}^{\overline{y}} u(w(y)^H - r) f(y|e_H) \mathrm{d}y - \lambda_H \int_{\underline{y}}^{\hat{y}_H} u(r - w(y)^H) f(y|e_H) \mathrm{d}y - c \\ \geq \max_{e \in \{e_L, e_H\}} \left\{ CPT(e, w(y)^L, r, \lambda_L) - (\lambda_H - \lambda_L) \int_{\underline{y}}^{\hat{y}_L} u(r - w(y)^L) f(y|e) \mathrm{d}y \right\}, \qquad (C20) \end{aligned}$$

and

$$\begin{aligned} \int_{\hat{y}_L}^{\overline{y}} u(w(y)^L - r) f(y|e_H) \mathrm{d}y - \lambda_L \int_{\underline{y}}^{\hat{y}_L} u(r - w(y)^L) f(y|e_H) \mathrm{d}y - c \\ \geq \max_{e \in \{e_L, e_H\}} \left\{ CPT(e, w(y)^H, r, \lambda_H) + (\lambda_H - \lambda_L) \int_{\underline{y}}^{\hat{y}_H} u(r - w(y)^H) f(y|e) \mathrm{d}y \right\}, \qquad (C21) \end{aligned}$$

respectively.

Note that $(C14)$ and $(C20)$ imply $(C17)$, so it must be that $(C17)$ slacks at the optimum while $(C14)$ binds. Moreover, since $(C18)$ and $(C19)$ show that the agent with λ_L derives profits when mimicking the agent with λ_H, then $(C20)$ and is always strictly satisfied and $(C21)$ binds at the optimum. Finally, denote by $w_{SB}(y)^i$ the contract from Proposition 2. From the proof of that proposition it is known that e_H generates high effort. This reduces the number of constraints to:

$$CPT(e, w_{SB}(y)^H, r, \lambda_H) = \overline{U}, \qquad (C22)$$

and

$$CPT(e, w_{SB}(y)^L, r, \lambda_L) = CPT(e, w_{SB}(y)^H, r, \lambda_H) + (\lambda_H - \lambda_L) \int_{\underline{y}}^{\hat{y}_H} u(r - w_{SB}(y)^H) f(y|e) \mathrm{d}y. \quad (C23)$$

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