ECONtribute
MARKETS & PUBLIC POLICY

ECONtribute
Discussion Paper No. 451

Moral Hazard and Disappointment Aversion

Alex Gershkov Benny Moldovanu Philipp Strack

November 2026 www.econtribute.de

UNIVERSITÄT BONN
UNIVERSITÄT ZU KÖLN

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2126/2 – 390838866 – is gratefully acknowledged.

Moral Hazard and Disappointment Aversion

Alex Gershkov, Benny Moldovanu and Philipp Strack[*†]

November 10, 2026

Abstract

We study the classical moral-hazard problem where both the agent and the principal are disappointment and risk averse in the sense of Gul [1991]. If only the agent is risk averse, the optimal contract consists of at most three levels of payment: a standard fixed wage for middle outcomes, equivalent to the agent's total perceived compensation including bonus and penalties, a penalty wage/termination clause for low outcomes, and a bonus paid only for the highest outcome. When both agent and principal are disappointment averse, the optimal contract consists of five distinct parts, three as above and two linearly increasing in outcome.

1 Introduction

In this paper we study the classical moral hazard problem with hidden actions and only change the specification of preferences for the agent and for the principal. These are assumed to have a piece-wise linear utility over monetary outcomes and to be disappointment- and risk-averse in the sense of Gul [1991]. An important aspect of human psychology captured by Gul's preferences and relevant for moral-hazard situations is that an agent who decides how much effort to exert anticipates the potential disappointment he will feel if his realized utility is less than his reference utility. This is relevant in incentive

[*]Gershkov wishes to thank the Israel Science Foundation for financial support (grant 1972/26). Moldovanu acknowledges financial support from the DFG (German Research Foundation) via Germany's Excellence Strategy - EXC 2047/1 - 390685813, EXC 2126/1-390838866 and the CRC TR-224 (project B01). Strack gratefully acknowledges support from the Cowles foundation. Gershkov: Department of Economics and the Federmann Center for the Study of Rationality, The Hebrew University of Jerusalem, and Department of Economics, University of Southampton, alexg@huji.ac.il; Moldovanu: Department of Economics, University of Bonn, mold@uni-bonn.de; Strack, Department of Economics, Yale University, philipp.strack@gmail.com.

[†]We wish to thank David Dillenberger, Daniel Gottlieb, Yoram Halevy, Johannes Hörner, Ryota Iijima, Botond Köszegi, Daniel Kraehmer, Stephan Lauermann, Colin Raymond, Ilya Segal, Rani Spiegler, Rakesh Vohra and all participants at the "Behavioral Mechanism Design Conference", Bonn 2026, for their very helpful remarks.

contracts because any reward provided by the principal conditional on a good outcome brings with it the possibility of disappointment - for both agent and principal - if this outcome does not realize.

Our results yield a simple explanation for the optimality and robustness of contracts that are neither too performance-sensitive, nor too minimalist (i.e., consisting of a basic wage and a bonus): if only the agent is disappointment and risk averse, then the optimal contract features a base wage, leads to the firing of the agent after bad outcomes and pays a fixed bonus after good outcomes; if both agent and principal are disappointment and risk averse, the optimal contract consists of several regions where the contract's slope is either zero or one. Although incentives can always be established in our framework with a two-tiered, pure bonus scheme, such a scheme turns the basic wage into a disappointing one because the overall utility (that includes a performance sensitive part) is then higher than the basic salary. This disappointment - more heavily weighted by the agent - calls for a higher compensation for exceptional performance, yielding an overall higher level compensation. The principal can both incentivize the agent and offer additional insurance to our risk-averse agent (while saving money !) by creating a contract with three wage levels, where the middle tier is geared towards the overall level of utility experienced by the agent from the contract (including bonus and penalties). In this way, only the penalty associated with low outcomes leads to disappointment and the (relatively lower) bonus leads to elation. When the principal is also disappointment- and risk averse, the above scheme exposes her to too much risk, and this feature leads to additional contract regions where the agent fully bears the risk.

Gul's theory of disappointment aversion is one of the most influential models in decision making under risk[1]. It posits that individuals evaluate the various outcomes of a risky action relative to an endogenous reference point: the ex-ante value of the action itself. Such a reference point balances the negative emotional responses to outcomes falling short of expectations - these are heavier weighted - against the positive emotional responses to outcomes exceeding expectations[2]. The theory is simple and is derived axiomatically: Gul replaces the independence axiom, well-known from its role in the characterization of expected utility, with a weaker form that only requires independence to hold on a smaller class of lotteries that preserve certain elation and disappointment sets. Gul's general theory encompasses expected utility as a special case, and is consistent with observed behavior in the famous Allais paradox. Moreover, it is, in a sense, a minimalist extension of expected utility since it only adds to it a unique real parameter $\beta > 0$. In our setting

[1]For alternative formulations of disappointment aversion see Delquie and Cillo [2006], Jia et al[2001] and Loomes and Sugden [1986]

[2]Observe the implicit nature of this specification. For an explicit representation of the preferences see Cerreia-Vioglio et al. [2020].

with piecewise linear utility over monetary outcomes, Gul's certainty equivalent of an uncertain outcome equals the (uniquely defined) value for which the expected elation divided by the expected disappointment equals $(1+\beta) > 1$. Note that the expectation of a random variable satisfies the analogous property where the ratio is 1.

We start our study with a simple geometric condition that is both necessary and sufficient for an action to be implementable for any possible cost vector. For this to happen, it is necessary and sufficient that the outcome probability vector associated to that action is a vertex of the convex hull of the set of probability vectors associated to all actions. This condition is equivalent to the assumption that, by repeatedly observing the outcomes of the agent's action, the principal can reject the hypothesis that the agent is taking a specific action if in fact the agent is not taking that action. Moreover, we show that binary contracts are always sufficient for implementation.

Our first main result (Theorem 1) is about the case where the principal is risk neutral but the agent is disappointment- and risk-averse. Under the well-known monotone likelihood ratio property (MLRP), the optimal contract is piece-wise linear with discrete jumps among its various parts, and is supported on at most three levels of payment: 1) A standard wage for middle outcomes, equivalent from the agent's point of view to the complete lottery formed by his compensation packages, including possible bonuses and penalties; 2) A (possibly random) malus or penalty wage/termination for low outcomes, where randomization occurs only at the highest outcome in this group of low outcomes and is between two wage levels[3] ; 3) A bonus that is only paid for the highest possible outcome.

The above findings are supported by available empirical evidence. Bonus-penalty systems that aim to align the agents' behavior with a desired goal are often observed in contracts beyond employment ones. For example, higher quality/delivery before deadline earns a bonus and delays incur a penalty in construction contracts. Jensen and Murphy [1990] document that over 90 percent of all large manufacturing firms had bonus plans in 1983, and 87 percent of firms with bonus plans paid bonuses for 1983 performance. The median bonus award for CEOs in the Conference Board's survey used by these authors was about 50 percent of base salary; over 20 percent of the surveyed firms report CEO bonuses exceeding 70 percent of base salary.

From a practical point of view, we also show (Proposition 2) that the optimal contract is fully determined by only two parameters: the size of the bonus awarded for exceptional performance and the overall probability of incurring a penalty, e.g., this can

[3]Randomization may improve the performance of contracts for risk-averse agents because it can relax incentive constraints. It turns out that the optimal contracts derived here use randomization very sparingly: a binary lottery at a few threshold outcomes that separate the contract's different regions, i.e. the penalty region from the basic wage region. All other payments are deterministic.

be interpreted as a low wage that is determined by the limited liability constraint, or as termination. Our results yield an intuitive, simple testable prediction: the ratio between the additional, performance-related pay and the basic wage is proportional (with a factor $(1+\beta)$ that measures the agent's disappointment- and risk-aversion) to the ratio between the probability of administering the penalty and the probability of obtaining the extra performance-related pay. Thus we should observe higher normalized bonuses (as percentage of base wage) in employment relations where average tenures are shorter and where the probability of obtaining a bonus is smaller.

Existing empirical work finds that firms with stronger disciplinary environments (higher turnover-performance sensitivity, stronger boards, greater predicted dismissal risk) also tend to employ a stronger performance-based compensation. But, theoretical models have some difficulty explaining why these two incentive instruments that appear to be substitutes are often used together. Our model with disappointment-averse managers provides a natural rationale: increasing dismissal risk raises the downside disappointment associated with failure, which must be offset by greater upside rewards to satisfy participation while maintaining incentives. Thus, performance-related compensation and dismissal risk are here **complements** in an optimal contract!

Peters and Wagner [2014] empirically analyze a very large data set of CEO contracts and employment features such as tenure length with particular attention to "forced turnover", aka dismissals. A main finding is that an increase of 1% in the probability of dismissal is associated with a 10% increase in total compensation. Since the basic wage varies very little in their data set, most of this increase can be associated to the part of the total compensation that is performance-related, such as bonus, stocks, etc... This is consistent with our finding that agents must be adequately compensated for turnover risk.

Proposition 3 extends the characterization of the optimal contract to interior actions, including cases where upward or non-local constraints bind. The existence of an optimal contract requires there the imposition of an upper bound $\bar{t}$ on transfers. The optimal contract in these cases still uses at most three payment levels: zero (the limited-liability bound), the base wage equal to the agent's certainty equivalent, and the bonus $\bar{t}$. We note that when upward constraints bind, the outcomes at which the agent is fired / receives the bonus need not lie below/above a single threshold. Finally, a-priori restricting the contract to be monotone can change the allocation of payments across outcomes and the cost of implementation, but does not require additional payment levels (Proposition 4).

Our last main result (Proposition 5) is about the case where both agent and principal are disappointment- and risk-averse. The optimal contract continues to be piece-wise linear with some discrete jumps but consists now of at most five distinct parts: the

penalty/basic wage/bonus scheme described above and, in addition, two separate areas where the agent's pay is linearly increasing in outcome. More precisely, the contract begins at zero (which is here the limited liability constraint) for poor performance and only starts rising after a specific outcome hurdle is met[4]. After rising, the agent's pay is capped at a maximum amount for a certain range of outputs[5]. Finally, there is a reactivated "high-performance" region where the agent's payoff rises with output, including, as before, a distinct upward jump at the highest possible outcome[6].

Real life instances of such constructions are, for example, Elon Musk's 2018 Tesla Performance Award[7] and the waterfall distribution plans found in contracts between the general partner (GP) and the limited partners (LP) in private equity and venture capital firms.

Edmans et al. [2017] observe that lower-level managers and employees are largely paid through fixed compensation and moderate bonuses while senior executives receive a much larger fraction of performance-contingent compensation. Commenting on this broad pattern, Gerhart and Milkovitch [1990] argue that "the higher in an organizational hierarchy a job is, the greater potential impact on organizational performance its incumbent is likely to have... there is also more likely to be a perceived link between employee's effort and organizational performance. This stronger link may offer organizations more opportunity to use bonuses and long term incentives among high-level managers than among other employees".

Viewed through this lens, we posit a risk-neutral principal in her interaction with low/middle manager or employee whose relative influence on the total outcome of the enterprise is small, and a disappointment- and risk-averse principal in her interaction with a top executive whose activities have a significant impact on the enterprise's outcome[8].

The remainder of the paper is organized as follows. Section 1.1 reviews the related literature. Section 2 and Appendix A briefly describe Gul's theory of disappointment aversion. Section 3 presents the moral hazard problem. Section 4 establishes the conditions for implementability, and Section 5 derives the optimal contracts. Section 6

[4]This can be replicated by a long call option that is initially "out of the money".

[5]In practice this is used to mitigate excessive risk-taking or windfall gains.

[6]In practice this is achieved by secondary tier of deep out of money options designed for extraordinary performance.

[7]Musk received no base salary, and compensation kicked in only after Tesla hit an initial market capitalization milestone of $100 billion and other operational goals. The contract was further split in 12 tranches: once a specific market cap tranche was hit the compensation was capped until the next target. Finally, if performance broke a massive, seemingly impossible "stretch" target a whole new block of stock options unlocked. See United States Securities and Exchange Commission [2018].

[8]Activist hedge funds (e.g., Pershing Square Capital or Trian) also fit this setting. Instead of owning tiny slices of hundreds of companies, they take massive, concentrated positions in a handful of target firms to force operational changes. A single failure can severely damage the fund's overall performance and trigger investor redemptions.

concludes. All proofs are relegated to Appendices B and C.

1.1 Related Literature

The empirical literature on the structure of compensation packages is extremely large. Murphy [1999] and Leonard [1990] and Baker, Jensen and Murphy [1988] systematically document executive salary, annual bonus, stock options, and turnover and how the relative importance of these instruments changes as a function of the level in the firm's hierarchy.

In their comprehensive survey of both theory and empirical findings Edmans et al. [2017] note that the value of pay to a risk-averse executive is potentially much lower than the nominal one since executives receive performance-linked pay and have often large holdings of company stock and options that are highly correlated with their firm-specific human capital. These authors also document how performance-based grants have become more complex over time. Earlier studies documented relatively simple performance-vesting grants (with zero vesting up to a threshold and full vesting at the threshold). After 2010, Edmans et al find that a majority of schemes for CEO compensation consist of a number of securities that vary in a piecewise linear way with the performance metric(s). Their Figure 11 illustrates the payoff structure of a typical formula-based bonus plan: no bonus is paid until performance reaches a lower threshold, at which point the payoff jumps discretely. On the upside, the bonus is capped at a second threshold beyond which the payoff does not increase. In the "incentive zone", the bonus increases in performance. In the middle of the incentive zone is a "target" performance level at which a "target" bonus is awarded. This structure is also found in the part of the compensation consisting of stock/options (their Fig. 10). Finally, Edmans et al [2017] also estimate that around half of CEO turnovers in public U.S. firms are "performance-induced", i.e., would not have occurred had performance been better and note that they are unaware of any moral hazard model that yields quantitative predictions for the optimal rate of firing. In fact, our results agree well with their general empirical findings.

Chakraborty et al. [2009] study the relationship between CEO compensation incentives and performance related turnover. Their theoretical model predicts that a steeper slope of the compensation contract may be associated with higher likelihood of performance related termination and confirms this finding empirically. The idea is that low outcomes despite high incentives is a negative signal about the agent's ability.

Batt and Colvin [2011] also argue that organizations that make greater use of performance incentives have significantly higher dismissal rates/higher total turnover. In their empirical study of U.S. call-centers these authors find that monitoring intensity and commission pay are positively associated with dismissals.

The theoretical literature on moral hazard problems with hidden actions under expected utility is very rich and contains several seminal contributions, e.g., Mirrlees [1976], Holmström [1979], Shavell [1979], Grossman and Hart [1983], Jewitt [1988], Rogerson [1985] and Innes [1990] [9].

The present sharp results characterizing optimal contracts in the presence of various frictions are in stark contrast to the results obtained under expected utility where, roughly speaking, optimal pay is very sensitive to performance, sometimes in counter-intuitive ways. In their textbook dedicated to the subject, Bolton and Dewatripont [2005], page 157 summarize the main findings of the classical model and conclude:

> ...the most that can be said about optimal incentive schemes, when no restrictions are put on probabilities except for requiring that they are all positive, is that the transfers cannot be decreasing everywhere in performance, nor increasing everywhere faster than performance... These are rather weak predictions. Just as sharp characterizations of the optimal incentive scheme are not obtainable in general, simple comparative statics results are unavailable... Finally, it should come as no surprise that sharp characterizations of the agent's second-best action are also not obtainable.."

The discrepancy between theoretical results and the practical design of incentives systems where most observed systems have a piece-wise linear (and overall non-linear) structure has been a continued source of puzzlement for economists. Prendergast [1999] surveys the theory along its applications to labour contracts and writes:

> "Yet many observed contracts are nonlinear, where, for example, discrete bonuses are offered for exceeding some performance threshold. For instance, Kevin J. Murphy (2000) highlights the importance of such bonus contracts for executives. Perhaps the most important form of nonlinearity concerns the threat of being fired, where wages vary little with performance but where poor performance is punished by dismissal. Rather remarkably, the theoretical literature has made little progress in understanding the observed (nonlinear) shape of optimal contracts."

In a classic contribution, Holmström and Milgrom (1987) assume that both principal and agent have CARA utility, and that the agent controls the drift of a Brownian motion in continuous time. Although the principal can make payments that depend on the entire history, the optimal contract is simply a linear function of the endpoint. The main

[9] Textbook expositions of the theory can be found in Bolton and Dewatripont [2005] and Salanie [2005], among others.

structural feature behind their result is the stationary structure of the model with CARA utilities: at each point in time, the optimal incentives going forward are independent of the previous history, and this leads to linearity.

An important theoretical avenue towards reconciling theory and empirical observations is the analysis of models where it is assumed that the agent's perceived utility is different from classical expected utility. For a comprehensive survey of behavioral contract theory see Kőszegi [2014].

Carroll [2015] offers another explanation for linear contracts: he considers a static moral hazard problem with risk neutral agent and principal where the principal knows some, but not all actions available to the agent. The principal evaluates contracts by their worst-case performance (including the unknown actions the agent might potentially take) and the resulting optimal contract is always linear.

De Meza and Webb [2007] looked at several moral hazard problems involving some form of the agent's loss-aversion and the MLRP property. One of their results (Proposition 3) looks at Gul's specification with a general form of utility over monetary outcomes. The authors consider only deterministic contracts, two effort levels and no limited liability constraint. An optimal feature, that we also replicate here, is the existence of an interval of outcomes where the pay to the agent does not vary with performance. It turns out that, in their framework, the solution for an agent having an endogenous reference point a la Gul coincides with the solution under an exogenous reference point. This feature is driven by the fact that, at their optimum, both the individual rationality constraint and the unique incentive constraint must bind, thus completely determining the solution. We note that in our model with many actions and with a limited liability constraint, the individual rationality constraint is generally lax at the optimum, and the agent receives here an "information" rent.

Herweg et al. [2010] use Köszegi and Rabin's [2006] formulation of loss-aversion with a linear utility over monetary outcomes. Thus, they consider a moral hazard model where the agent pairwise compares each wage with any other wage that he could have received instead, where each comparison is weighted by the occurrence probability of the alternative wage: one and the same wage is being perceived as both a gain relative to some wages and a loss relative to others, leading to a stochastic reference point (a phenomenon the authors call "mixed feelings"). This important feature is absent in our model where all wages are compared to a single other payment, Gul's endogenous certainty equivalent. Herweg et al. [2010] also conjecture that, because wage differentiation is less costly in the absence of mixed feelings one should expect the corresponding optimal contract to be more differentiated than their binary one, but still simpler than a fully contingent contract such as the optimal one under expected utility. This intuition agrees well with

our results that are able to explain the more complex structure of observed contracts in practice. In Section 3 we discuss the limitations of their result and illustrate settings where the optimal contract under Köszegi and Rabin's formulation of loss aversion becomes as complex as possible, i.e. contains as many different payments as the number of possible outcomes.

Chaigneau et al [2025] study optimal contracts for an agent who believes that she deserves a given share of output, the "perceived" fair wage. Their main finding is that a binary (base wage and bonus) scheme is optimal[10]. If his wage is below that, he suffers a disutility that is linear in the discrepancy. The principal is risk neutral, and both parties are protected by limited liability. Thus, in contrast to the present paper, the reference point is increasing in output. This leads to a rather different optimal contract, in which pay-performance sensitivity is positive even if incentives are not a concern; if incentives are a concern, pay-performance sensitivity is driven by the fair wage rather than by the trade-off between incentives and risk sharing. These authors show that "unfairness" can be a powerful motivator: if output is below a threshold, the optimal contract pays a zero wage; once output crosses that threshold, the CEO is paid the fair wage.

Dittman et al. [2010] consider a moral hazard problem where the agent is loss averse and has an exogenous reference point. Under a specialized technical condition that ensures the validity of the first-order approach, their optimal contract is simple: it pays a fixed low wage for low outcomes below a threshold, and pays the exogenous reference wage plus a monotonically increasing part above that threshold. The authors note that the obtained non-linear contract can be implemented only by using a continuum of securities, and therefore cannot be directly compared to observed contracts that are piecewise linear.

2 Gul's Theory of Disappointment-Aversion

Assuming piece-wise linear utility over prizes[11] Gul's certainty equivalent of a lottery $\mathbf{t}$, $CE[\mathbf{t}]$, is defined as a weighted average of the value of prizes that are above it and those that are below it - note the circularity in this definition! In the Appendix we show that, for our purposes, a convenient form to express $CE[\mathbf{t}]$ is given by:

$$CE[\mathbf{t}] = \mathbb{E}[\mathbf{t}] - \beta\,\mathbb{E}[(CE[\mathbf{t}] - \mathbf{t})^{+}].$$

[10] In this framework, Ho [2025] shows that binary contracts that include some stochasticity can reduce the principal's expenditure.

[11] We note that Gul's theory is more general and allows for a more general utility functions over prizes.

This formulation clarifies that the certainty equivalent is always smaller than the lottery's expectation whenever $\beta > 0$. The agent is then both disappointment- and risk-averse in the usual sense of aversion to mean preserving spreads.

Bellini et al [2025] offer an axiomatization of the preferences used here (their Theorem 5) and elegantly connect Gul's theory to both the pioneering econometric work of Newey and Powell [1987] on asymmetric least square estimation and expectiles, and to the theory of elicitable and coherent risk measures.

Observe that $\mathbb{E}[\mathbf{t}]$, the expectation of a lottery $\mathbf{t}$, is the unique solution v to the equation

$$\mathbb{E}[(\mathbf{t}-v)^+] = \mathbb{E}[(v-\mathbf{t})^+] \Leftrightarrow \frac{\mathbb{E}[(\mathbf{t}-v)^+]}{\mathbb{E}[(v-\mathbf{t})^+]} = 1.$$

Bellini et al [2025] (Proposition 1) offer an intuitive representation of Gul's certainty equivalent $CE[\mathbf{t}]$ as given by the unique solution v to the equation

$$\mathbb{E}[(\mathbf{t}-v)^+] = (1+\beta)\,\mathbb{E}[(v-\mathbf{t})^+] \Leftrightarrow \frac{\mathbb{E}[(\mathbf{t}-v)^+]}{\mathbb{E}[(v-\mathbf{t})^+]} = 1+\beta.$$

In other words, Gul's certainty equivalent equals the threshold v where the expected elation $\mathbb{E}[(\mathbf{t}-v)^+]$ **overcompensates** the expected disappointment $\mathbb{E}[(v-\mathbf{t})^+]$ precisely by a factor $(1+\beta) > 1$.

Finally we note that Gul's certainty equivalent with piece-wise linear utility over prizes is both positively homogeneous and translation invariant, i.e. $CE[\gamma\mathbf{t}] = \gamma CE[\mathbf{t}]$ for $\gamma > 0$, and $CE[\mathbf{t}+a] = CE[\mathbf{t}] + a$ for any constant a. In other words, the resulting utility functions belong the class of utilities satisfying constant risk aversion (see Safra and Segal [1998]).

3 The Moral Hazard Problem

We consider here the classical moral hazard problem with hidden actions and only change the specification of preferences for the agent. A disappointment-averse agent with parameter $\beta > 0$ has a set of costly actions $A = \{a_1, \ldots, a_K\}$. We assume that

$$0 = c(a_1) < c(a_2) < \cdots < c(a_K),$$

where $c(a) \geq 0$ is the cost of action a. The principal wants here to implement a_K, the costliest action We consider the implementation of interior actions in Section 5.3.

The taken action is not observable by the principal, who is risk-neutral. The principal only observes the action's realized outcome x, out of a finite set of possible outcomes $X =$

$\{x_1, ..., x_n\}$ where $n \geq 3$. Without loss of generality we assume that $x_1 < x_2 < ... < x_n$. We denote by

$$p_a(x_i) > 0$$

the probability of outcome x_i conditional on the agent's action being a. We denote by $\mathbf{p}_a$ the vector of probabilities $(p_a(x_1), .., p_a(x_n))$

We also assume that family of distributions $(\mathbf{p}_{a_1}, \ldots, \mathbf{p}_{a_K})$ is ordered in the strict MLRP order. That is, for every $r > s$, the ratio

$$\frac{p_{a_r}(x_i)}{p_{a_s}(x_i)}$$

is strictly increasing in i.

4 Implementable Actions

A contract specifies a (possibly random) vector $\tau = (\tau_1, \tau_2, .., \tau_n)$, where τ_j is the principal's random payment (wage) to the agent in case the realized outcome is x_j. We assume that the transfers $\tau_1, \tau_2, \ldots$ have finite expected values. The principal's optimal contract must avoid excessive disappointment - which is very costly for the agent and hence must be compensated by the principal - and, as usual also under expected utility, balance the insurance motive with the incentive needs.

A given contract τ together with a chosen action $a \in A$ determine together a random compensation for the agent: this is described here by a random variable $\mathbf{t}_a$ where

$$\mathbb{P}[\mathbf{t}_a \leq t] = \sum_{i=1}^{n} \mathbb{P}[\tau_i \leq t] \, p_a(x_i).$$

Finally, we consider here the case with limited liability where $\mathbf{t}_a \geq 0$ for all $a \in A$, that is each random compensation $\mathbf{t}_a$ has a non-negative support[12]

We say that an action a_j is implementable if there is a contract τ with induced compensation vectors $\mathbf{t}_{a_i} \geq 0$, such that for all i

$$CE\left[\mathbf{t}_{a_j}\right] - c(a_j) \geq CE\left[\mathbf{t}_{a_i}\right] - c(a_i)$$

where $CE[\mathbf{t}_{a_l}]$ denotes the certainty equivalent of the (random) transfer $\mathbf{t}_{a_l}$ when the agent takes action a_l. While for general non expected-utility preferences it might be optimal for the agent to deviate to a mixed action, it suffices to consider pure deviations for Gul preferences as they satisfy betweenness.

[12]Setting the limited liability at zero is a normalization.

We say that an action a_j is *statistically identifiable* if the associated distribution p_{a_j} can not be generated by any mixed strategy over actions other than i, i.e. $\mathbf{p}_{a_j} \notin \text{co}\{\mathbf{p}_{a_i} : i \neq j\}$. The precise condition ensuring that an action is implementable is cost-sensitive. If an action is implementable for every increasing, non-negative cost vector we say that it is *uniformly implementable*. Below we establish that identifiability is necessary and sufficient for an action to be uniformly implementable.

Proposition 1. *Fix an action a_j. The following statements are equivalent:*

(i) Action a_j is identifiable.

(ii) Action a_j is uniformly implementable.

(iii) Action a_j is uniformly implementable using a contract that uses exactly 2 payment levels.

Proposition 1 establishes that, under the identifiability condition all actions are uniformly implementable. Furthermore, bonus contracts with only two payment levels are enough to implement any action. Thus, the potential appearance of additional payments in an optimal contract is due to optimality conditions rather than being necessary for implementation. Hermalin and Katz [1991] showed that the same identifiability condition (i) is equivalent to uniform implementability under expected utility.

5 The Optimal Contract

Throughout the paper we say that a contract is an optimal contract implementing a given action a if it provides the agent an incentive to take action a and if it minimizes the expected cost to the principal $\mathbb{E}[t_a]$. An optimal contract τ implementing the costliest action a_K solves

$$\begin{aligned} &\min \mathbb{E}\left[\mathbf{t}_{a_K}\right] \\ &s.t. \\ CE\left[\mathbf{t}_{a_K}\right] - c(a_K) &\geq CE\left[\mathbf{t}_{a_j}\right] - c\left(a_j\right) \quad \forall j \neq K && (IC_j) \\ CE\left[\mathbf{t}_{a_K}\right] - c(a_K) &\geq 0 && (IR) \\ \mathbf{t}_{a_K} &\geq 0. && (LL) \end{aligned}$$

The (IR) constraint is here redundant as it follows from $(IC_1), (LL)$, our assumption that $c(a_1) = 0$, and the fact that Gul preferences are monotonic in the FOSD order. We will thus henceforth ignore the participation constraint.[13]

[13]Formally, $CE\left[\mathbf{t}_{a_K}\right] - c(a_K) \geq CE\left[\mathbf{t}_{a_1}\right] - c\left(a_1\right) = CE\left[\mathbf{t}_{a_1}\right] \geq 0$.

In general, there might be multiple optimal contracts. Throughout the main text, we restrict attention to optimal contracts that are *risk-undominated* in the sense that there does not exist another optimal contract giving the agent the same value while exposing him to less risk (in the mean-preserving spread sense).[14]

The assumed strict MLRP order of the distributions is sufficient to ensure the implementability of the highest action a_K. The following theorem completely characterizes the principal's optimal contract that implements the highest action a_K. This contract induces a payment structure consisting of only three pay levels: a "regular" or baseline positive wage for middle outcomes, a penalty (or malus) for low outcomes, and a bonus for the highest outcome.

Theorem 1. *Any optimal contract implementing a_K, uses either two or three different payments. There exists an $m < n$ such that any optimal contract*

1. *Pays* 0 *with probability* 1 *if a low outcome $x_1, \ldots, x_{m-1}$ is realized*[15]*;*
2. *Pays a binary lottery with payoffs* 0 *and $CE[\mathbf{t}_{a_K}]$ if outcome x_m is realized;*
3. *Pays $CE[\mathbf{t}_{a_K}]$ with probability* 1 *if one of the outcomes $x_{m+1}, \ldots, x_{n-1}$ is realized and*
4. *Pays a non-random transfer $\tau_n > CE[\mathbf{t}_{a_K}]$ if the highest outcome x_n is realized.*

The proof is given in Appendix B and proceeds in four steps. First, a scaling argument shows that at least one downward incentive constraint must bind. Otherwise, all transfers could be reduced proportionally while preserving the downward constraints, and lowering the principal's expected payment. Second, any transfer strictly between 0 and the certainty equivalent $w = CE[t_{a_K}]$ can be replaced by a binary lottery over $\{0, w\}$ that preserves both the target action's expected transfer and its certainty equivalent. This replacement strictly lowers the certainty equivalent of every lower action.[16] It would therefore make all downward constraints slack, contradicting the first step. Third, the MLRP ordering implies that every payment above w at an outcome below the highest outcome x_n can be shifted to x_n while preserving the target action's expected payment and making every lower action less attractive. Thus elation payments occur only at the

[14]This is a natural selection, resulting from adding a small amount of strict risk aversion to the agent's preferences. In the Appendix we provide versions of our results without this assumption. Theorem 1 slightly changes as the payment after the highest outcome can be random in the optimal contract. Proposition 2, 3, and 4 do not change even if we remove this assumption.

[15]The set of outcomes where 0 is paid for sure is empty if $m = 1$.

[16]The first two lemmas in the proof are used also to establish Proposition 3 about implementation of interior actions. They hold even if transfers are bounded and the agent can deviate not only to lower, but also to higher actions.

highest outcome. Finally, another MLRP rearrangement argument moves the probability of a zero payment toward lower outcomes. Consequently, zero is paid below a single threshold outcome, possibly with randomization at the threshold and w is paid at all intermediate outcomes.

The distinction between two and three payment levels in Theorem 1 reflects the trade-off between insurance and incentives. For a risk-neutral agent, the well known optimal contract is binary: it pays a bonus only at the highest outcome and zero otherwise. This binary structure remains optimal for small values of β. Three payment levels become optimal if intermediate outcomes are sufficiently informative about effort, or if the agent is sufficiently disappointment averse. Introducing then a base wage equal to the agent's certainty equivalent insures and protects the agent from disappointment at intermediate outcomes and can thereby reduce the cost of implementation. Proposition 6 in the Appendix characterizes when exactly two or three payment levels are optimal; Corollary 1 displays a transparent sufficient condition for three payment levels to be optimal independently of costs:

$$\max_{i<K} \frac{p_{a_i}(x_{n-1})}{p_{a_K}(x_{n-1})} \leq \min\{1, \beta\}\,.$$

Intuitively, the second highest outcome must be a sufficiently informative signal of effort and the disappointment aversion needs to be strong enough.

In our model, we assumed linear utility over monetary outcomes. However, analogously to Herweg et al. [2010], the continuity of preferences together with the presence of a finite number of outcomes imply that adding a small curvature to the utility over money does not change here the characterization of the optimal contract.

5.1 The Bonus and the Penalty Probability

The optimal contract derived above is characterized only by two parameters: the size of the bonus payment τ_n and the overall probability of the penalty (i.e., zero compensation) conditional on the implemented action a_K. Fixing this last probability uniquely determines the range of outcomes $x_1, .., x_{m-1}$ where the punishment is administered with probability 1, the necessary randomization at the threshold outcome x_m, and the range of outcomes $x_{m+1}, .., x_{n-1}$ where the basic wage is paid with probability one.

Consider an optimal contract implementing the highest action a_K, and let

$$w := CE[\mathbf{t}_{a_K}]$$

denote the basic wage, let

$$\pi_0 := \mathbb{P}_{a_K}\left[\mathbf{t}_{a_K} = 0\right]$$

denote the probability of the zero payment under the implemented action and finally let

$$B := \mathbb{E}\left[\tau_n - w\right] = \mathbb{E}[(\tau_n - w)^+]$$

denote the performance-related payment in excess of the regular wage, where the expectation is over any randomization in the payment following the realization of the highest outcome. The second equality holds because the optimal lottery at the highest outcome uses only payments weakly above w.

There are different combinations of the two parameters (B, π_0) that generate the same utility to the agent from taking action a_K and the principal needs to choose the particular pair that minimizes the transfer to the agent while still implementing action a_K[17].

Proposition 2 (Expected bonus and firing probability)**.** *If $w > 0$, then*

$$\frac{B}{w} = (1+\beta)\frac{\pi_0}{p_{a_K}(x_n)}. \tag{1}$$

Proof. See Appendix B. □

Proposition 2 yields a clear, testable prediction of our model: the ratio between the additional, performance-related pay and the basic wage is proportional (with a factor $1+\beta$ that measures the agent's disappointment- and risk-aversion) to the ratio between the probability of administering the penalty and the probability of obtaining the extra performance-related pay. Thus higher bonuses as a percentage of base wage should be observed in those employment situations where average tenures are shorter and where the probability of obtaining a bonus is smaller[18].

Ang et al. [2005] calibrated a number of data generating processes to a large US data set on stock and bond returns assuming that investors are disappointment averse and find very reasonable portfolios associated to moderate disappointment aversion ($\beta \in [0.3, 1]$) with utility functions exhibiting quite low curvature. Assuming that β is in that interval, the relation between dismissal probability and performance pay described in Peters and Wagner [2014] would emerge from our formula above whenever the probability for the exceptional performance and for the bonus to materialize is approximately $\frac{1+\beta}{10}$, and thus in the interval $[0.13, 0.2]$.

[17]Note that the payoffs from all other actions are also pinned down by these two instruments.

[18]Recall here Batt and Colvin [2011, p. 20 and Appendix 1], study of U.S. call-centers where they define an index of performance pay and find that this index is associated with higher dismissal rates [2011, Table 2, model 2b]. They also report that dismissal rates increase with the fraction of pay at risk [2011, p. 30].

5.2 Discussion

We gather below several notable structural remarks about the optimal mechanism and its interpretation:

Randomization Although we allow for random mechanisms, the optimal contract makes a minimal use of this option: randomization occurs only at the unique treeshold outcome x_m that separates the set of outcomes where the agent is punished for low performance from the set of outcomes where he receives a fixed, positive wage that does not vary with performance.

In principle, it is possible to use randomization at the highest outcome where a bonus is paid, i.e., announcing for example that, if the highest outcome realizes, a certain bonus will be paid with some probability less than 1 and otherwise the agent gets the regular wage. While this approach is sometimes used in practice, from a theoretical point of view it is here without loss of generality to only consider a deterministic bonus. This follows because all payments in an optimal binary lottery at the highest outcome must be weakly higher than the (endogenous) certainty equivalent.

The Agent's Rent Recall that in the case of an observable action, the optimal compensation is outcome-independent and fixed at $c(a_K)$. Since the cost of the least costly action a_1 is zero, and since even when the agent chooses that action he gets a positive wage with a positive probability, we obtain by IC_1 that the IR constraint does not bind, i.e. $CE[t_{a_K}] > c(a_K)$. Therefore the agent necessarily gets here an information rent.

Limited Liability The precise limited liability level, here taken to be zero, doesn't play an important role other than specifying the lower bound on the punishment. Any other liability limit yields analogous results. Moreover, the probability of punishment is (weakly) increasing in this bound. In the limit where the limited liability bound goes to minus infinity, the model converges to the one without limited liability where a punishment (i.e., a wage below the certainty equivalent) occurs only when the lowest outcome x_1 is realized. In that case, there is no optimal contract because decreasing the probability of punishment while increasing its size is always beneficial for the principal. Observe also that, if the limited liability bound decreases to minus infinity, while the cost of the lowest effort remains fixed at, say, $c(a_1) = 0$, then, at some point, the IR constraint will bind. We obtain then that $CE[\mathbf{t}_{a_K}] = c(a_K)$ and the structure of the optimal contract in the Theorem doesn't change. The punishment is imposed at some lowest outcomes, exactly as with zero limited liability.

Upper Bound on Transfers In addition to the bound from below defining the limited liability constraint, we sometimes observe in practice a bound from above on transfers, $\bar{t} > 0$, and also payments at the bound on a range of high outcomes. In our model the imposition of such a bound yields two distinct cases:

1. If the upper bound $\bar{t}$ is larger than the optimal calculated bonus $E[\tau_n]$, then the optimal contract without bound - which pays a non-random bonus only at the highest outcome - remains optimal. As we establish in Lemma 7, a sufficient condition for this to happen is that the upper bound on the transfer $\bar{t}$ satisfies[19]

$$\bar{t} \geq \max_{i<K} \frac{c(a_K) - c(a_i)}{g(p_{a_K}(x_n)) - g(p_{a_i}(x_n))}$$

where $g(q) := \frac{q}{1+\beta(1-q)}$.

2. If the bound is lower than the optimal calculated bonus, then the optimal contract indeed pays $\bar{t}$ on several high outcomes. That is, there are $m, l \in \{1, ..., n\}$ with $1 \leq m < l \leq n$ such that, the optimal contract pays 0 with probability 1 if one of the outcomes $x_1, \ldots, x_{m-1}$ is realized; pays a binary lottery with payoffs 0 and $CE[t_{a_K}]$ if outcome x_m is realized; pays $CE[t_{a_K}]$ with probability 1 if one of the outcomes $x_{m+1}, \ldots, x_{l-1}$ is realized; pays a binary lottery with payoffs $CE[t_{a_K}]$ and $\bar{t}$ if outcome x_l is realized, and pays $\bar{t}$ with probability 1 if one of the outcomes $x_{l+1}, \ldots, x_n$ is realized.

The EU Case with a Piecewise Linear Bernoulli Function A comparison can be made between our setting and the EU framework with a piecewise linear, concave Bernoulli utility where the agent is risk-averse. In that case (for the analogous situations with one kink, as analyzed here), the optimal contract will also have a support with three payments: zero, a bonus and a middle wage. However, there the wage is defined solely by the kink in the Bernoulli utility function and is completely independent of the parameters of the underlying moral hazard problem such as the costs and the probability distributions over outcomes. Changing the parameters of the problem may change there the optimal action the principal wants to implement; yet, the wage in the optimal contract remains unchanged. This is a very unappealing property that cannot reflect contracts used in practice.

A Comparison to Yaari and Reference Dependent Preferences As mentioned in the introduction, Herweg et al. [2010] derived an optimal binary contract (i.e., a basic wage and a bonus) in a framework where the agent has a continuum of possible actions and loss-averse preferences with a stochastic reference point a la Köszegi-Rabin [2006].

[19] The term on the righ-hand-side is finite by the strict MLRP assumption.

We note here that their specification of preferences coincides in fact with a risk-averse *dual* utility a la Yaari [1987] where the probability weighting function g is convex and quadratic, of the form

$$g(p) = (2 - \lambda)p + (\lambda - 1)p^2$$

where $1 \leq \lambda \leq 2$ is the parameter of loss and risk aversion. It is interesting to note that Yaari's dual utility theory incorporates rank dependence because the associated probability distortion depends on the payment's rank in the overall scheme, but there is **no** explicit reference point (neither deterministic nor stochastic).

The elegant result proven by Herweg et al about the optimality of binary schemes closely follows from their first order approach that allows them to reduce the problem to a linear maximization with an IR and a unique IC constraint. In order to make the first-order approach valid, Herweg et al impose several relatively strong assumptions on the primitives and on the class of contracts they analyze. In particular, all probability vectors associated to the possible outcomes as a function of effort ($\mathbf{p}_{a_K}$ in our notation - these are functions in their case with a continuum of effort levels) need to be colinear: an extra unit of effort moves all probabilities in exactly the same way, independently of the initial effort. This basically reduces their model to one with two actions, high and low with costs which are convex in the probability of taking the higher action. In Appendix D we display an example a la Herweg et al where the strict MLRP holds but where their colinearity condition on probabilities does not. The optimal contract contains then as many payments as there are outcomes. Thus, neither Yaari nor Choice-acclimating preferences will always lead to a small number of payment levels.

5.3 The Optimal Contract Implementing Interior Actions

Suppose now that the principal wants to optimally implement an interior action a_r, $1 < r < K$, in a framework where transfers are bounded both from below and from above:[20]

$$0 \leq \tau_h \leq \bar{t} < \infty \text{ for every } h.$$

Here MLRP by itself is not enough to ensure implementability. As shown in the proof of Proposition 1, any interior action is implementable under a weak identifiability condition given a sufficiently high upper bound $\bar{t}$ on transfers. That proof also gives a formula for how large $\bar{t}$ needs to be in order to be able to implement a specific action. Recall also that if the necessary bonus for implementation without an upper bound exceeds $\bar{t}$, then the optimal contract pays $\bar{t}$ at all outcomes above a threshold, and may randomize between

[20]Note that the lowest action a_1 can always be implemented by a constant transfer of zero.

the baseline wage and $\bar{t}$ at the threshold outcome.

If no upward incentive constraint binds when implementing an interior action a_r, the "capped" version of Theorem 1 immediately gives the optimal contract: one can simply ignore all actions above a_r and treat a_r as the highest action. The next Proposition shows that, even in the case where some upward incentive constraints bind, the contract still uses at most three payment levels: zero (the limited-liability bound), the baseline wage equal to the agent's certainty equivalent, and the bonus $\bar{t}$.

Proposition 3. *Fix an implementable, interior action a_r, $1 < r < K$.*

(i) There is an optimal contract τ implementing a_r such that its support is contained in $\{0, CE[\mathbf{t}_{a_r}], \bar{t}\}$.

(ii) If an upward incentive constraint binds at every optimal contract, then every optimal contract is of this form.

The main difference to the optimal contract implementing the highest action is the allocation of the three payments across outcomes. When upward constraints bind, the outcomes at which the agent is fired/receives the bonus need not lie below/above a single threshold. The reason is that a deviation to a higher action in a monotone contract might enable the agent to receive the bonus with a higher probability but without generating significant benefit for the designer. To optimally deter such deviations, the principal might pay the bonus at interior outcomes. We note that this phenomenon is not a peculiarity of the model with Gul's preferences: it also appears under EU.

Optimal Monotone Contract We call a contract *monotone* if its associated payment lotteries conditional on the realized outcome are ordered by first-order stochastic dominance across outcomes. We noted above that optimal contracts implementing interior actions need not be monotone (even under MLRP). Restricting attention to monotone contracts is natural in practice: for example such contracts leave the agent with no incentive to destroy output in cases where he may have the opportunity to do so. Our next result considers optimal monotone contracts. The monotonicity restriction can change the allocation of payments across outcomes and the cost of implementation, but it does not require additional payment levels.

Proposition 4. *Fix an interior action a_r, $1 < r < K$, that is implementable by a feasible, monotone contract. There exists a contract τ that is optimal among all the monotone contracts implementing a_r such that the support of τ is contained in the set*

$$\{0, w, \bar{t}\},$$

where $w := CE[\mathbf{t}_{a_r}]$. *If an upward incentive constraint binds at every optimal monotone contract, then every optimal monotone contract has a support contained in the above set.*

Remark 1 (Computing the Optimal Contract)**.** *Due to the three-payment structure established in Proposition 3 and 4 it is easy to compute the (monotone) optimal contract: it turns out that the resulting optimization problem is linear when fixing the target action* a_r *and the certainty equivalent of the agent* $CE[\mathbf{t}_{a_r}]$. *We formally derive the optimal contract in Proposition 7 in the Appendix.*

5.4 Both Principal and Agent are Disappointment- and Risk-Averse

Assume now that both the agent (a) and the principal (p) are disappointment-averse with parameters $\beta^a > \beta^p > 0$, so that the agent is more averse than the principal. We comment below on the reverse case where $\beta^p > \beta^a > 0$**.** In this section, we assume that the principal seeks to provide incentives in order to implement the highest action, a_K. The principal's problem is

$$\begin{aligned} &\max_{\tau} CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] \\ &\qquad s.t. \\ CE^a\left[\mathbf{t}_{a_K}\right] - c\left(a_K\right) &\geq CE^a\left[\mathbf{t}_{a_j}\right] - c\left(a_j\right),\ \forall j \neq K && (\mathrm{IC}_j) \\ CE^a\left[\mathbf{t}_{a_K}\right] - c\left(a_K\right) &\geq 0 && (\mathrm{IR}) \\ \mathbf{t}_{a_K} &\geq 0 && (\mathrm{LL}) \end{aligned}$$

where we denote by CE^a the certainty equivalent of the agent, and by CE^p the certainty equivalent of the principal.

Proposition 5. *There exists an optimal contract with the following structure: there exist* $m_1 \leq m_2 \leq m_3 \leq n$ *such that the optimal contract:*

1. *pays* 0 *with probability 1 at outcomes* $x_1, ..., x_{m_1-1}$.
2. *pays* $x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]$ *with probability 1 at outcomes* $x_{m_1+1}, ..., x_{m_2-1}$,
3. *pays* $CE^a\left[\mathbf{t}_{a_K}\right]$ *with probability 1 at outcomes* $x_{m_2+1}, ..., x_{m_3-1}$,
4. *pays* $x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]$ *with probability 1 at outcomes* $x_{m_3+1}, ..., x_{n-1}$,

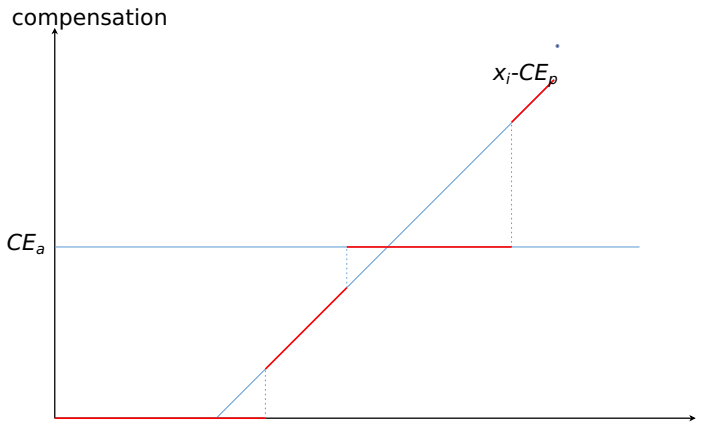


Figure 1: Compensation as a function of the realized outcome.

5. *pays the highest transfer* $\tau_n \geq \max_i\{CE^a\left[\mathbf{t}_{a_K}\right],\ x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\}$ *at the highest outcome* x_n.

6. *At outcomes* x_{m_1}, x_{m_2} *and* x_{m_3} *the transfers consist of lotteries with at most two outcomes.*

The proof of the above Proposition (see Appendix C) proceeds via several Lemmas. We first show that there exists an optimal contract where, after any outcome $x_i \neq x_n$, the compensation is either deterministic or a lottery with support

$$\left\{0, CE^a\left[\mathbf{t}_{a_K}\right], x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right\}.$$

We then show that each randomization has a binary structure. Each Lemma starts with a contract that satisfies all constraints and shows how it can be (weakly) improved from the principal's perspective if it does not satisfy certain properties.

Observe that the previous contract characterized in Proposition 1 implements the costliest action a_K independently of the risk and disappointment preferences of the principal. That contract, however, exposes the principal to substantial risk and disappointment. For risk- and disappointment-neutral principal the cheapest way to incentivise the agent is to punish him at low outcomes and to pay a bonus at the top. Low realizations of the output necessarily lead to payments in the disappointment region of the principal's utility. As a consequence, it is less costly for a risk- and disappointment-averse principal to reduce the payment for middle outcomes as long as they are in the principal's disappointment region. Similarly, if the bonus payment at the top output is in the principal's disappointment region, then reducing it and increasing the payment at the next highest outcomes also may increase the principal's payoff.

Finally, we mention that the proof of the above Proposition does not invoke the assumption that $c(a_1) = 0$. With this assumption, a simpler proof of Lemma 8 , using the same argument as in Section 4, shows that the assumption $\beta^a > \beta^p$ can be dispensed with. Thus the general structure of the optimal contract obtained above continues to hold also for the opposite case $\beta^a \leq \beta^p$. In that case, some portions of the contract may be degenerate so that the contract pays either 0 or $x - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]$.

6 Conclusion

The present study extends to an important family of non-expected utility models our understanding of the inherent trade-off between insurance and incentives appearing in moral hazard situations. It is motivated by some apparent gaps between theoretical results and real-life incentive contracts.

We studied the classical moral hazard problem with hidden actions where the agent and the principal are disappointment- and risk-averse in the sense of Gul [1991]. Gul's influential theory of disappointment aversion assumes that individuals evaluate the various outcomes of a risky action relative to an endogenous reference point: the ex-ante value of the action itself. Outcomes falling short of expectations are heavier weighted than outcomes exceeding expectations.

If only the agent is disappointment and risk averse, then the optimal contract features at most three compensation levels: a base wage, a penalty level determined by the limited liability constraint after unsatisfactory performance, and a bonus paid for exceptional performance. If both agent and principal are disappointment and risk averse, the optimal contract consists of several regions where the contract's slope is either zero or one, such that the agent's compensation is comparatively more exposed to outcome variability. The case where the principal is risk neutral better fits an interaction with a low/middle manager or employee whose relative influence on the enterprise is small, while the case where both agent and principal are disappointment- and risk-averse reflects the interaction with a top executive whose actions significantly impact the organization.

Our model also provides a testable prediction about the complementarity in the optimal contract between incentive tools that are often perceived as substitutes: the dismissal probability and the performance bonus. More precisely, we predict higher normalized bonuses (as percentage of base wage) in employment relations where average tenures are shorter and where the probability of an exceptional outcome is smaller.

Finally, we note here that, due to the non-expected utility preferences, the use of dynamic contracts may increase here the principal's payoff (see for example Mora and Strack [2025]). But, such contracts are not simple, involve complex lotteries and are not

observed in practice.

7 Appendix A (Gul's Disappointment Aversion and the Certainty Equivalent)

Gul's original formulation of the certainty equivalent assuming a piece-wise linear utility over monetary outcomes is given by:

$$CE\left[\mathbf{t}\right]=\frac{\alpha}{1+\left(1-\alpha\right)\beta}\mathbb{E}\left[\mathbf{t}|\mathbf{t}\geq CE\left[\mathbf{t}\right]\right]+\frac{\left(1-\alpha\right)\left(1+\beta\right)}{1+\left(1-\alpha\right)\beta}\mathbb{E}\left[\mathbf{t}|t<CE\left[\mathbf{t}\right]\right]$$

where $\alpha=\mathbb{P}\left[t\geq CE\left[\mathbf{t}\right]\right]$ and where β is the additional, new and unique parameter, specific to this theory. The above formulation is equivalent to

$$\begin{aligned}
0 &= \frac{\alpha}{1+\left(1-\alpha\right)\beta}\mathbb{E}\left[\mathbf{t}-CE\left[\mathbf{t}\right]|\mathbf{t}\geq CE\left[\mathbf{t}\right]\right] \\
&\quad +\frac{\left(1-\alpha\right)\left(1+\beta\right)}{1+\left(1-\alpha\right)\beta}\mathbb{E}\left[\mathbf{t}-CE\left[\mathbf{t}\right]|\mathbf{t}<CE\left[\mathbf{t}\right]\right] \\
&\Leftrightarrow \quad 0=\mathbb{E}[\mathbf{t}]-CE\left[\mathbf{t}\right]+\left(1-\alpha\right)\beta\mathbb{E}\left[\mathbf{t}-CE\left[\mathbf{t}\right]|\mathbf{t}<CE\left[\mathbf{t}\right]\right] \\
&\Leftrightarrow \quad CE\left[\mathbf{t}\right]=\mathbb{E}[\mathbf{t}]-\left(1-\alpha\right)\beta\mathbb{E}\left[CE\left[\mathbf{t}\right]-\mathbf{t}|\mathbf{t}<CE\left[\mathbf{t}\right]\right] \\
&\Leftrightarrow \quad CE[\mathbf{t}]=\mathbb{E}[\mathbf{t}]-\beta\mathbb{E}[(CE[\mathbf{t}]-\mathbf{t})^{+}]
\end{aligned}$$

The empirical and experimental evidence on disappointment aversion is mixed. Camerer and Ho [1994] show in their survey of experimental studies that Gul's theory explains observed behavior better than standard expected utility[21]. Disappointment-averse preferences have attracted considerable attention in finance, particularly in the context of inter-temporal consumption–investment decisions. Much of this literature adopts a generalization of Gul's model introduced by Routledge and Zin [2010], that incorporates an additional parameter, allowing the threshold between disappointment and elation to differ from the certainty equivalent. However, Delikouras [2017] shows that the standard version of Gul's model with a piecewise linear utility function explains more than 90% of the observations in a dataset that combines information on returns of financial assets and aggregate consumption.

[21]Disappointment aversion (and in general a class of expectation-based reference-dependent preferences) can explain otherwise puzzling individual behavior even in the high-stakes decisions (see Hassidim et al [2021] and Dreyfuss, et al [2022]). Abdellaoui and Bleichrodt [2007] find that Loomes and Sugden's [1986] model of disapointment better fits their data.

8 Appendix B (Risk-Neutral Principal)

Proof of Proposition 1. (i) implies (iii). Assume that $\mathbf{p}_{\mathbf{a}_j} \notin \text{co}\{\mathbf{p}_{\mathbf{a}_i} : i \neq j\}$. Because this convex hull is compact, the Strict Hyperplane Separation Theorem yields a vector $\mathbf{y} \in \mathbb{R}^n$ such that

$$\mathbf{p}_{\mathbf{a}_j} \cdot \mathbf{y} > \mathbf{p}_{\mathbf{a}_i} \cdot \mathbf{y} \quad \text{for every } i \neq j.$$

Because $\mathbf{p}_{\mathbf{a}_i}$ is a probability vector for all i, the vector $\mathbf{y}$ cannot be a constant. Normalize y by setting

$$q_m = \frac{y_m - \min_\ell y_\ell}{\max_\ell y_\ell - \min_\ell y_\ell} \in [0,1] \text{ and define } \Pi_i = \sum_{m=1}^{n} p_{a_i}(x_m) q_m.$$

Since all probability vectors sum to one, this positive affine normalization preserves the strict inequalities, so that $\Pi_j > \Pi_i$ for every $i \neq j$.

Conditional on outcome x_m, let the contract pay a bonus $B \geq 0$ with probability q_m and pay zero otherwise. Under action a_i, the probability of receiving B is Π_i, and hence the equivalence between Gul's and Yaari's preferences for **binary** lotteries yields

$$CE[\mathbf{t}_{a_i}] = Bg(\Pi_i), \text{ where } g(Q) := \frac{Q}{1+\beta(1-Q)}.$$

Note also that

$$g'(Q) = \frac{1+\beta}{[1+\beta(1-Q)]^2} > 0.$$

Thus, the incentive constraint corresponding to action a_i is

$$B[g(\Pi_j) - g(\Pi_i)] \geq c(a_j) - c(a_i).$$

The smallest non-negative bonus satisfying all these constraints is

$$\bar{t}_j(\mathbf{q}) = \max\left\{0, \max_{i \neq j} \frac{c(a_j) - c(a_i)}{g(\Pi_j) - g(\Pi_i)}\right\}. \tag{2}$$

The bonus is finite because every denominator is positive. Therefore, for every upper bound on transfers $\bar{t} > \bar{t}_j(\mathbf{q})$, setting $B = \bar{t}$ implements a_j with a feasible binary contract while making all incentive constraints strict. Moreover, the normalization of the separating vector $\mathbf{y}$ implies that $q_m = 0$ for at least one outcome and that $q_m = 1$ for at least one other outcome. Hence the contract uses exactly the two payments 0 and $\bar{t}$, proving (iii).

(iii) implies (ii). This is immediate.

(ii) implies (i). We prove the contra-positive. Suppose that $\mathbf{p}_{\mathbf{a}_j} = \sum_{i \neq j} \alpha_i \mathbf{p}_{\mathbf{a}_i}$, where $\alpha_i \geq 0$ and $\sum_{i \neq j} \alpha_i = 1$. Under any contract, the compensation lottery induced by a_j is the corresponding mixture, with weights α_i, of the compensation lotteries induced by the other actions. The strict MLRP condition implies that $p_{a_i}(x_n)$ is strictly increasing in i. Hence the lowest and highest actions are identifiable, and a non-identifiable action must satisfy $1 < j < K$. Any mixture representing its distribution must assign positive weight to actions below a_j, so that

$$L := \sum_{i<j} \alpha_i > 0.$$

Set $\varepsilon = L/[2(1+\beta)K]$ and choose costs $c_i := c(a_i)$ given by

$$c_i = \begin{cases} \varepsilon(i-1), & i < j, \\ 1 + \varepsilon(i-1), & i \geq j. \end{cases}$$

These costs satisfy $c_1 = 0 < c_2 < \cdots < c_K$. Moreover, $c_j - c_i \geq 1$ for $i < j$, whereas $c_i - c_j = \varepsilon(i-j)$ for $i > j$.

Suppose, towards a contradiction, that a contract implements a_j, and write $w_i = CE[\mathbf{t}_{a_i}]$. Define

$$F_i(v) = \mathbb{E}[(\mathbf{t}_{a_i} - v) - \beta(v - \mathbf{t}_{a_i})^+].$$

The certainty-equivalent equation gives $F_i(w_i) = 0$. Here $w_j = CE[\mathbf{t}_{a_j}]$, whereas $F_i(w_j)$ evaluates the payment lottery induced by a_i at the threshold w_j. Since the slope of $y \mapsto y - \beta(-y)^+$ lies between 1 and $1+\beta$, for every $v \geq w$ we have

$$-(1+\beta)(v-w) \leq F_i(v) - F_i(w) \leq -(v-w).$$

Incentive compatibility requires $w_i \leq w_j + c_i - c_j$. For $i < j$, this implies $w_j - w_i \geq c_j - c_i \geq 1$, so the preceding slope bound gives

$$F_i(w_j) \leq F_i(w_i) - (w_j - w_i) \leq -1.$$

For $i > j$, put $d = c_i - c_j = \varepsilon(i-j) > 0$. Then $w_j \geq w_i - d$, so monotonicity and the slope bound imply

$$F_i(w_j) \leq F_i(w_i - d) \leq F_i(w_i) + (1+\beta)d = (1+\beta)\varepsilon(i-j).$$

Consequently,

$$F_i(w_j) \leq \begin{cases} -1, & i < j, \\ (1+\beta)\varepsilon(i-j), & i > j. \end{cases}$$

On the other hand, the mixture representation of $\mathbf{p}_{a_j}$ and the linearity of the expectation operator imply

$$\begin{aligned} 0 = F_j(w_j) &= \sum_{i \neq j} \alpha_i F_i(w_j) \\ &\leq -L + (1+\beta)\varepsilon \sum_{i>j} \alpha_i (i-j) \\ &\leq -L + (1+\beta)\varepsilon K = -\frac{L}{2} < 0, \end{aligned}$$

a contradiction. Hence a_j fails to be implementable for this strictly increasing cost vector, proving necessity. The same argument also establishes the equivalence when the admissible cost vectors are only required to be weakly increasing. □

The next six Lemmas will be used to establish Theorem 1:

Lemma 1 (Implementability under a sufficiently large transfer bound)**.** *The highest action a_K is implementable for every $\bar{t} \geq \bar{t}_0$.*

Proof. Consider the binary contract that pays T if x_n is realized and zero otherwise. Under action a, the induced transfer equals T with probability $p_a(x_n)$ and zero otherwise. Its certainty equivalent is given by

$$CE[\mathbf{t}_a] = Tg(p_a(x_n)).$$

This certainty equivalent lies in $[0, T]$, so that the defining certainty-equivalent equation becomes

$$CE[\mathbf{t}_a] = p_a(x_n)T - \beta(1 - p_a(x_n))CE[\mathbf{t}_a].$$

The function g is strictly increasing (see also the proof of Proposition 1) and setting $T = \bar{t}_0$ gives, for every $i < K$,

$$\begin{aligned} CE[\mathbf{t}_{a_K}] - CE[\mathbf{t}_{a_i}] &= \bar{t}_0 \left(g(p_{a_K}(x_n)) - g(p_{a_i}(x_n)) \right) \\ &\geq c(a_K) - c(a_i). \end{aligned}$$

Thus all incentive constraints hold. The participation constraint follows from the constraint against a_1, limited liability, and $c(a_1) = 0$. Finally, this binary contract is feasible whenever $\bar{t} \geq \bar{t}_0$. □

We say that a contract implementing action a is *suboptimal* if there exists another contract implementing the same action at a lower cost to the principal.

Lemma 2 (A downward constraint binds)**.** *Fix a target action a_r, $1 < r \leq K$. Every contract that implements a_r such that there is no binding downward incentive constraint (i.e., a constraint corresponding to an action a_i, $i < r$) is suboptimal.*

Proof. Recall that Gul's certainty equivalent is homogeneous:

$$CE[\lambda \mathbf{t}] = \lambda \, CE[\mathbf{t}] \text{ for every } \lambda \geq 0.$$

Suppose by contradiction that in an optimal contract τ all downward incentive constraints are slack. For $\lambda < 1$ sufficiently close to one, consider the scaled contract $\lambda\tau$. Scaling preserves admissibility, while the principal's expected payment under a_r strictly decreases.

For every $i < r$, slackness and homogeneity imply

$$\lambda \left(CE[\mathbf{t}_{a_r}] - CE[\mathbf{t}_{a_i}] \right) > c(a_r) - c(a_i)$$

for λ sufficiently close to one. Thus, every downward constraint remains satisfied. For every $j > r$, the upward constraint can be written as

$$CE[\mathbf{t}_{a_j}] - CE[\mathbf{t}_{a_r}] \leq c(a_j) - c(a_r).$$

If the left-hand side of the above inequality is non-positive, then scaling preserves the constraint. If it is positive, multiplication by $\lambda < 1$ relaxes the constraint. Finally, the downward constraint corresponding to action a_1 implies

$$CE[\lambda \mathbf{t}_{a_r}] - c(a_r) > CE[\lambda \mathbf{t}_{a_1}] - c(a_1) = CE[\lambda \mathbf{t}_{a_1}] \geq 0,$$

so that participation remains satisfied. The scaled contract is feasible and strictly cheaper, and the original contract was thus suboptimal. □

Lemma 3 (No intermediate payments below the certainty equivalent)**.** *Fix a target action a_r, $1 < r \leq K$, and a contract τ implementing a_r, and let $w := CE[\mathbf{t}_{a_r}]$. If τ pays with positive probability a transfer in the open interval $(0, w)$ then it is suboptimal.*

Proof. Suppose that after some outcome x_h the contract pays a transfer in $(0, w)$ with positive probability. Conditional on a realization where $s \in (0, w)$ is paid , replace s by the binary lottery that pays w with probability s/w and 0 otherwise. Leave all other transfers unchanged, and denote the induced compensation under action a by $\tilde{\mathbf{t}}_a$. The replacement is feasible because $w \leq \bar{t}$.

For every action a the replacement preserves the conditional mean. Under the target action it also preserves the value of $\mathbb{E}[(w - \mathbf{t}_a)^+]$ since

$$\left(1 - \frac{s}{w}\right) w = w - s.$$

Consequently, $\mathbb{E}[\tilde{\mathbf{t}}_a] = \mathbb{E}[\mathbf{t}_a]$ for every a and $CE[\tilde{\mathbf{t}}_{a_r}] = CE[\mathbf{t}_{a_r}] = w$.

Now fix a lower action a_i and set $u_i = CE[\mathbf{t}_{a_i}]$. We obtain

$$0 < u_i \leq w - \big(c(a_r) - c(a_i)\big) < w.$$

Incentive compatibility and the strict increase in cost give the upper bound in the equation above. The strict lower bound follows because transfers are all non-negative, all outcome probabilities are positive, and the affected positive transfer occurs with positive probability: For every $s \in (0, w)$ affected by the replacement, its contribution to the loss relative to the reference point u_i strictly increases. Indeed, if $s \geq u_i$, the increase is

$$\mathbb{E}[(u_i - \tilde{\mathbf{t}}_{a_i}))^+ | \mathbf{t}_{a_i} = s] = \left(1 - \frac{s}{w}\right) u_i > 0 = \mathbb{E}[(u_i - \mathbf{t}_{a_i}))^+ | \mathbf{t}_{a_i} = s],$$

whereas if $s < u_i$, it is

$$\begin{aligned}\mathbb{E}[(u_i - \tilde{\mathbf{t}}_{a_i}))^+ | \mathbf{t}_{a_i} = s] - \mathbb{E}[(u_i - \mathbf{t}_{a_i}))^+ | \mathbf{t}_{a_i} = s] &= \left(1 - \frac{s}{w}\right) u_i - (u_i - s) \\ &= s\left(1 - \frac{u_i}{w}\right) > 0.\end{aligned}$$

It follows from the certainty-equivalent equation that $CE[\tilde{\mathbf{t}}_{a_i}] < CE[\mathbf{t}_{a_i}]$ for every $i < r$.

For a higher action a_j, the replacement made above is a mean-preserving spread. Thus, for every u, it weakly increases the expectation $\mathbb{E}[(u - \mathbf{t}_{a_j})^+]$ while preserving the mean, and hence $CE[\tilde{\mathbf{t}}_{a_j}] \leq CE[\mathbf{t}_{a_j}]$ for every $j > r$.The modified contract is therefore feasible and has the same expected cost under the target action as the original contract, but all downward incentive constraints are slack. Thus, Lemma 2 implies that τ is suboptimal. □

Lemma 4 (No disappointing payment at the highest outcome)**.** *Let τ implement a_K, and write $w := CE[\mathbf{t}_{a_K}]$. If $\mathbb{P}[\tau_n < w] > 0$, then τ is suboptimal.*

Proof. Define

$$d := \mathbb{E}[(w - \tau_n)^+], \qquad D := \sum_i p_{a_K}(x_i)\mathbb{E}[(w - \tau_i)^+], \qquad e_i := \mathbb{E}[(\tau_i - w)^+].$$

The certainty-equivalent equation gives $\sum_i p_{a_K}(x_i)e_i = (1 + \beta)D$. Since $d > 0$, set

$\theta := p_{a_K}(x_n)d/D \in (0,1]$.

At x_n, replace every payment below w by w. At every outcome, replace every payment $s > w$ by $w + (1-\theta)(s-w)$, leaving all other payments unchanged. Both expected disappointment and expected elation under a_K are multiplied by $1-\theta$, so its certainty equivalent remains w. Write $\tilde{\mathbf{t}}_a$ for the transfer lottery under the modified contract and action a. Raising low payments at x_n increases expected compensation, while reducing payments above w decreases it. The net change under a_K is

$$\begin{aligned}
\mathbb{E}[\tilde{\mathbf{t}}_{a_K}] - \mathbb{E}[\mathbf{t}_{a_K}] &= p_{a_K}(x_n)\mathbb{E}[(w-\tau_n)^+] - \theta \sum_i p_{a_K}(x_i)\mathbb{E}[(\tau_i - w)^+] \\
&= p_{a_K}(x_n)d - \theta \sum_i p_{a_K}(x_i)e_i \\
&= p_{a_K}(x_n)d - \frac{p_{a_K}(x_n)d}{D}(1+\beta)D \\
&= -\beta p_{a_K}(x_n)d < 0.
\end{aligned}$$

To verify incentive compatibility, fix a lower action a and let $\tilde{w} := CE[\mathbf{t}_a] < w$ be its certainty equivalent under the original contract. Define

$$\Phi_a(z) := \mathbb{E}[\mathbf{t}_a] - z - \beta\mathbb{E}[(z-\mathbf{t}_a)^+],$$

and let $\tilde{\Phi}_a$ denote the corresponding function after the modification. These functions are strictly decreasing, with their respective certainty equivalents as their unique zeros:

$$\Phi_a(\tilde{w}) = 0, \qquad \tilde{\Phi}_a\big(CE[\tilde{\mathbf{t}}_a]\big) = 0.$$

In the comparison below, we evaluate both functions at the same value $\tilde{w} = CE[\mathbf{t}_a]$, the certainty equivalent under the original contract. The sign of $\tilde{\Phi}_a(\tilde{w})$ then determines whether the new certainty equivalent lies above or below this original value. Since $\tilde{w} < w$, reducing payments above w does not affect expected disappointment relative to $\tilde{w}$. Raising payments below w at x_n to w removes all such disappointment at that outcome. Thus

$$\mathbb{E}[(\tilde{w} - \tilde{\mathbf{t}}_a)^+] - \mathbb{E}[(\tilde{w} - \mathbf{t}_a)^+] = -p_a(x_n)\mathbb{E}[(\tilde{w} - \tau_n)^+].$$

Combining this with the change in expected transfers gives

$$\begin{aligned}\tilde{\Phi}_a(\tilde{w}) - \Phi_a(\tilde{w}) &= \mathbb{E}[\tilde{\mathbf{t}}_a] - \mathbb{E}[\mathbf{t}_a] \\ &\quad - \beta\Big(\mathbb{E}[(\tilde{w} - \tilde{\mathbf{t}}_a)^+] - \mathbb{E}[(\tilde{w} - \mathbf{t}_a)^+]\Big) \\ &= p_a(x_n)d - \theta \sum_i p_a(x_i)e_i + \beta p_a(x_n)\mathbb{E}[(\tilde{w} - \tau_n)^+] \\ &\leq p_a(x_n)d - (1+\beta)p_a(x_n)d + \beta p_a(x_n)\mathbb{E}[(\tilde{w} - \tau_n)^+] \\ &= \beta p_a(x_n)\big(\mathbb{E}[(\tilde{w} - \tau_n)^+] - d\big) < 0.\end{aligned}$$

The weak inequality follows from MLRP:

$$\theta \sum_i p_a(x_i)e_i \geq \theta \frac{p_a(x_n)}{p_{a_K}(x_n)} \sum_i p_{a_K}(x_i)e_i = (1+\beta)p_a(x_n)d.$$

The strict inequality follows from $\tilde{w} < w$ and $\mathbb{P}[\tau_n < w] > 0$. Thus $\tilde{\Phi}_a(\tilde{w}) < 0 = \Phi_a(\tilde{w})$. Since $\tilde{\Phi}_a$ is strictly decreasing, its zero $CE[\tilde{\mathbf{t}}_a]$ lies below $\tilde{w}$, so every lower action's certainty equivalent falls. The modification preserves participation and incentive compatibility while strictly reducing expected compensation. □

Lemma 5. *Let τ implement action a_K. If the highest outcome x_n is not the only outcome at which a transfer above $w := CE[\mathbf{t}_{a_K}]$ is paid with positive probability, then the contract is suboptimal.*

Proof. Suppose that at some outcome x_i with $i < n$ there is a transfer $r > w$ made with positive probability. Replace this transfer by w and simultaneously increase the transfer at x_n by

$$\frac{p_{a_K}(x_i)}{p_{a_K}(x_n)}\mathbb{P}[\tau_i = r](r - w).$$

Let $\widetilde{\tau}$ denote the resulting contract. By construction,

$$\mathbb{E}[\tilde{\mathbf{t}}_{a_K}] = \mathbb{E}[\mathbf{t}_{a_K}].$$

Moreover, the target action's certainty equivalent cannot decrease. Indeed, if $CE[\tilde{\mathbf{t}}_{a_K}] < w$, then

$$\begin{aligned}w - CE[\tilde{\mathbf{t}}_{a_K}] &= \beta\left(\mathbb{E}[(CE[\tilde{\mathbf{t}}_{a_K}] - \tilde{\mathbf{t}}_{a_K})^+] - \mathbb{E}[(w - \mathbf{t}_{a_K})^+]\right) \\ &\leq \beta\left(\mathbb{E}[(w - \tilde{\mathbf{t}}_{a_K})^+] - \mathbb{E}[(w - \mathbf{t}_{a_K})^+]\right) \leq 0,\end{aligned}$$

yielding a contradiction.

Now fix a lower action $a \in \{a_1, \ldots, a_{K-1}\}$ and let

$$\Delta := \mathbb{P}[\tau_i = r](r - w).$$

The change in the expected transfer after this action is taken is

$$\begin{aligned}\mathbb{E}[\tilde{\mathbf{t}}_a - \mathbf{t}_a] &= -p_a(x_i)\Delta + p_a(x_n)\frac{p_{a_K}(x_i)}{p_{a_K}(x_n)}\Delta \\ &= \Delta p_a(x_n)\left(\frac{p_{a_K}(x_i)}{p_{a_K}(x_n)} - \frac{p_a(x_i)}{p_a(x_n)}\right) < 0.\end{aligned}$$

where the last inequality follows from the strict MLRP condition:

$$\frac{p_{a_K}(x_i)}{p_a(x_i)} < \frac{p_{a_K}(x_n)}{p_a(x_n)}.$$

Both transfers involved in the perturbation are weakly above w, whereas $CE[\mathbf{t}_a] < w$. Hence the disappointment term under a is unchanged, and $CE[\tilde{\mathbf{t}}_a] < CE[\mathbf{t}_a]$ for every $a < a_K$ Thus, if the original contract was optimal, the modified contract is optimal and makes every downward incentive constraint slack, contradicting Lemma 2. □

Lemma 6 (The zero-payment region is a lower interval)**.** *Every contract implementing a_K is suboptimal if there does not exist $m \in \{1, \ldots, n-1\}$ such that a transfer of zero:*

1. *is paid with probability one at outcomes $x_1, \ldots, x_{m-1}$;*[22]
2. *is paid with probability $q \in (0, 1]$ at outcome x_m; and*
3. *is paid with probability zero at outcomes $x_{m+1}, \ldots, x_n$.*

Proof. Let m be the largest index in $\{1, \ldots, n-1\}$ such that $\mathbb{P}[\tau_m = 0] > 0$. Suppose there exists $j < m$ such that $\mathbb{P}[\tau_j = 0] < 1$. By Lemmas 3 and 5, the transfer at outcomes x_j and x_m is a lottery over $\{0, w\}$, where $w = CE[\mathbf{t}_{a_K}]$.

Decrease the probability of paying zero at x_m by $\epsilon > 0$ and increase the probability of paying zero at x_j by

$$\epsilon\frac{p_{a_K}(x_m)}{p_{a_K}(x_j)}.$$

For ϵ sufficiently small this defines a feasible contract $\widetilde{\tau}$. The adjustment preserves both the expected transfer and the certainty equivalent under a_K.

[22] If $m = 1$, this set of outcomes is empty.

For any lower action a, the change in expected transfer is

$$\begin{aligned}\mathbb{E}[\tilde{\mathbf{t}}_a - \mathbf{t}_a] &= \epsilon w \left(p_a(x_m) - p_a(x_j) \frac{p_{a_K}(x_m)}{p_{a_K}(x_j)} \right) \\ &= \epsilon w p_a(x_j) \left(\frac{p_a(x_m)}{p_a(x_j)} - \frac{p_{a_K}(x_m)}{p_{a_K}(x_j)} \right) < 0,\end{aligned}$$

where the inequality follows from the strict MLRP condition. Moreover, because zero is the only transfer below $CE[\mathbf{t}_a]$, the loss relative to the old certainty equivalent strictly increases:

$$\begin{aligned}&\mathbb{E}[(CE[\mathbf{t}_a] - \tilde{\mathbf{t}}_a)^+] - \mathbb{E}[(CE[\mathbf{t}_a] - \mathbf{t}_a)^+] \\ &= \epsilon CE[\mathbf{t}_a] \left(p_a(x_j) \frac{p_{a_K}(x_m)}{p_{a_K}(x_j)} - p_a(x_m) \right) > 0.\end{aligned}$$

It follows that $CE[\tilde{\mathbf{t}}_a] < CE[\mathbf{t}_a]$ for every $a < a_K$. The modified contract is optimal and makes every downward constraint slack, contradicting Lemma 2. □

Proof of Theorem 1: For future reference we define

$$\bar{t}_0 := \max_{i<K} \frac{c(a_K) - c(a_i)}{g(p_{a_K}(x_n)) - g(p_{a_i}(x_n))}, \tag{3}$$

where $g : [0,1] \to [0,1]$ is defined by

$$g(p) := \frac{p}{1 + \beta(1-p)}. \tag{4}$$

By the strict MLRP condition we know that $p_{a_K}(x_n) > p_{a_i}(x_n)$ for all $i \neq K$, and that $\bar{t}_0$ is finite. In particular, the condition for implementability from Proposition 1 is satisfied.

As established in Lemmas 1 to 6, every optimal contract (even those which are risk-dominated) has the structure displayed in Theorem 1 at all values below the highest outcome. However, as long as the expectation of τ_K remains unchanged, an optimal contract could randomize over transfers above $CE[t_{a_K}]$ at the highest outcome. Any such randomization leaves the certainty equivalent of the agent unchanged, but exposes the agent to a strictly greater risk in the mean preserving spread sense. Thus, any contract for which τ_K is random is risk-dominated**.** □

Proposition 6 (Two versus three payment levels)**.** *Under the assumptions of Theorem 1, let I_0 be the maximizers in* (3)*, with g defined in* (4)*, and set $A_i := 1 + \beta(1 - p_{a_i}(x_n))$.*

Every risk-undominated optimum uses three payment levels if and only if, for every $i \in I_0$,

$$\frac{p_{a_K}(x_n)p_{a_i}(x_{n-1}) - p_{a_i}(x_n)p_{a_K}(x_{n-1})}{p_{a_K}(x_n) - p_{a_i}(x_n)} < \beta\left(\frac{p_{a_K}(x_{n-1})}{A_K} + \frac{p_{a_i}(x_{n-1})}{A_i}\right).$$

Otherwise, the binary contract paying $\bar{t}_0$ at x_n and zero elsewhere is optimal, and then every risk-undominated optimum uses only two payments.

Proof: In the first step we establish that it is enough to check whether there exists a contract with three payment levels that implements a_K at the cost of the optimal binary contract. For $i < K$, denote

$$\rho_i := \frac{p_{a_i}(x_n)}{p_{a_K}(x_n)}, \qquad \ell_i := \frac{p_{a_i}(x_{n-1})}{p_{a_K}(x_{n-1})}, \qquad H_i := \frac{\ell_i - \rho_i}{1-\rho_i} - \frac{\beta}{A_K} - \frac{\beta\ell_i}{A_i}.$$

The inequality in the statement is equivalent to $H_i < 0$. Let $C_0 := p_{a_K}(x_n)\bar{t}_0$ be the cost of the binary contract in the statement. If a cheaper three-payment contract implements a_K, homogeneity lets us scale its payments up in order to obtain another such contract costing exactly C_0. Thus , in order to rule out that a three payment contract is optimal, it suffices to rule out the existence of a three-payment contract with cost C_0.

By Theorem 1, it suffices to consider contracts with payments in $\{0, w, T\}$, where $w = CE[\mathbf{t}_{a_K}]$ and $T > w$ is paid only at x_n. Let s_i be the probability of receiving w under a_i, for $i = 1, \ldots, K$. In the binary case is $s_K = 0$, whereas with three payments we have $s_K > 0$. At expected cost C_0, the target's certainty-equivalent equation determines

$$w = \frac{C_0}{A_K - \beta s_K}, \qquad T = \frac{w[A_K - (1+\beta)s_K]}{p_{a_K}(x_n)}.$$

In the next steps we shall rewrite the IC constraints for all actions a_i with $i \in I_0$. Binary incentive compatibility and full support imply $w \geq C_0/A_K > c(a_K)$. Fix $i \in I_0$ and set

$$\bar{w}_i := w - c(a_K) + c(a_i) \in (0, w).$$

The incentive constraint against a_i requires $CE[\mathbf{t}_{a_i}] \leq \bar{w}_i$. Since $\bar{w}_i < w < T$, zero is the only payment below $\bar{w}_i$, and its probability under a_i is $1 - p_{a_i}(x_n) - s_i$. The certainty-equivalent residual decreases with its argument, so this constraint is equivalent to

$$ws_i + p_{a_i}(x_n)T \leq \bar{w}_i\big[1 + \beta(1 - p_{a_i}(x_n) - s_i)\big] = \bar{w}_i(A_i - \beta s_i).$$

Substituting $\bar{w}_i$ and $p_{a_K}(x_n)T = w[A_K - (1+\beta)s_K]$, and using $A_i - \rho_i A_K = (1+\beta)(1-\rho_i)$, gives

$$(1+\beta)w(1 - \rho_i + \rho_i s_K - s_i) \geq [c(a_K) - c(a_i)](A_i - \beta s_i).$$

In the binary contract, $s_K = s_i = 0$ and $w = C_0/A_K$. Since $i \in I_0$, its incentive constraint binds there, yielding

$$(1+\beta)C_0(1-\rho_i) = [c(a_K) - c(a_i)]A_K A_i.$$

Using this identity to eliminate the cost difference, and substituting $w = C_0/(A_K - \beta s_K)$, the preceding incentive inequality becomes

$$1 + \frac{\rho_i s_K - s_i}{1-\rho_i} \geq \left(1 - \frac{\beta s_K}{A_K}\right)\left(1 - \frac{\beta s_i}{A_i}\right).$$

Expanding the product, cancelling the constant term, and collecting terms gives that the IC constraint for action a_i with $i \in I_0$ is equivalent to

$$s_K\left(\frac{\beta}{A_K} + \frac{\rho_i}{1-\rho_i}\right) \geq s_i\left(\frac{1}{1-\rho_i} - \frac{\beta}{A_i} + \frac{\beta^2 s_K}{A_K A_i}\right).$$

The coefficient multiplying s_i is positive, since $A_i - \beta(1-\rho_i) = 1 + \beta\rho_i(1 - p_{a_K}(x_n)) > 0$. MLRP gives $s_i \geq \ell_i s_K$. Consequently, when $s_K > 0$, the displayed constraint requires

$$H_i + \frac{\beta^2 \ell_i s_K}{A_K A_i} \leq 0.$$

This condition is also sufficient for that constraint when $s_i = \ell_i s_K$.

Only if part: If $H_i \geq 0$ for some $i \in I_0$, this condition fails for every $s_K > 0$. Hence, every three-payment contract with cost equal or below the cost of the optimal binary contract C_0 will violate an incentive constraint. The binary contract is therefore optimal, and every risk-undominated optimum uses two payments, including cases where $H_i = 0$.

If part: If $H_i < 0$ for every $i \in I_0$, choose a sufficiently small $s_K > 0$ and pay w with probability $s_K/p_{a_K}(x_{n-1})$ at x_{n-1}, and zero otherwise below x_n. Then $s_i = \ell_i s_K$ for every $i < K$, so every initially binding incentive constraint becomes slack. Initially slack constraints remain slack by continuity. Scaling all payments down slightly therefore preserves incentive compatibility and reduces cost below C_0. Theorem 1 implies that every risk-undominated optimum uses three payments. □

Corollary 1 (A sufficient condition for three payments)**.** *Under the assumptions of Theorem 1, every risk-undominated optimal contract implementing a_K uses three payments if*

$$\max_{i<K} p_{a_i}(x_{n-1}) \leq \min\{1, \beta\}\, p_{a_K}(x_{n-1}).$$

Proof. Using the notation of the proof of Proposition 6, the condition gives $\ell_i \leq 1$ and $\ell_i \leq \beta$ for every $i < K$. Hence $(\ell_i - \rho_i)/(1-\rho_i) \leq \ell_i$, while full support implies

$A_i, A_K < 1 + \beta$. Therefore,

$$H_i < \ell_i - \frac{\beta(1+\ell_i)}{1+\beta} = \frac{\ell_i - \beta}{1+\beta} \leq 0.$$

Proposition 6 applies to every possible binding deviation. □

Lemma 7 (A sufficiently large transfer bound does not bind)**.** *There exists an optimal contract for the problem with $\bar{t} = \infty$ that is feasible for every $\bar{t} \geq \bar{t}_0$. Consequently, for every $\bar{t} \geq \bar{t}_0$, both the bounded and the unbounded problems have the same value.*

Proof. The binary contract from Lemma 1 concentrates all compensation on x_n and leads to an expected payment of $p_{a_K}(x_n)\bar{t}_0$. Since any contract which has higher expected cost is suboptimal we can without loss restrict attention to contracts satisfying,

$$\mathbb{E}[\mathbf{t}_{a_K}] \leq p_{a_K}(x_n)\bar{t}_0.$$

Let $w = CE[\mathbf{t}_{a_K}]$. By limited liability,

$$\mathbb{E}[\mathbf{t}_{a_K}] = \sum_{i=1}^{n} p_{a_K}(x_i)\mathbb{E}[\tau_i] \geq p_{a_K}(x_n)\mathbb{E}[\tau_n].$$

Combining the two inequalities gives that one can without loss restrict attention to contract satisfying $\mathbb{E}[\tau_n] \leq \bar{t}_0$. Also, $w \leq \mathbb{E}[\mathbf{t}_{a_K}] \leq \bar{t}_0$, because the certainty equivalent is always weakly below the mean and because $p_{a_K}(x_n) \leq 1$.

By Lemmas 3, 5, and 6, we can without loss restrict attention to contracts such that payments below x_n belong to $\{0, w\}$ and payments at x_n are weakly above w (as any other contract is suboptimal). Replace the possibly random payment at x_n by the deterministic payment $\mathbb{E}[\tau_n]$. This preserves all expected payments and certainty equivalents, while making every transfer weakly smaller than $\bar{t}_0$. As the set of (random) contracts where all payments lie in the compact interval $[0, \bar{t}_0]$ is compact an optimal contract exists. Furthermore, this optimum of the problem without an upper-bound is feasible for every $\bar{t} \geq \bar{t}_0$, and such a bound does not change the optimal value. □

Proof of Proposition 2: By the structure of the optimal contract, the only payment below w is the zero payment, and that payment occurs with probability π_0 under action a_K. The only payment above w is made at the highest outcome x_n. Since $w = CE[\mathbf{t}_{a_K}]$, Gul's certainty-equivalent equation gives

$$\mathbb{E}\left[(\mathbf{t}_{a_K} - w)^+\right] = (1+\beta)\,\mathbb{E}\left[(w - \mathbf{t}_{a_K})^+\right].$$

The left-hand side is $p_{a_K}(x_n)B$, while the right-hand side is $(1+\beta)\pi_0 w$. Dividing by $p_{a_K}(x_n)w > 0$ gives the result in (1). □

Proof of Proposition 3. Because the outcome set is finite and because all transfers lie in the compact interval $[0, \bar{t}]$, the set of feasible randomized contracts implementing a_r is compact and non-empty. The objective and the incentive constraints are continuous, so implementability implies the existence of an optimal contract. Fix one such contract and write $w = CE[\mathbf{t}_{a_r}]$.

By Lemma 3, no optimal contract uses a transfer in $(0, w)$. If $w = \bar{t}$, the result is immediate. Suppose therefore that $w < \bar{t}$, and consider a transfer $s \in (w, \bar{t})$ paid with positive probability after some outcome. Replace it, conditional on the same event, by the lottery that pays $\bar{t}$ with probability

$$\frac{s-w}{\bar{t}-w}$$

and w with the complementary probability. The conditional mean and the target action's expected gain above w are unchanged. Hence both $CE[\mathbf{t}_{a_r}]$ and the principal's expected payment under a_r are unchanged.

Every lower action has a certainty equivalent strictly below w, so that the replacement leaves its certainty equivalent unchanged. For a higher action a_i, the replacement is a mean-preserving spread and therefore weakly lowers its certainty equivalent. Thus, applying the replacement to every transfer in $(w, \bar{t})$ yields an optimal contract supported on $\{0, w, \bar{t}\}$.

For the final claim, suppose that an optimal contract uses a transfer in $(w, \bar{t})$. Every binding upward constraint is associated with an action a_i satisfying

$$CE[\mathbf{t}_{a_i}] = w + c(a_i) - c(a_r) > w.$$

The replacement strictly lowers the certainty equivalent of every such action and therefore makes all previously binding upward constraints slack; constraints that were already slack remain slack. The transformed contract is consequently an optimal contract for which all upward constraints are slack, contrary to the maintained assumption. Thus no optimal contract can use a transfer in $(w, \bar{t})$. □

Proof of Proposition 4. The set of monotone contracts form a closed subset of the set of feasible randomized contracts. Hence monotone implementability and compactness imply the existence of a contract that is optimal within this subset. Moreover, multiplying every payment lottery by the same scalar $\lambda \in [0, 1]$ preserves the first-order stochastic-dominance order. Therefore the proof of Lemma 2 applies within this class: at least one

downward incentive constraint binds at every optimal monotone contract.

Fix such a contract and write $w = CE[\mathbf{t}_{a_r}]$. Consider first the stochastic kernel

$$K^-(s) = \begin{cases} \left(1 - \frac{s}{w}\right)\delta_0 + \frac{s}{w}\delta_w, & 0 < s < w, \\ \delta_s, & \text{otherwise.} \end{cases}$$

The case $w = 0$ is immediate. For $w > 0$, the lotteries $K^-(s)$ are increasing in s in the sense of first-order stochastic dominance. Indeed, for every increasing function f, the function $s \mapsto \int f(y)K^-(s, dy)$ is increasing. Consequently, applying this same kernel to the payment lottery at every outcome preserves the monotonicity of the contract. The calculations in the proof of Lemma 3 show that the transformation preserves the target action's certainty equivalent and expected transfer, weakly lowers every higher action's certainty equivalent, and strictly lowers every lower action's certainty equivalent whenever a payment in $(0, w)$ occurs. Such a payment would therefore produce an optimal monotone contract at which all downward incentive constraints are slack, a contradiction. Thus, no optimal monotone contract uses a payment in $(0, w)$.

If $w = \bar{t}$, the first conclusion follows. Otherwise, consider the kernel

$$K^+(s) = \begin{cases} \frac{\bar{t} - s}{\bar{t} - w}\delta_w + \frac{s - w}{\bar{t} - w}\delta_{\bar{t}}, & w < s < \bar{t}, \\ \delta_s, & \text{otherwise.} \end{cases}$$

This kernel is also increasing in s in the sense of first-order stochastic dominance, and hence preserves monotonicity when applied at every outcome. As in the proof of Proposition 3, it preserves the target action's certainty equivalent and expected transfer, leaves every lower action's certainty equivalent unchanged, and weakly lowers every higher action's certainty equivalent. Applying it to an optimal monotone contract therefore gives an optimal monotone contract supported on $\{0, w, \bar{t}\}$.

Finally, suppose that an optimal monotone contract uses a payment in $(w, \bar{t})$. Applying K^+ strictly relaxes every binding upward constraint and leaves all other constraints satisfied. It consequently produces an optimal monotone contract at which all upward constraints are slack. Under the assumption in the final sentence of the corollary, this is impossible. Hence every optimal monotone contract has then a support contained in the set $\{0, w, \bar{t}\}$. □

The last Proposition in this Appendix relates to Remark 1 in the main text. Fix a target action a_r and write $w = CE[\mathbf{t}_{a_r}]$. By the three-payment proposition, let π_h and z_h be the probabilities of paying 0 and $\bar{t}$ at outcome x_h; the remaining probability is

assigned to w. Denote by

$$\Pi_i := \sum_{h=1}^{n} p_{a_i}(x_h)\pi_h, \text{ and } Z_i := \sum_{h=1}^{n} p_{a_i}(x_h)z_h.$$

For every $i \neq r$, define the cost difference $d_i := c(a_r) - c(a_i)$.

Proposition 7. *The cheapest contract implementing a_r solves*

$$\begin{aligned}
\min_{w,\pi,\mathbf{z}} \quad & w(1+\beta\Pi_r) \\
\textit{subject to} \quad & (\bar{t}-w)Z_r = (1+\beta)w\Pi_r, \\
& \big[(1+\beta)w - \beta d_i\big]\Pi_i - (\bar{t}-w)Z_i \geq d_i, && i < r, \\
& (\bar{t}-w-\beta d_i)Z_i - (1+\beta)w\Pi_i \leq -(1+\beta)d_i, && i > r, \\
& \pi_h, z_h \geq 0, \qquad \pi_h + z_h \leq 1, && h = 1,\dots,n, \\
& c(a_r) \leq w \leq \bar{t}.
\end{aligned}$$

For fixed w, this is a finite linear program in $(\pi, \mathbf{z})$; the full problem is a one-dimensional minimization over w. For monotone contracts, add the constraints:

$$\pi_h \geq \pi_{h+1}, \qquad z_h \leq z_{h+1}, \qquad h = 1,\dots,n-1.$$

Proof: The expected transfer under a_i is $\mu_i = w - w\Pi_i + (\bar{t}-w)Z_i$. The function

$$u \mapsto \mu_i - u - \beta\mathbb{E}_{a_i}[(u - \mathbf{t}_{a_i})^+]$$

is strictly decreasing and vanishes at $CE[\mathbf{t}_{a_i}]$. Hence

$$CE[\mathbf{t}_{a_i}] \leq u \iff \mu_i - u - \beta\mathbb{E}_{a_i}[(u - \mathbf{t}_{a_i})^+] \leq 0.$$

For the target action, setting $u = w$ gives

$$(\bar{t}-w)Z_r = (1+\beta)w\Pi_r, \text{ and } \mathbb{E}_{a_r}[\mathbf{t}_{a_r}] = w(1+\beta\Pi_r).$$

The IC constraint for action a_i is

$$CE[\mathbf{t}_{a_i}] \leq w - d_i.$$

If $i < r$, then $d_i > 0$ and only zero lies below $w - d_i$, so

$$\mathbb{E}_{a_i}[(w - d_i - \mathbf{t}_{a_i})^+] = (w - d_i)\Pi_i.$$

If $i > r$, then $d_i < 0$. Provided that $w - d_i < \bar{t}$, both zero and w lie below $w - d_i$, and therefore

$$\mathbb{E}_{a_i}[(w - d_i - \mathbf{t}_{a_i})^+] = w\Pi_i - d_i(1 - Z_i).$$

Substitution gives the two IC inequalities in the program below. If $i > r$ and $w - d_i \geq \bar{t}$, the upward constraint is automatic and its displayed inequality is redundant.

Consequently, the cheapest contract implementing a_r solves

$$\begin{aligned}
\min_{w,\mathbf{q},\mathbf{z}} \quad & w(1 + \beta\Pi_r) \\
\text{subject to} \quad & (\bar{t} - w)Z_r = (1 + \beta)w\Pi_r, \\
& \big[(1 + \beta)w - \beta d_i\big]\Pi_i - (\bar{t} - w)Z_i \geq d_i, && i < r, \\
& (\bar{t} - w - \beta d_i)Z_i - (1 + \beta)w\Pi_i \leq -(1 + \beta)d_i, && i > r, \\
& \pi_h, z_h \geq 0, \qquad \pi_h + z_h \leq 1, && h = 1, \ldots, n, \\
& c(a_r) \leq w \leq \bar{t}.
\end{aligned}$$

For fixed w, this is a finite linear program in $(\pi, \mathbf{z})$ and the full problem is a one-dimensional minimization over w. For monotone contracts, we add to the constraints

$$\pi_h \geq \pi_{h+1}, \text{ and } z_h \leq z_{h+1}, \qquad h = 1, \ldots, n-1. \quad \square$$

9 Appendix C (Proof of Proposition 5)

The Proof of Proposition 5 follows by combining several Lemmas. First, Lemma 8 shows that at least one incentive constraint must bind in an optimal contract. Lemmas 9-12 show that, for any contract that implements the highest acton a_K, there exist a contract that also implements a_K and generates (weakly) higher utility to the principal such that all payments besides the payment after the highest outcome, x_n, are concentrated on the set $\{0, CE^a[t_{a_K}], x_i - CE^p\left[(\mathbf{x} - \mathbf{t})_{a_K}\right]\}$. Lemmas 13-15 show that for any contract with such a support, the principal's payoff can be improved unless the compensation scheme has the following monotone structure: there exist three cutoffs such that: 1) if the realized outcome is below the lowest outcome the compensation is 0; 2) if the realized output is between the lowest and the middle cutoff the compensation is $x_i - CE^p\left[(\mathbf{x} - \mathbf{t})_{a_K}\right]$; 3) if the realized outcome is between the middle and the highest cutoff the compensation is $CE^a[t_{a_K}]$; 4) if the outcome is above the highest cutoff, but different from the highest outcome, the compensation is $x_i - CE^p\left[(\mathbf{x} - \mathbf{t})_{a_K}\right]$. If the outcome coincides with one of the cutoffs, then the compensation is a lottery over the neighboring intervals.

Lemma 8. *In every optimal contract implementing a_K, at least one of the incentive constraints IC_j must bind.*

Proof: Assume, by contradiction, that there is an optimal contract with no IC binding constraints. We show that, if this is the case, the optimal solution is a fixed wage at $c(a_K)$ - but, such contract cannot implement the highest action.

Ignoring the IC constraints, the optimal contract solves

$$\max_{\tau} CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]$$
$$s.t.\ CE^a\left[\mathbf{t}_{a_K}\right] - c(a_K) \geq 0 \quad (IR)$$

where the distributions over outcomes (and hence also over transfers) are specified by p_{a_K}.

Recall that

$$CE^a\left[\mathbf{t}_{a_K}\right] = \mathbb{E}\left[\mathbf{t}_{a_K}\right] - \beta^a \mathbb{E}\left[\left(CE^a\left[\mathbf{t}_{a_K}\right] - \mathbf{t}_{a_K}\right)^+\right].$$

At the optimum, the IR constraint binds, as otherwise we can decrease all the payments by $\varepsilon > 0$ and increase the principal's utility. Hence, $CE^a\left[\mathbf{t}_{a_K}\right] = c(a_K)$.

We now claim that the solution to the above problem is a fixed payment, $\tau_i = c(a_K)$, for all i. That is, the principal absorbs all the risk and the (more) disappointment averse agent obtains full insurance. Assume that this is not the case. That is, there are two outcomes x_i and x_j such that $\mathbb{P}\left[\tau_i < c(a_K)\right] > 0$ and $\mathbb{P}\left[\tau_j > c(a_K)\right] > 0$. Denote by r_i a realization of τ_i such that $r_i < c(a_K)$ and denote by r_j a realization of τ_j with $r_j > c(a_K)$. Consider the following modification of the compensation scheme, and define $\tilde{\mathbf{t}}$ to be the new associated random variable expressing the payment to the agent: i) instead of transfer of r_i, the compensation is $c(a_K)$ with probability $\alpha > 0$, and with probability $1-\alpha$ it is unchanged and equals r_i; ii) instead of transfer of r_j, the compensation is $c(a_K)$ with probability $\gamma > 0$, and with probability $1-\gamma$ it is unchanged and equals r_j. iii) all other transfers are not changed. We choose α and γ such that the agent's utility remains the same, that is:

$$\begin{aligned} & p_{a_K}(x_i)\,\mathbb{P}\left[\tau_i = r_i\right] r_i + p_{a_K}(x_j)\,\mathbb{P}\left[\tau_j = r_j\right] r_j - \beta^a p_{a_K}(x_i)\,\mathbb{P}\left[\tau_i = r_i\right]\left(c(a_K) - r_i\right) \\ = \;& p_{a_K}(x_i)\,\mathbb{P}\left[\tau_i = r_i\right]\left(\alpha c(a_K) + (1-\alpha) r_i\right) + p_{a_K}(x_j)\,\mathbb{P}\left[\tau_j = r_j\right]\left(\gamma c(a_K) + (1-\gamma) r_j\right) \\ & -\beta^a p_{a_K}(x_i)(1-\alpha)\,\mathbb{P}\left[\tau_i = r_i\right]\left(c(a_K) - r_i\right). \end{aligned}$$

Therefore, the probabilities α and γ must satisfy

$$\gamma p_{a_K}(x_j)\,\mathbb{P}\left[\tau_j = r_j\right]\left(r_j - c(a_K)\right) = \alpha p_{a_K}(x_i)\,\mathbb{P}\left[\tau_i = r_i\right]\left(c(a_K) - r_i\right)\left(1+\beta^a\right). \qquad (5)$$

The above modification decreases the expected compensation $\mathbb{E}\left[\mathbf{t}_{a_K}\right]$ since the change in the agent's utility is given by

$$\begin{aligned}
& p_{a_K}(x_i)\,\mathbb{P}\left[\tau_i = r_i\right](\alpha c(a_K) + (1-\alpha) r_i) + p_{a_K}(x_j)\,\mathbb{P}\left[\tau_j = r_j\right](\gamma c(a_K) + (1-\gamma) r_j) \\
& -p_{a_K}(x_i)\,\mathbb{P}\left[\tau_i = r_i\right] r_i - p_{a_K}(x_j)\,\mathbb{P}\left[\tau_j = r_j\right] r_j \\
= \;& p_{a_K}(x_i)\,\mathbb{P}\left[\tau_i = r_i\right]\alpha(c(a_K) - r_i) + p_{a_K}(x_j)\,\mathbb{P}\left[\tau_j = r_j\right]\gamma(c(a_K) - r_j) \\
= \;& p_{a_K}(x_i)\,\mathbb{P}\left[\tau_i = r_i\right]\alpha(c(a_K) - r_i) - \alpha p_{a_K}(x_i)\,\mathbb{P}\left[\tau_i = r_i\right](c(a_K) - r_i)(1+\beta^a) \\
= \;& -\alpha p_{a_K}(x_i)\,\mathbb{P}\left[\tau_i = r_i\right](c(a_K) - r_i)\,\beta^a < 0,
\end{aligned}$$

where the third line follows from (5). We now show that the modification increases the principal's utility. Recall that

$$CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] = \mathbb{E}\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] - \beta^p \mathbb{E}\left[\left(CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] - (\mathbf{x}-\mathbf{t})_{a_K}\right)^+\right].$$

As we already showed, the modification increases $\mathbb{E}\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]$ by $\beta^a \alpha p_{a_K}(x_i)\left[\tau_i = r_i\right](c(a_K) - r_i)$. We show now that $CE^p\left[\left(\mathbf{x}-\tilde{\mathbf{t}}\right)_{a_K}\right] > CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]$. Suppose that $CE^p\left[\left(\mathbf{x}-\tilde{\mathbf{t}}\right)_{a_K}\right] \leq CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]$. We then obtain:

$$\begin{aligned}
& CE^p\left[\left(\mathbf{x}-\tilde{\mathbf{t}}\right)_{a_K}\right] = \mathbb{E}\left[\left(\mathbf{x}-\tilde{\mathbf{t}}\right)_{a_K}\right] - \beta^p \mathbb{E}\left[\left(CE^p\left[\left(\mathbf{x}-\tilde{\mathbf{t}}\right)_{a_K}\right] - \left(\mathbf{x}-\tilde{\mathbf{t}}\right)_{a_K}\right)^+\right] \\
\geq\; & \mathbb{E}\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] + \beta^a \alpha p_{a_K}(x_i)\,\mathbb{P}\left[\tau_i = r_i\right](c(a_K) - r_i) - \beta^p \mathbb{E}\left[\left(CE^p(\mathbf{x}-\mathbf{t})_{a_K} - \left(\mathbf{x}-\tilde{\mathbf{t}}\right)_{a_K}\right)^+\right] \\
=\; & \mathbb{E}\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] + \beta^a \alpha p_{a_K}(x_i)\,\mathbb{P}\left[\tau_i = r_i\right](c(a_K) - r_i) - \beta^p \mathbb{E}\left[\left(CE^p(\mathbf{x}-\mathbf{t})_{a_K} - (\mathbf{x}-\mathbf{t})_{a_K}\right)^+\right] \\
& -\beta^p p_{a_K}(x_i)\,\mathbb{P}\left[\tau_i = r_i\right]\alpha\left[\left(CE^p(\mathbf{x}-\mathbf{t})_{a_K} - (x_i - c(a_K))\right)^+ - \left(CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] - (x_i - r_i)\right)^+\right] \\
& -\beta^p p_{a_K}(x_j)\,\mathbb{P}\left[\tau_j = r_j\right]\gamma\left[\left(CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] - (x_j - c(a_K))\right)^+ - \left(CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] - (x_j - r_j)\right)^+\right] \\
\geq\; & \mathbb{E}\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] + \beta^a \alpha p_{a_K}(x_i)\,\mathbb{P}\left[\tau_i = r_i\right](c(a_K) - r_i) - \beta^p \mathbb{E}\left[\left(CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] - (\mathbf{x}-\mathbf{t})_{a_K}\right)^+\right] \\
& -\beta^p p_{a_K}(x_i)\,\mathbb{P}\left[\tau_i = r_i\right]\alpha\left[\left(CE^p(\mathbf{x}-\mathbf{t})_{a_K} - (x_i - c(a_K))\right)^+ - \left(CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] - (x_i - r_i)\right)^+\right] \\
>\; & \mathbb{E}\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] - \beta^p \mathbb{E}\left[\left(CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] - (\mathbf{x}-\mathbf{t})_{a_K}\right)^+\right] = CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right],
\end{aligned}$$

where the second inequality follows because $r_j > c(a_K)$ implies

$$\left(CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] - (x_j - c(a_K))\right)^+ - \left(CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] - (x_j - r_j)\right)^+ \leq 0,$$

and where the third inequality follows from

$$\beta^a\left(c\left(a_K\right)-r_i\right)>\beta^p\left(c\left(a_K\right)-r_i\right)$$
$$\geq-\beta^p\left(CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]-\left(x_i-c\left(a_K\right)\right)\right)^+ +\beta^p\left(CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]-\left(x_i-r_i\right)\right)^+.$$

The above chain of the inequalities leads to a contradiction. Therefore, $CE^p\left[\left(\mathbf{x}-\tilde{\mathbf{t}}\right)_{a_K}\right]>CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]$.

Hence, the proposed modification increases the principal's utility, yielding a contradiction to the optimality of the original contract. Finally, although we did not use here limited liability, the constant payment of $c\left(a_K\right)\geq 0$ satisfies this constraint. □

Fix now a contract τ that implements a_K. In Lemmas 9-12 below we show that if this contract contains payoffs elements outside the set $\{0, CE[t_{a_K}], x_i - CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]\}$, then there exists an alternative contract that also implements a_K and generates a (weakly) higher payoff to the principal where thosepayoffs are replaced with payoffs in the set $\{0, CE[t_{a_K}], x_i - CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]\}$.

Lemma 9. *1. If there exists an outcome x_i such that $x_i - CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]\leq 0$ and*

$$\mathbb{P}\left[\tau_i\in\left(0, CE^a\left[\mathbf{t}_{a_K}\right]\right)\right]>0,$$

there there exists another contract $\hat{\tau}$ that implements a_K in which

$$\mathbb{P}\left[\hat{\tau}_i\in\left(0, CE^a\left[\hat{\mathbf{t}}_{a_K}\right]\right)\right]=0 \text{ and}$$
$$\hat{\tau}_j=\tau_j \text{ for any } j\neq i.$$

Moreover, $CE^p\left[\left(\mathbf{x}-\hat{\mathbf{t}}\right)_{a_K}\right]\geq CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]$.

2. If there exists an outcome x_i such that $x_i - CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]>0$ and

$$\mathbb{P}\left[\tau_i\in\left(0, \min\left\{CE^a\left[\mathbf{t}_{a_K}\right],\ x_i-CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]\right\}\right)\right]>0,$$

then there exists another contract $\hat{\tau}$ that implements a_K in which

$$\mathbb{P}\left[\hat{\tau}_i\in\left(0, \min\left\{CE^a\left[\hat{\mathbf{t}}_{a_K}\right],\ x_i-CE^p\left[\left(\mathbf{x}-\hat{\mathbf{t}}\right)_{a_K}\right]\right\}\right)\right]=0 \text{ and}$$
$$\hat{\tau}_j=\tau_j \text{ for any } j\neq i.$$

Moreover, $CE^p\left[\left(\mathbf{x}-\hat{\mathbf{t}}\right)_{a_K}\right]\geq CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]$.

Proof: This is similar to the proof of Lemma 3, Theorem 1. We add it here for completeness, with a slight adaptation (mainly at point 2).

1. Consider the modification of the payments offered in Lemma 3. Observe that in this case, all perturbed payments belong to the disappointment interval of the principal. Therefore, the modification offered increases the agent's incentives (as shown there) and it also preserves the principal's utility since it keeps the expected transfer the same.

2. Consider a contract τ and suppose that at some outcome x_i such that $x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] > 0$ there is a transfer r such that $0 < r < A := \min\left\{CE^a\left[\mathbf{t}_{a_K}\right], x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right\}$ made with positive probability. Then the principal can replace r with a binary lottery with payments 0 and A such that the expected transfer, conditional on the realization of outcome x_i, is unchanged: the payment A is then offered with probability $\frac{r}{A}$. Keep all other payments from the original contract τ unchanged, and denote the new compensation (that follows from the new contract) in case of effort a by $\hat{\mathbf{t}}_a$.

For every action $a \in \{a_1, \ldots, a_K\}$ the principal's expected transfer is unchanged, $\mathbb{E}[\hat{\mathbf{t}}_a] = \mathbb{E}[\mathbf{t}_a]$, and the same holds for the principal's disappointment term

$$\mathbb{E}\left[(CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] - (x-\mathbf{t})_{a_K})^+\right].$$

Therefore $CE^p\left[\left(\mathbf{x}-\hat{\mathbf{t}}\right)_{a_K}\right] = CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]$. For the targeted action a_K, the agent's disappointment term $\mathbb{E}[(CE^a[\mathbf{t}_{a_K}]-\mathbf{t}_{a_K})^+]$ is also unchanged, so that $CE^a[\hat{\mathbf{t}}_{a_K}] = CE^a[\mathbf{t}_{a_K}]$.

Fix now any action $a \in \{a_1, \ldots, a_{K-1}\}$. Since the original contract τ implements a_K and since costs are strictly increasing, it holds that

$$CE^a[\mathbf{t}_a] - c(a) \leq CE^a[\mathbf{t}_{a_K}] - c(a_K),$$

and hence that

$$CE^a[\mathbf{t}_a] \leq CE^a[\mathbf{t}_{a_K}] - [c(a_K) - c(a)] < CE^a[\mathbf{t}_{a_K}].$$

If $r \geq CE^a[\mathbf{t}_a]$, then

$$\mathbb{E}[(CE^a[\mathbf{t}_a] - \mathbf{t}_a)^+ \mid \mathbf{t}_a = r] = 0$$

and therefore

$$\mathbb{E}[(CE^a[\mathbf{t}_a] - \hat{\mathbf{t}}_a)^+ \mid \mathbf{t}_a = r] - \mathbb{E}[(CE^a[\mathbf{t}_a] - \mathbf{t}_a)^+ \mid \mathbf{t}_a = r] > 0\,.$$

If instead $r < CE^a[\mathbf{t}_a]$ and $A > CE^a[\mathbf{t}_a]$, then

$$\begin{aligned}
&\mathbb{E}[(CE^a[\mathbf{t}_a] - \hat{\mathbf{t}}_a)^+] - \mathbb{E}[(CE^a[\mathbf{t}_a] - \mathbf{t}_a)^+] \\
&= p_a(x_i)\mathbb{P}[\tau_i = r] \cdot \left(\left(1 - \frac{r}{A}\right) CE^a[\mathbf{t}_a] - (CE^a[\mathbf{t}_a] - r)\right) \\
&= p_a(x_i)\mathbb{P}[\tau_i = r] \cdot r \left(1 - \frac{CE^a[\mathbf{t}_a]}{A}\right) > 0\,.
\end{aligned}$$

If $r < CE^a[\mathbf{t}_a]$ and $A < CE^a[\mathbf{t}_a]$, then

$$\mathbb{E}[(CE^a[\mathbf{t}_a] - \hat{\mathbf{t}}_a)^+] - \mathbb{E}[(CE^a[\mathbf{t}_a] - \mathbf{t}_a)^+] = 0$$

Thus, for every action $a \in \{a_1, \ldots, a_{K-1}\}$ we have

$$\mathbb{E}[(CE^a[\mathbf{t}_a] - \tilde{\mathbf{t}}_a)^+] \geq \mathbb{E}[(CE^a[\mathbf{t}_a] - \mathbf{t}_a)^+]\,.$$

Suppose we had $CE^a[\hat{\mathbf{t}}_a] > CE^a[\mathbf{t}_a]$. Then

$$\begin{aligned}
CE^a[\hat{\mathbf{t}}_a] &= \mathbb{E}[\hat{\mathbf{t}}_a] - \beta^a\, \mathbb{E}[(CE^a[\hat{\mathbf{t}}_a] - \hat{\mathbf{t}}_a)^+] \\
&< \mathbb{E}[\hat{\mathbf{t}}_a] - \beta^a\, \mathbb{E}[(CE^a[\mathbf{t}_a] - \hat{\mathbf{t}}_a)^+] \\
&\leq \mathbb{E}[\mathbf{t}_a] - \beta^{\,a} \mathbb{E}[(CE^a[\mathbf{t}_a] - \mathbf{t}_a)^+] = CE^a[\mathbf{t}_a]\,,
\end{aligned}$$

which is a contradiction. Hence $CE^a[\hat{\mathbf{t}}_a] = CE^a[\mathbf{t}_a]$ for every lower action $a < a_K$, the new contract still implements a_K, the principal's expected utility under a_K is unchanged, and every IC constraint still holds. If $CE^a[\hat{\mathbf{t}}_a] < CE^a[\mathbf{t}_a]$ for every lower action $a < a_K$ there exists $\lambda < 1$ such that replacing the contract $\hat{\tau}$ with contract $\lambda\hat{\tau}$ still implements a_K and increases the principal's expected utility. □

Lemma 10. *If there exists an outcome x_i where $CE^a\left[\mathbf{t}_{a_K}\right] > x_i - CE^p\left[(\mathbf{x} - \mathbf{t})_{a_K}\right] > 0$ and*

$$\mathbb{P}\left[\tau_i \in \left(x_i - CE^p\left[(\mathbf{x} - \mathbf{t})_{a_K}\right], CE^a\left[\mathbf{t}_{a_K}\right]\right)\right] > 0,$$

then there exists another contract $\tilde{\tau}$ that implements a_K in which

$$\mathbb{P}\left[\tilde{\tau}_i \in \left(x_i - CE^p\left[\left(\mathbf{x} - \tilde{\mathbf{t}}\right)_{a_K}\right], CE^a\left[\tilde{\mathbf{t}}_{a_K}\right]\right)\right] = 0 \text{ and}$$
$$\tilde{\tau}_j = \tau_j \quad \text{for any } j \neq i.$$

Moreover, $CE^p\left[\left(\mathbf{x} - \tilde{\mathbf{t}}\right)_{a_K}\right] \geq CE^p\left[(\mathbf{x} - \mathbf{t})_{a_K}\right]$.

Proof This is similar to the proof of Lemma 3, Theorem 1. We add the proof here for completeness, with a slight adaptation.

Consider a contract τ and suppose that at some outcome x_i where $CE^a\left[\mathbf{t}_{a_K}\right] > x_i - CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right] > 0$ there is a transfer r such that $x_i - CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right] < r < CE^a\left[\mathbf{t}_{a_K}\right]$ that is made with positive probability. Then the principal can replace r with a binary lottery with payments $x_i - CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]$ and $CE^a\left[\mathbf{t}_{a_K}\right]$ such that the expected transfer, conditional on the realization of outcome x_i, is unchanged: the payment $CE^a\left[\mathbf{t}_{a_K}\right]$ is offered with probability $\frac{r-\left(x_i-CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]\right)}{CE^a\left[\mathbf{t}_{a_K}\right]-\left(x_i-CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]\right)} \in (0,1)$. Keep all other payments from the original contract τ unchanged, and denote the new compensation (that follows from the new contract $\tilde{\tau}$) in case of effort a by $\tilde{\mathbf{t}}_a$.

The principal's expected share of the output remains the same: $\mathbb{E}\left[\left(\mathbf{x}-\tilde{\mathbf{t}}\right)_{a_K}\right] = \mathbb{E}\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]$. Moreover, since all the changes are in the disappointment part of the principal's payment, the principal's utility remaines unchanged.

For the targeted action a_K, the agent's disappointment term $\mathbb{E}[(CE^a[\mathbf{t}_{a_K}] - \mathbf{t}_{a_K})^+]$ is also unchanged, so that $CE^a[\tilde{\mathbf{t}}_{a_K}] = CE^a[\mathbf{t}_{a_K}]$.

Fix now any action $a \in \{a_1, \ldots, a_{K-1}\}$. Since the original contract τ implements a_K and since costs are strictly increasing, it holds that

$$CE^a[\mathbf{t}_a] - c(a) \leq CE^a[\mathbf{t}_{a_K}] - c(a_K),$$

and hence that

$$CE^a[\mathbf{t}_a] \leq CE^a[\mathbf{t}_{a_K}] - [c(a_K) - c(a)] < CE^a[\mathbf{t}_{a_K}]\,.$$

If $r > x_i - CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right] \geq CE^a[\mathbf{t}_a]$, then $CE^a[\mathbf{t}_a] = CE^a[\tilde{\mathbf{t}}_a]$ as all the changes are in the elation part of the agent's payments.

If $r \geq CE^a[\mathbf{t}_a] > x_i - CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]$, then

$$\mathbb{E}[(CE^a[\mathbf{t}_a] - \mathbf{t}_a)^+ \mid \mathbf{t}_a = r] = 0$$

and therefore

$$\mathbb{E}[(CE^a[\mathbf{t}_a] - \tilde{\mathbf{t}}_a)^+ \mid \mathbf{t}_a = r] - \mathbb{E}[(CE^a[\mathbf{t}_a] - \mathbf{t}_a)^+ \mid \mathbf{t}_a = r] > 0\,.$$

If $CE^a[\mathbf{t}_a] > r > x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]$, then

$$
\begin{aligned}
&\mathbb{E}[(CE^a[\mathbf{t}_a] - \tilde{\mathbf{t}}_a)^+] - \mathbb{E}[(CE^a[\mathbf{t}_a] - \mathbf{t}_a)^+] \\
&= p_a(x_i)\mathbb{P}[\tau_i = r]\times \\
&\left(\left(1 - \frac{r - \left(x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right)}{CE^a\left[\mathbf{t}_{a_K}\right] - \left(x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right)}\right)\left(CE^a[\mathbf{t}_a] - \left(x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right)\right) - \left(CE^a[\mathbf{t}_a] - r\right)\right) \\
&= p_a(x_i)\mathbb{P}[\tau_i = r]\times \\
&[CE^a[\mathbf{t}_a] - \left(x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right) \\
&- \left(r - \left(x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right)\right)\frac{CE^a[\mathbf{t}_a] - \left(x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right)}{CE^a\left[\mathbf{t}_{a_K}\right] - \left(x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right)} - \left(CE^a[\mathbf{t}_a] - r\right)] \\
&> p_a(x_i)\mathbb{P}[\tau_i = r]\times \\
&\left(CE^a[\mathbf{t}_a] - \left(x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right) - \left(r - \left(x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right)\right) - \left(CE^a[\mathbf{t}_a] - r\right)\right) \\
&= 0,
\end{aligned}
$$

where the inequality follows as $CE^a\left[\mathbf{t}_{a_K}\right] > CE^a[\mathbf{t}_a]$ and $CE^a[\mathbf{t}_a] - \left(x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right) > 0$.

Thus, for every action $a \in \{a_1, \ldots, a_{K-1}\}$ we have

$$\mathbb{E}[(CE^a[\mathbf{t}_a] - \tilde{\mathbf{t}}_a)^+] \geq \mathbb{E}[(CE^a[\mathbf{t}_a] - \mathbf{t}_a)^+]\,.$$

The rest is similar to the last part of the proof of Lemma 9.2. □

Lemma 11. *If there exists an outcome x_i such that $x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] > CE^a\left[\mathbf{t}_{a_K}\right]$, and*

$$\mathbb{P}\left[\tau_i \in \left(CE^a\left[\mathbf{t}_{a_K}\right],\ x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right)\right] > 0,$$

then there exists another contract $\tilde{\tau}$ that implements a_K in which

$$\mathbb{P}\left[\tilde{\tau}_i \in \left(CE^a\left[\tilde{\mathbf{t}}_{a_K}\right],\ x_i - CE^p\left[\left(\mathbf{x}-\tilde{\mathbf{t}}\right)_{a_K}\right]\right)\right] = 0 \text{ and}$$
$$\tilde{\tau}_j = \tau_j \quad \text{for any } j \neq i..$$

Moreover, $CE^p\left[\left(\mathbf{x}-\tilde{\mathbf{t}}\right)_{a_K}\right] = CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]$.

Proof: Consider a contract τ and suppose that at some outcome x_i such that $x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] > CE^a\left[\mathbf{t}_{a_K}\right]$ there is a transfer $CE^a\left[\mathbf{t}_{a_K}\right] < r < x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]$ made with positive probability. Then the principal can replace r by a binary lottery with payments $CE^a\left[\mathbf{t}_{a_K}\right]$ and $x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]$ such that the expected transfer, conditional

on the realization of outcome x_i, is unchanged: the payment is $CE^a\left[\mathbf{t}_{a_K}\right]$ with probability

$$\frac{r - x_i + CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]}{CE^a\left[\mathbf{t}_{a_K}\right] + CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] - x_i}$$

and the payment is $x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]$ with probability

$$\frac{CE^a\left[\mathbf{t}_{a_K}\right] - r}{CE^a\left[\mathbf{t}_{a_K}\right] + CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] - x_i}.$$

Keep all other payments from the original contract τ unchanged, and denote by $\tilde{\mathbf{t}}_a$ the new compensation that follows from the new contract in case of effort a .

For every action $a \in \{a_1, \ldots, a_K\}$ the principal's expected transfer is unchanged and the same holds for the principal's disappointment term $\mathbb{E}\left[(CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] - (\mathbf{x}-\mathbf{t})_{a_K})^+\right]$. Therefore $CE^p\left[\left(\mathbf{x}-\tilde{\mathbf{t}}\right)_{a_K}\right] = CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]$. Since the original contract is incentive compatible, $CE^a\left[\mathbf{t}_{a_K}\right] \geq CE^a\left[\mathbf{t}_a\right]$ holds for any action $a \in \{a_1, \ldots, a_{K-1}\}$. For any action a, the agent's disappointment term $\mathbb{E}\left[(CE^a[\mathbf{t}_a] - \mathbf{t}_{a_K})^+\right]$ is also unchanged, so that $CE^a[\tilde{\mathbf{t}}_a] = CE^a[\mathbf{t}_a]$. Therefore, the new contract also implements a_K and $CE^p\left[\left(\mathbf{x}-\tilde{\mathbf{t}}\right)_{a_K}\right] = CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]$. □

Lemma 12. *If in contract τ, the highest outcome x_n is not the only outcome at which a transfer higher than*

$$\max\left\{CE^a[\mathbf{t}_{a_K}],\ x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right\}$$

is paid with positive probability, then there exists another contract $\hat{\tau}$ (with the induced new compensation $\hat{\mathbf{t}}_a$) that implements a_K and $CE^p\left[\left(\mathbf{x}-\hat{\mathbf{t}}\right)_{a_K}\right] > CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]$.

Proof: Suppose that at some outcome x_i with $i < n$ there is a transfer $r > \max\left\{CE^a[\mathbf{t}_{a_K}],\ x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right\}$ made with positive probability. Then the principal can define a new contract $\widetilde{\tau}$ where she replaces r by $\max\left\{CE^a[\mathbf{t}_{a_K}], x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right\}$ at $\widetilde{\tau}_i$ and simultaneously increases τ_n by the amount

$$\frac{p_{a_K}(x_i)}{p_{a_K}(x_n)}\mathbb{P}[\tau_i = r]\delta,$$

where $\delta = r - \max\left\{CE^a[\mathbf{t}_{a_K}], x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right\} > 0$. By construction, this modification leaves the expected transfer under the target action unchanged, i.e. $\mathbb{E}[\tilde{\mathbf{t}}_{a_K}] = \mathbb{E}[\mathbf{t}_{a_K}]$, and hence it keeps the principal's expected payoff unchanged $\mathbb{E}[\left(\mathbf{x}-\tilde{\mathbf{t}}\right)_{a_K}] = \mathbb{E}[(\mathbf{x}-\mathbf{t})_{a_K}]$. The principal's disappointment term $\mathbb{E}\left[(CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] - (\mathbf{x}-\mathbf{t})_{a_K})^+\right]$ either decreases (if $\mathbb{P}\left[\tau_n < x_n - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right] > 0$) or remains the same (if $\mathbb{P}[\tau_n <$

$x_n - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]] = 0$). Moreover, the value to the agent when he takes the target action cannot decrease. Indeed, if $CE^a[\mathbf{t}_{a_K}] > CE^a[\tilde{\mathbf{t}}_{a_K}]$, then

$$\begin{aligned} CE^a[\mathbf{t}_{a_K}] - CE^a[\tilde{\mathbf{t}}_{a_K}] &= \beta^a\left(\mathbb{E}\left[(CE^a[\tilde{\mathbf{t}}_{a_K}] - \tilde{\mathbf{t}}_{a_K})^+\right] - \mathbb{E}\left[(CE^a[\mathbf{t}_{a_K}] - \mathbf{t}_{a_K})^+\right]\right) \\ &\leq \beta^a\left(\mathbb{E}\left[(CE^a[\mathbf{t}_{a_K}] - \tilde{\mathbf{t}}_{a_K})^+\right] - \mathbb{E}\left[(CE^a[\mathbf{t}_{a_K}] - \mathbf{t}_{a_K})^+\right]\right) \\ &\leq 0 \end{aligned}$$

which yields a contradiction.

Now fix any lower action $a \in \{a_1, \ldots, a_{K-1}\}$, and let

$$\Delta := \mathbb{P}[\tau_i = r]\delta\,.$$

The change in expected transfer under action a is

$$\begin{aligned} \mathbb{E}_a[(\tilde{\mathbf{t}}_a - \mathbf{t}_a)] &= -p_a(x_i)\Delta + p_a(x_n)\frac{p_{a_K}(x_i)}{p_{a_K}(x_n)}\Delta \\ &= \Delta\, p_a(x_n)\left(\frac{p_{a_K}(x_i)}{p_{a_K}(x_n)} - \frac{p_a(x_i)}{p_a(x_n)}\right) < 0\,. \end{aligned}$$

where the last inequality follows from the strict MLRP ordering that requires

$$\frac{p_{a_K}(x_i)}{p_a(x_i)} < \frac{p_{a_K}(x_n)}{p_a(x_n)}\,.$$

Moreover, both transfers involved in the above perturbation are weakly above $CE^a[\mathbf{t}_{a_K}]$, while $CE^a[\mathbf{t}_a] < CE^a[\mathbf{t}_{a_K}]$. Hence the agent's disappointment term under action a is unchanged, and therefore $CE^a[\tilde{\mathbf{t}}_a] < CE^a[\mathbf{t}_a]$ for every lower action a.

Thus, the new contract still implements a_K, yields the same expected payment under a_K, and makes every IC constraint slack. Therefore, there exists $0 < \lambda < 1$ such that the contract $\hat{\tau} = \lambda\widetilde{\tau}$ implements action a_K and $CE^p\left[\left(\mathbf{x}-\hat{\mathbf{t}}\right)_{a_K}\right] > CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]$. □

The above Lemmas imply together that, with the exception of the highest outcome x_n, all payments are concentrated on the support with three elements $\{0, CE^a[t_{a_K}], x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\}$. We now restrict attention to contracts with such support and show that the randomization occurring in an optimal contract is relatively limited.

Definition 1. *Let x_{j^*} be the highest outcome with the property that*

$$CE^a\left[\mathbf{t}_{a_K}\right] > x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right],$$

and let x_{j_*} be the highest outcome with the property

$$x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] \leq 0.$$

The next Lemma shows that a payment zero is used only on a set of low outcomes. That is, there exists a single cutoff such that the compensation is zero only for outcomes below this cutoff. At the cutoff, the compensation is zero with positive probability.

Lemma 13. *In an optimal mechanism if for some $l \in \{1,...,n\}$ it holds that $\mathbb{P}[\tau_l = 0] > 0$, then for any $i < l$ it holds that $\mathbb{P}[\tau_i = 0] = 1$*

Proof: Assume this is not the case. That is, there exists s and i with $i < s$ such that $\mathbb{P}[\tau_s = 0] > 0$ and $\mathbb{P}[\tau_i = 0] < 1$. Consider the following perturbation: increase $\mathbb{P}[\tau_i = 0]$ by $\varepsilon > 0$ and decrease $\mathbb{P}[\tau_s = 0]$ by δ. A few cases below describe additional perturbations depending on s and i:

1. $i < s \leq j_*$. Then the perturbation decreases $\mathbb{P}[\tau_i = CE^a[\mathbf{t}_{a_K}]]$ by ε and increases $\mathbb{P}[\tau_s = CE^a[\mathbf{t}_{a_K}]]$ by δ, such that the expected utility under the target action remains the same. Observe that all the changes are in the disappointment term, so it doesn't effect the principal's utility. Furthermore, the proof of Lemma 6, Theorem 1 implies that the change releases the IC constraint. Hence, a contradiction that the original contract was optimal.

2. $i \leq j_* < s \leq j^*$. Then the perturbation decreases $\mathbb{P}[\tau_i = CE^a[\mathbf{t}_{a_K}]]$ by ε and increases $\mathbb{P}[\tau_s = CE^a[\mathbf{t}_{a_K}]]$ by δ. Define by $\widetilde{\tau}$ the new contract, and by $\tilde{\mathbf{t}}_a$ the new compensation scheme under action a. The perturbation is chosen such that the expected payment under the target action is unchanged:

$$\mathbb{E}\left[\tilde{\mathbf{t}}_{a_K} - \mathbf{t}_{a_K}\right] = p_{a_K}(x_i)(-\varepsilon)CE^a[\mathbf{t}_{a_K}] + p_{a_K}(x_s)\delta CE^a[\mathbf{t}_{a_K}] = 0$$
$$\delta = \frac{p_{a_K}(x_i)}{p_{a_K}(x_s)}\varepsilon.$$

Observe that since $j_* < s$ we have this change decreases the payment to the agent at the disappointment region of the principal's utility and increases the payment to the agent at the elation region of the principal's utility. Hence,

$$\mathbb{E}\left[(CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] - \left(\mathbf{x}-\tilde{\mathbf{t}}\right)_{a_K})^+\right] - \mathbb{E}\left[(CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] - (\mathbf{x}-\mathbf{t})_{a_K})^+\right]$$
$$= p_{a_K}(x_i)(-\varepsilon)CE^a[\mathbf{t}_{a_K}] + p_{a_K}(x_s)\delta\left(CE^a[\mathbf{t}_{a_K}] - x_s + (CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right) <$$
$$p_{a_K}(x_i)(-\varepsilon)CE^a[\mathbf{t}_{a_K}] + p_{a_K}(x_s)\delta CE^a[\mathbf{t}_{a_K}] = 0$$

We show now that this perturbation increases the principal's utility. Suppose we have

$CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] \geq CE^p\left[\left(\mathbf{x}-\tilde{\mathbf{t}}\right)_{a_K}\right]$. Then

$$\begin{aligned} CE^p\left[\left(\mathbf{x}-\tilde{\mathbf{t}}\right)_{a_K}\right] &= \mathbb{E}\left[\left(\mathbf{x}-\tilde{\mathbf{t}}\right)_{a_K}\right] - \beta^p\mathbb{E}\left[(CE^p\left[\left(\mathbf{x}-\tilde{\mathbf{t}}\right)_{a_K}\right] - \left(\mathbf{x}-\tilde{\mathbf{t}}\right)_{a_K})^+\right] \\ &\geq \mathbb{E}\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] - \beta^p\mathbb{E}\left[(CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] - \left(\mathbf{x}-\tilde{\mathbf{t}}\right)_{a_K})^+\right] \\ &> \mathbb{E}\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] - \beta^p\mathbb{E}\left[(CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] - (\mathbf{x}-\mathbf{t})_{a_K})^+\right] = CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right], \end{aligned}$$

which is a contradiction. Hence, $CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] < CE^p\left[\left(\mathbf{x}-\tilde{\mathbf{t}}\right)_{a_K}\right]$.

Observe that the perturbation is in the disappointment region of the agent's utility under the target action, and hence, $CE^a\left[\mathbf{t}_{a_K}\right] = CE^a\left[\tilde{\mathbf{t}}_{a_K}\right]$. By the same argument as in the proof of Lemma 6, Theorem 1, the perturbation decreases the utility from any lower action, and hence it releases the IC constraint. Hence $CE[\tilde{\mathbf{t}}_a] < CE[\mathbf{t}_a]$ for every lower action a, and thus $\widetilde{\tau}$ is an optimal contract where all IC constraints are slack, which is a contradiction to Lemma 8.

3. $i \leq j_* \leq j^* < s$. Then the perturbation decreases $\mathbb{P}\left[\tau_i = CE^a\left[\mathbf{t}_{a_K}\right]\right]$ by ε and increases $\mathbb{P}\left[\tau_s = CE^a\left[\mathbf{t}_{a_K}\right]\right]$ by δ. Define by $\widetilde{\tau}$ the new contract, and by $\tilde{\mathbf{t}}_a$ the new compensation scheme under action a. The perturbation is chosen such that the expected payment under the target action is unchanged:

$$\mathbb{E}\left[\tilde{\mathbf{t}}_{a_K} - \mathbf{t}_{a_K}\right] = p_{a_K}(x_i)(-\varepsilon)CE^a\left[\mathbf{t}_{a_K}\right] + p_{a_K}(x_s)\delta CE^a\left[\mathbf{t}_{a_K}\right] = 0$$
$$\delta = \frac{p_{a_K}(x_i)}{p_{a_K}(x_s)}\varepsilon$$

Observe that this change decreases the payment to the agent at the disappointment region of the principal's utility and increases the payment to the agent at the elation region of the principal's utility. Hence, similar to case 2 the change increases the principal's utility. Furthermore, similarly to case 2 the perturbation preserves the agent's utility from the target action, and due to the MLRP decreases the agent's utility from any lower action. Hence, the perturbation increases the principal's utility. Hence, a contradiction that the original contract was optimal.

4. $j_* < i \leq j^* < s$. If $\mathbb{P}\left[\tau_i = CE^a\left[\mathbf{t}_{a_K}\right]\right] > 0$, then the perturbation decreases $\mathbb{P}\left[\tau_i = CE^a\left[\mathbf{t}_{a_K}\right]\right]$ by ε and increases $\mathbb{P}\left[\tau_s = CE^a\left[\mathbf{t}_{a_K}\right]\right]$ by δ such that the expected transfer under the target action remains the same. That is,

$$\mathbb{E}\left[\tilde{\mathbf{t}}_{a_K} - \mathbf{t}_{a_K}\right] = p_{a_K}(x_i)(-\varepsilon)(CE^a\left[\mathbf{t}_{a_K}\right]) + p_{a_K}(x_s)\delta CE^a\left[\mathbf{t}_{a_K}\right] = 0$$

Similarly to the proof of Lemma 6, Theorem 1 the change releases the IC constraint. Observe that since the change keeps the expected transfer to the agent under the tar-

get action unchanged, and decreases the disappointment payment of the principal, the perturbation increases the principal's utility.

If $\mathbb{P}\left[\tau_i = CE^a\left[\mathbf{t}_{a_K}\right]\right] = 0$, then $\mathbb{P}\left[\tau_i = x_i - CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]\right] > 0$. Then choose the perturbation that decreases $\mathbb{P}\left[\tau_i = x_i - CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]\right]$ by ε and increases $\mathbb{P}\left[\tau_s = CE^a\left[\mathbf{t}_{a_K}\right]\right]$ by δ such that the expected transfer under the target action remains the same. That is,

$$\mathbb{E}\left[\tilde{\mathbf{t}}_{a_K} - \mathbf{t}_{a_K}\right] = p_{a_K}\left(x_i\right)\left(-\varepsilon\right)\left(x_i - CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]\right) + p_{a_K}\left(x_s\right)\delta CE^a\left[\mathbf{t}_{a_K}\right] = 0$$

Define by $\widetilde{\tau}$ the new contract, and by $\tilde{\mathbf{t}}_a$ the new compensation scheme under action a. Since all these perturbation are in the elation region of the principal, the principal's utility remains the same. Since all the perturbation are in the disappointment region of the agent under the target action, $CE^a\left[\mathbf{t}_{a_K}\right] = CE^a\left[\tilde{\mathbf{t}}_{a_K}\right]$.

For the same argument as in the proof of Lemma 6 , Theorem 1, this perturbation decreases the agent's certainty equivalence from any lower action. Hence $CE[\tilde{\mathbf{t}}_a] < CE[\mathbf{t}_a]$ for every lower action a, and thus $\widetilde{\tau}$ is an optimal contract where all IC constraints are slack, which is a contradiction to Lemma 8.

5. $j_* < i < s \leq j^*$. If $\mathbb{P}\left[\tau_i = CE^a\left[\mathbf{t}_{a_K}\right]\right] > 0$, then the perturbation decreases $\mathbb{P}\left[\tau_i = CE^a\left[\mathbf{t}_{a_K}\right]\right]$ by ε and increases $\mathbb{P}\left[\tau_s = x_s - CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]\right]$ by δ such that the expected transfer under the target action remains the same. Since all the perturbations are in the disappointment part of the agent's utility under the target action, we have $CE[\tilde{\mathbf{t}}_{a_K}] = CE[\mathbf{t}_{a_K}]$. For the same argument as in the proof of Lemma 6 , Theorem 1, this perturbation decreases the agent's certainty equivalence from any lower action. Observe that this change decreases the payment to the agent at the disappointment region of the principal's utility and increases the payment to the agent at the elation region of the principal's utility. Hence, similar to case 2 the change increases the principal's utility. Hence, a contradiction that the original contract was optimal.

If $\mathbb{P}\left[\tau_i = CE^a\left[\mathbf{t}_{a_K}\right]\right] = 0$, then $\mathbb{P}\left[\tau_i = x_i - CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]\right] > 0$. Then, the additional perturbation decreases $\mathbb{P}\left[\tau_i = x_i - CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]\right]$ by ε and increases $\mathbb{P}[\tau_s = x_s - CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]]$ by δ such that the expected transfer under the target action remains the same. Since all the perturbations are in the disappointment part of the agent's utility under the target action, we have $CE[\tilde{\mathbf{t}}_{a_K}] = CE[\mathbf{t}_{a_K}]$. For the same argument as in the proof of Lemma 6 , Theorem 1, this perturbation decreases the agent's certainty equivalence from any lower action. Since all these perturbation are in the elation region of the principal, the principal's utility remains the same. Hence, if the original contract was optimal, the new contract is optimal as well.

6. $j^* < i < s$. If $\mathbb{P}\left[\tau_i = CE^a\left[\mathbf{t}_{a_K}\right]\right] > 0$, then the perturbation decreases $\mathbb{P}\left[\tau_i = CE^a\left[\mathbf{t}_{a_K}\right]\right]$ by ε and increases $\mathbb{P}\left[\tau_s = CE^a\left[\mathbf{t}_{a_K}\right]\right]$ by δ such that the expected transfer under the tar-

get action remains the same. Since all these perturbation are in the elation region of the principal, the proof of Lemma 6, Theorem 1 implies that the perturbation releases the IC constraints, hence this case cannot be part of the optimal contract.

If $\mathbb{P}\left[\tau_i = CE^a\left[\mathbf{t}_{a_K}\right]\right] = 0$, then $\mathbb{P}\left[\tau_i = x_i - CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]\right] > 0$, and in case of realization of output x_i the payment is a binary lottery between 0 and $x_i - CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right] > CE^a\left[\mathbf{t}_{a_K}\right]$. Replace this binary lottery with its utility-preserving (for the agent under the target action) contraction. That is, if $\mathbb{P}\left[\tau_i = 0\right] = \alpha > 0$ and $\mathbb{P}\left[\tau_i = x_i - CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]\right] = 1-\alpha$, perturb the payments as follows: $\mathbb{P}\left[\tilde{\tau}_i = CE^a\left[\mathbf{t}_{a_K}\right]\right] = \gamma > 0$, $\mathbb{P}\left[\tilde{\tau}_i = x_i - CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]\right] = \delta < 1-\alpha$ and $\mathbb{P}\left[\tilde{\tau}_i = 0\right] = 1-\gamma-\delta < \alpha$ such that

$$\begin{aligned} & (1-\alpha)\left(x_i - CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]\right) - \beta^a \alpha CE^a\left[\mathbf{t}_{a_K}\right] \\ = \;& \gamma CE^a\left[\mathbf{t}_{a_K}\right] + \delta\left(x_i - CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]\right) - \beta^a\left(1-\gamma-\delta\right)CE^a\left[\mathbf{t}_{a_K}\right], \end{aligned}$$

which we can rewrite as

$$(1-\alpha-\delta)\left(x_i - CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]\right) = \left(\beta^a\alpha + \gamma - \beta^a\left(1-\gamma-\delta\right)\right)CE^a\left[\mathbf{t}_{a_K}\right].$$

As the agent is risk averse – and one payment is in the elation interval, while another is at the disappointment interval – the perturbation decreases the expected payment to the agent under the target action. Since all the perturbed payments belong to the principal's elation interval, this perturbation increases the principal's utility. Fix any lower action $a \in \{a_1, ..., a_{K-1}\}$. We show now that

$$\mathbb{E}[\mathbf{t}_a] - \beta^a\mathbb{E}[(CE^a[\mathbf{t}_a] - \mathbf{t}_a)^+] > \mathbb{E}[\tilde{\mathbf{t}}_a] - \beta^a\mathbb{E}[(CE^a[\mathbf{t}_a] - \tilde{\mathbf{t}}_a)^+]$$

For the chosen perturbation we have

$$\begin{aligned} & \mathbb{E}[\mathbf{t}_a] - \beta^a\mathbb{E}[(CE^a[\mathbf{t}_a] - \mathbf{t}_a)^+] - \mathbb{E}[\tilde{\mathbf{t}}_a] + \beta^a\mathbb{E}[(CE^a[\mathbf{t}_a] - \tilde{\mathbf{t}}_a)^+] \\ = \;& p_a\left(x_i\right)\{(1-\alpha)\left(x_i - CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]\right) - \beta^a\alpha CE^a\left[\mathbf{t}_a\right] - \gamma CE^a\left[\mathbf{t}_{a_K}\right] \\ & -\delta\left(x_i - CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]\right) + \beta^a\left(1-\gamma-\delta\right)CE^a\left[\mathbf{t}_a\right]\} \\ = \;& p_a\left(x_i\right)\{(1-\alpha-\delta)\left(x_i - CE^p\left[\left(\mathbf{x}-\mathbf{t}\right)_{a_K}\right]\right) - \gamma CE^a\left[\mathbf{t}_{a_K}\right] \\ & -\beta^a\left(\alpha - \left(1-\gamma-\delta\right)\right)CE^a\left[\mathbf{t}_a\right]\} \\ = \;& p_a\left(x_i\right)\{\left(\beta^a\alpha + \gamma - \beta^a\left(1-\gamma-\delta\right)\right)CE^a\left[\mathbf{t}_{a_K}\right] \\ & -\gamma CE^a\left[\mathbf{t}_{a_K}\right] - \beta^a\left(\alpha - \left(1-\gamma-\delta\right)\right)CE^a\left[\mathbf{t}_a\right]\} > 0 \end{aligned}$$

where the last inequality follows since As $CE^a\left[\mathbf{t}_{a_K}\right] > CE^a\left[\mathbf{t}_a\right]$ for any $a < a_K$.

Suppose that $CE^a[\tilde{\mathbf{t}}_a] \geq CE^a[\mathbf{t}_a]$. Then

$$\begin{aligned} CE^a[\tilde{\mathbf{t}}_a] &= \mathbb{E}[\tilde{\mathbf{t}}_a] - \beta^a \mathbb{E}[(CE^a[\tilde{\mathbf{t}}_a] - \tilde{\mathbf{t}}_a)^+] \\ &\leq \mathbb{E}[\tilde{\mathbf{t}}_a] - \beta^a \mathbb{E}[(CE^a[\mathbf{t}_a] - \tilde{\mathbf{t}}_a)^+] \\ &< \mathbb{E}[\mathbf{t}_a] - \beta^a \mathbb{E}[(CE^a[\mathbf{t}_a] - \mathbf{t}_a)^+] = CE^a[\mathbf{t}_a]\,, \end{aligned}$$

which is a contradiction. Hence the perturbation increases the principal's utility and preserves the incentives. Contradiction that the original scheme was part of the optimal contract. □

The next Lemma shows that if a payment $x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]$ is made after an outcome $x_i > x_{j^*+1}$, then there exists an interval of outcomes with upper bound x_{n-1} where this payment is made with positive probability at the interval's lower bound and with probability 1 after the other outcomes in this interval.

Lemma 14. *In an optimal contract, if*

$$\mathbb{P}\left[\tau_i = x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right] > 0,$$

holds for outcome $x_i > x_{j^*+1}$, then for all $l \in \{i+1, ..., n-1\}$ holds

$$\mathbb{P}\left[\tau_l = x_l - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right] = 1.$$

Proof: If there exist $i > j^* + 1$ and $l \in \{i+1, ..., n-1\}$ such that

$$\mathbb{P}\left[\tau_i = x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right] > 0,$$

and

$$\mathbb{P}\left[\tau_l = x_l - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right] < 1,$$

consider the following modification of the transfers: decrease $\mathbb{P}\left[\tau_i = x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right]$ by ε , increase $\mathbb{P}\left[\tau_i = CE^a\left[\mathbf{t}_{a_K}\right]\right]$ by ε, increase $\mathbb{P}\left[\tau_l = x_l - CE^p\left[(x-t)_{a_K}\right]\right]$ by δ , and decrease $\mathbb{P}\left[\tau_l = CE^a\left[\mathbf{t}_{a_K}\right]\right]$ by δ $\left(\mathbb{P}\left[\tau_l = CE^a\left[\mathbf{t}_{a_K}\right]\right]\right) > 0$ because $l > i$ and $\mathbb{P}\left[\tau_i = 0\right] < 1$, and hence $\mathbb{P}\left[\tau_l = 0\right] = 0$) such that the expected transfer to the agent under the target action remains the same:

$$p_{a_K}(x_i)\,\varepsilon\left(-x_i + CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] + CE^a\left[\mathbf{t}_{a_K}\right]\right) + p_{a_K}(x_l)\,\delta\left(x_l - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] - CE^a\left[\mathbf{t}_{a_K}\right]\right) = 0$$

$$\delta\left(x_l - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] - CE^a\left[\mathbf{t}_{a_K}\right]\right) = -\frac{p_{a_K}(x_i)}{p_{a_K}(x_l)}\varepsilon\left(-x_i + CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] + CE^a\left[\mathbf{t}_{a_K}\right]\right).$$

Since all these changes are in the principal's elation region, the perturbation doesn't

change the principal's utility. Moreover, since for any alternative action $a \in \{a_1, ..., a_{K-1}\}$ the transfer changes are in the agent's elation region, the change in the agent's utility under an alternative action a is

$$\begin{aligned}
& p_a(x_i)\,\varepsilon\left(-x_i + CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] + CE^a\left[\mathbf{t}_{a_K}\right]\right) + p_a(x_l)\,\delta\left(x_l - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] - CE^a\left[\mathbf{t}_{a_K}\right]\right) \\
= \; & p_a(x_i)\,\varepsilon\left(-x_i + CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] + CE^a\left[\mathbf{t}_{a_K}\right]\right) \\
& -p_a(x_l)\,\frac{p_{a_K}(x_i)}{p_{a_K}(x_l)}\varepsilon\left(-x_i + CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] + CE^a\left[\mathbf{t}_{a_K}\right]\right) \\
= \; & \varepsilon\left(-x_i + CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] + CE^a\left[\mathbf{t}_{a_K}\right]\right) p_a(x_l)\left(\frac{p_a(x_i)}{p_a(x_l)} - \frac{p_{a_K}(x_i)}{p_{a_K}(x_l)}\right) < 0,
\end{aligned}$$

where the inequality follows by MLRP, and because $x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] > CE^a\left[\mathbf{t}_{a_K}\right]$ holds for $i > j^*$. Therefore, the perturbation turns every IC constraint into slack, contradicting Lemma 8. □

The Next Lemma shows that the transfer $CE^a\left[\mathbf{t}_{a_K}\right]$ can be paid only in the middle interval, with potential randomization only at its boundary.

Lemma 15. *In an optimal contract, if for $i < j^*$ holds $\mathbb{P}\left[\tau_i = CE^a\left[\mathbf{t}_{a_K}\right]\right] > 0$, then for all $s \in \{i+1, ..., j^*\}$ holds $\mathbb{P}\left[\tau_s = CE^a\left[\mathbf{t}_{a_K}\right]\right] = 1$.*

Proof: Suppose this is not the case, that is, there exists an optimal contract in which $i < s \leq j^*$ with $\mathbb{P}\left[\tau_i = CE^a\left[\mathbf{t}_{a_K}\right]\right] > 0$ and $\mathbb{P}\left[\tau_s = CE^a\left[\mathbf{t}_{a_K}\right]\right] < 1$. Lemma 13 implies that $\mathbb{P}\left[\tau_s = 0\right] = 0$, and hence $\mathbb{P}\left[\tau_s = x_s - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right] > 0$. Consider the following perturbation: Decrease $\mathbb{P}\left[\tau_i = CE^a\left[\mathbf{t}_{a_K}\right]\right]$ by ε, increase $\mathbb{P}\left[\tau_i = x_s - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right]$ (from zero) by ε, and increase $\mathbb{P}\left[\tau_s = CE^a\left[\mathbf{t}_{a_K}\right]\right]$ by δ and decrease $\mathbb{P}\left[\tau_s = x_s - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right]$ by δ, such that the expected transfer to the agent under the target action remains unchanged:

$$\begin{aligned}
\mathbb{E}\left[\tilde{\mathbf{t}}_{a_K} - \mathbf{t}_{a_K}\right] = & \; p_{a_K}(x_i)\,\varepsilon\left(-CE^a\left[\mathbf{t}_{a_K}\right] + x_s - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right) \\
& + p_{a_K}(x_s)\,\delta\left(CE^a\left[\mathbf{t}_{a_K}\right] - x_s + CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right) = 0
\end{aligned}$$

That is,

$$p_{a_K}(x_i)\,\varepsilon = p_{a_K}(x_s)\,\delta$$

Observe that all the modifications are in the disappointment part of the principal's utility, and in the disappointment part of the agent's utility under the target action.

Fix any lower action $a \in \{a_1, ..., a_{K-1}\}$.

$$\begin{aligned}
\mathbb{E}\left[\tilde{\mathbf{t}}_a - \mathbf{t}_a\right] &= p_a(x_i)\,\varepsilon\left(-CE^a\left[\mathbf{t}_{a_K}\right] + x_s - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right) \\
&\quad + p_a(x_s)\,\delta\left(CE^a\left[\mathbf{t}_{a_K}\right] - x_s + CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right) \\
&= \left(CE^a\left[\mathbf{t}_{a_K}\right] - x_s + CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right)\left(p_a(x_s)\,\delta - p_a(x_i)\,\varepsilon\right) \\
&= \left(CE^a\left[\mathbf{t}_{a_K}\right] - x_s + CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right)\varepsilon p_a(x_s)\left(\frac{p_{a_K}(x_i)}{p_{a_K}(x_s)} - \frac{p_a(x_i)}{p_a(x_s)}\right) < 0
\end{aligned}$$

where the inequality follows from the strict MLRP and $CE^a\left[\mathbf{t}_{a_K}\right] \geq x_s - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]$ (with the last equality only for $s = j^*$). If $CE^a\left[\mathbf{t}_a\right] \leq x_s - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]$ then the perturbation doesn't change the disappointment term of the agent's utility in case of alternative action. If $CE^a\left[\mathbf{t}_a\right] > x_s - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]$, then

$$\begin{aligned}
&\mathbb{E}[(CE[\mathbf{t}_a] - \tilde{\mathbf{t}}_a)^+] - \mathbb{E}[(CE[\mathbf{t}_a] - \mathbf{t}_a)^+] \\
&= p_a(x_i)\,\varepsilon\left(CE^a\left[\mathbf{t}_a\right] - x_s + CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right) - p_a(x_s)\,\delta\left(CE^a\left[\mathbf{t}_a\right] - x_s + CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right) \\
&= \left(CE^a\left[\mathbf{t}_a\right] - x_s + CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right)\left(p_a(x_s)\,\delta - p_a(x_i)\,\varepsilon\right) \\
&= \left(CE^a\left[\mathbf{t}_a\right] - x_s + CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\right)\varepsilon p_a(x_s)\left(\frac{p_a(x_i)}{p_a(x_s)} - \frac{p_{a_K}(x_i)}{p_{a_K}(x_s)}\right) > 0
\end{aligned}$$

Suppose that $CE[\tilde{\mathbf{t}}_a] \geq CE[\mathbf{t}_a]$. Then

$$\begin{aligned}
CE[\tilde{\mathbf{t}}_a] &= \mathbb{E}[\tilde{\mathbf{t}}_a] - \beta\mathbb{E}[(CE[\tilde{\mathbf{t}}_a] - \tilde{\mathbf{t}}_a)^+] \\
&\leq \mathbb{E}[\tilde{\mathbf{t}}_a] - \beta\mathbb{E}[(CE[\mathbf{t}_a] - \tilde{\mathbf{t}}_a)^+] \\
&< \mathbb{E}[\mathbf{t}_a] - \beta\mathbb{E}[(CE[\mathbf{t}_a] - \mathbf{t}_a)^+] = CE[\mathbf{t}_a],
\end{aligned}$$

which is a contradiction. Hence $CE[\tilde{\mathbf{t}}_a] < CE[\mathbf{t}_a]$ for every lower action a, and thus $\widetilde{\tau}$ is an optimal contract where all IC constraints are slack, which is a contradiction to Lemma 8. □

Finally, our last Lemma shows the existence of an optimal contract in the class of contracts where all payments are concentrated on the set $\{0, CE^a[t_{a_K}], x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\}$.

Lemma 16. *Let $\mathbf{x}_{a_K}$ represent the random outcome conditional on the agent taking action a_K. There exists an optimal contract for the problem where $\bar{t} = \infty$ that is feasible for every $\bar{t} \geq \max\left\{\bar{t}_0, \mathbb{E}\left[\mathbf{x}_{a_K}\right]\right\}$. Consequently, the bounded and unbounded problems have the same value for every $\bar{t} \geq \max\left\{\bar{t}_0, \mathbb{E}\left[\mathbf{x}_{a_K}\right]\right\}$.*

Proof: The binary contract from Lemma 1 that concentrates all compensation on x_n incentivises the agent to exert effort a_K. Denote by $CE^p_{bin} < \infty$ the principal's utility from this contract. Since any contract that generates a utility to the principal that is

lower than CE^p_{bin} is suboptimal, we can, without loss of generality, restrict attention to contracts satisfying

$$CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] \geq CE^p_{bin}$$

This implies that

$$\mathbb{E}\left[\mathbf{x}_{a_K}\right] - \mathbb{E}\left[\mathbf{t}_{a_K}\right] - \beta^p \mathbb{E}\left[\left(CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] - (\mathbf{x}-\mathbf{t})_{a_K}\right)^+\right] \geq CE^p_{bin}$$

$$\mathbb{E}\left[\mathbf{t}_{a_K}\right] \leq \mathbb{E}\left[\mathbf{x}_{a_K}\right] - CE^p_{bin} - \beta^p \mathbb{E}\left[\left(CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right] - (\mathbf{x}-\mathbf{t})_{a_K}\right)^+\right] \leq \mathbb{E}\left[\mathbf{x}_{a_K}\right]$$

By the Lemmas in this Appendix we can restrict attention to contracts such that payments below x_n belong to the set $\{0, CE^a\left[\mathbf{t}_{a_K}\right], x_i - CE^p\left[(\mathbf{x}-\mathbf{t})_{a_K}\right]\}$. This implies that $\mathbb{E}\left[\tau_n\right] \leq \mathbb{E}\left[\mathbf{x}_{a_K}\right]$. Replace the possibly random payment at x_n by the deterministic payment $\mathbb{E}\left[\tau_n\right]$. The rest of the proof is similar to Lemma 7.

10 Appendix D. The Optimal Contract with Yaari Dual utility

In this Appendix we illustrate that if the agent's preferences are described by Yaari's dual utility function (or by Köszegi-Rabin [2006] loss aversion with linear utility over monetary outcomes), then the number of different payments in the optimal contract may be as large as the number of possible outcomes. The example below has four possible available actions and four possible outcomes. The case with n available actions and n possible outcomes, where the number of different payments in the optimal contract is n, is analogous and available upon request from the authors. Let $a_1 < a_2 < a_3 < a_4$. The outcome probabilities and effort costs are given by the following table:

Action	x_1	x_2	x_3	x_4	Cost c_i
a_1	0.600	0.250	0.100	0.050	517/1000
a_2	0.350	0.300	0.225	0.125	8097/8000
a_3	0.300	0.275	0.250	0.175	4721/4000
a_4	0.250	0.250	0.250	0.250	11/8

The distributions in the table satisfy the strict MLRP condition. The risk-neutral principal maximizes expected output minus expected compensation. The contract must satisfy a limited liability constraint, that is, all the payments need to be non-negative. The agent's probability distortion function is given by

$$g(p) = \frac{4}{5}p + \frac{1}{5}p^2,$$

that is strictly increasing and strictly convex, with $g(0) = 0$ and $g(1) = 1$.This corresponds to a Köszegi-Rabin loss averse utility function

$$\mathcal{U}_{KR}(\mathbf{t}) = \mathbb{E}(\mathbf{t}) + \int\int \mu\left(t - y\right) dF\left(t\right) dF\left(y\right)$$

where $\mathbf{y}$ is distributed as $\mathbf{t}$ and where

$$\mu(z) = \begin{cases} z & \text{if } z \geq 0 \\ \frac{6}{5}z & \text{if } z < 0 \end{cases}.$$

The agent has an outside option with zero utility and her utility from taking action a_i is given here by $CE_{Yaari}\left[\mathbf{t}_{a_i}\right] - c\left(a_i\right)$, where $CE_{Yaari}\left[\mathbf{t}_{a_i}\right]$ denotes the Yaari (dual) certainty equivalent of the transfer lottery. For nonnegative payments this is given by:

$$CE_{Yaari}\left[\mathbf{t}_{a_i}\right] = \int_0^\infty g\big(\mathbb{P}[\mathbf{t}_{a_i} \geq s]\big)\, ds.$$

Assume that the outcomes $\{x_1, ..., x_4\}$ are such that the principal seeks to implement the costliest action a_4. The optimal contract is given here by

$$t^*(x_1) = 0, \quad t^*(x_2) = 1, \quad t^*(x_3) = 2, \quad t^*(x_4) = 3.$$

Each of the three payment increments equals one. Therefore

$$CE_{Yaari}\left[\mathbf{t}^*_{a_i}\right]) = g\big(P_i\{x_2, x_3, x_4\}\big) + g\big(P_i\{x_3, x_4\}\big) + g\big(P_i\{x_4\}\big).$$

The relevant probabilities and the resulting Yaari utilities are given by:

Action	$P_i\{x_2, x_3, x_4\}$	$P_i\{x_3, x_4\}$	$P_i\{x_4\}$	$CE_{Yaari}\left[\mathbf{t}^*_{a_i}\right])$
a_1	0.400	0.150	0.050	4136/8000
a_2	0.650	0.350	0.125	8097/8000
a_3	0.700	0.425	0.175	9442/8000
a_4	0.750	0.500	0.250	11000/8000

Observe that in this contract

$$CE_{Yaari}\left[\mathbf{t}_{a_i}\right] - c\left(a_i\right) = 0 \qquad \text{for } i = 1, 2, 3, 4.$$

The participation constraint and all three incentive constraints bind. The expected payment of this contract is

$$\mathbb{E}[\mathbf{t}_{a_4}] = \frac{0 + 1 + 2 + 3}{4} = \frac{3}{2}.$$

The uniform distribution of outcomes resulting from the target action together with MLRP imply the monotonicity of the optimal contract. The suggested contract minimizes the risk premium associated with action a_4 paid by the principal, while still satisfying the IC and IR constraints. More specifically, there exist positive Lagrange multipliers of all the IC and the IR constraints and the contract described above is a global maximum.

References

[2007] Abdellaoui, M. & H. Bleichrodt (2007) Eliciting Gul's theory of disappointment aversion by the tradeoff method, *Journal of Economic Psychology* **28**, 631-645.

[2005] Ang, A., Bekaert, G., & Liu, J. (2005). Why stocks may disappoint. *Journal of Financial Economics*, **76**(3), 471-508.

[1988] Baker, G. P., Jensen, M. C., & Murphy, K. J. (1988). Compensation and incentives: Practice vs. theory. *The Journal of Finance*, **43**(3), 593-616.

[2011] Batt, R., & Colvin, A. J. S. (2011). An employment systems approach to turnover: Human resources practices, quits, dismissals, and performance. *Academy of Management Journal*, **54**(4), 695-717.

[2025] Bellini, F., Maccheroni, F., Mao, T., Wang, R., & Wu, Q. (2025). Disappointment aversion and expectiles. arXiv preprint arXiv:2508.05541.

[2005] Bolton, P. and Dewatripoint, M. (2005) *Contract Theory* , Cambridge MA: The MIT Press.

[1994] Camerer, C. F. & T.-H. Ho (1994) Violations of the betweenness axiom and nonlinearity in probability, *Journal of Risk and Uncertainty*, **8**, 167-196.

[2015] Carroll, Gabriel "Robustness and linear contracts (2015) ", *American Economic Review* **105** (2), 536-563

[2020] Cerreia-Vioglio, S., Dillenberger, D., & Ortoleva, P. (2020). An explicit representation for disappointment aversion and other betweenness preferences. *Theoretical Economics*, 15, 1509-1546.

[2025] Chaigneau, P., Edmans, A., & Gottlieb, D. (2025). A theory of fair CEO pay. *American Economic Review: Insights*, **7**(3), 306-324.

[2009] Chakraborty, A., Sheikh, S., & Subramanian, N. (2009). The relationship between incentive compensation and performance related CEO turnover. *Journal of Economics and Business*, **61**(4), 295-311.

[2006] 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.

[2017] Delikouras (2017). "Where's the kink? Disappointment events in consumption growth and equilibrium asset prices," *The Review of Financial Studies*, **30** (8), 2851-2889.

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

[2010] 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.

[2022] Dreyfuss, B., O. Heffetz, and M. Rabin. (2022). Expectations-based loss aversion may help explain seemingly dominated choices in strategy-proof mechanisms." *American Economic Journal: Microeconomics* **14** (4): 515–55.

[2017] Edmans, A., Gabaix, X., & Jenter, D. (2017). Executive compensation: A survey of theory and evidence. *The Handbook of the Economics of Corporate Governance* 1, 383-539.

[1990] Gerhart, B., & Milkovich, G. T. (1990). Organizational differences in managerial compensation and financial performance. *Academy of Management Journal*, **33**(4), 663-691.

[1991] Gul, F. (1991). A theory of disappointment aversion. *Econometrica* **59**(3), 667-686.

[1983] Grossman, S. J., & Hart, O. D. (1983). An analysis of the principal-agent problem. *Econometrica*, **51**(1).

[2021] Hassidim, A., Romm, A., and Shorrer, R. (2021) The limits of incentives in economic matching procedures. *Management Science* **67**(2), 951-963.

[2025] Ho, H. (2025). Loss aversion, moral hazard, and stochastic contracts. *Managerial and Decision Economics*, **46**(3), 1677-1685.

[1991] Hermalin, B., and M. Katz (1991) Moral Hazard and Verifiability: The Effects of Renegotiation in Agency. *Econometrica* **59**(6), 1735-1753.

[2010] Herweg, Fabian, Daniel Müller, and Philipp Weinschenk (2010). "Binary payment schemes: Moral hazard and loss aversion." *American Economic Review* **100(**5), 2451-2477.

[1979] Holmström, B. (1979). Moral hazard and observability. *The Bell Journal of Economics*, 74-91.

[1987] Holmstrom, B., & Milgrom, P. (1987). Aggregation and linearity in the provision of intertemporal incentives. *Econometrica:* Journal of the Econometric Society, 303-328.

[1990] Innes, R. (1990) Limited liability and incentive contracting with ex-ante action choices. *Journal of Economic Theory* **52**, 45-67.

[1990] Jensen, M. C., & Murphy, K. J. (1990). Performance pay and top-management incentives. *Journal of Political Economy*, **98**(2), 225-264.

[1988] Jewitt, I. (1988) Justifying the first-order approach to principal-agent problems. *Econometrica*: 1177-1190.

[2001] Jia, J., Dyer, J. S., & Butler, J. C. (2001). "Generalized disappointment models. *Journal of Risk and Uncertainty* **22**(1), 59-78.

[2022] Jung, J.Y. (2022). "Effects of changes in preferences in moral hazard problems," *Journal of Economic Theory* **205**, 105527.

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

[2014] Kőszegi, B. (2014) "Behavioral contract theory." *Journal of Economic Literature*, **52** (4), 1075-1118.

[1990] Leonard, J. S. (1990). Executive pay and firm performance. Industrial and labor relations Review, 43(3), 13S-29S.

[1980] Lewis, T. R. (1980). Bonuses and penalties in incentive contracting. *The Bell Journal of Economics*, 292-301.

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

[2016] Masatlioglu, Y., & Raymond, C. (2016). "A behavioral analysis of stochastic reference dependence". *American Economic Review* **106**(9), 2760-2782.

[1976] James A Mirrlees (1976) The Optimal structure of incentives and authority within an organization *The Bell Journal of Economics* **7(1),** 105 - 131.

[2012] Moldovanu, B., Sela, A., & Shi, X. (2012). Carrots and sticks: prizes and punishments in contests. *Economic Inquiry*, **50**(2), 453-462.

[2025] Mora, E. R., & Strack, P. (2025). Mechanism Design without Expected Utility Preferences, discussion paper Yale University

[1999] Murphy, K. J. (1999). Executive compensation. *Handbook of Labor Economics* vol. 3, 2485-2563.

[1998] Murphy, K. J. (2000). Performance standards in incentive contracts. *Journal of Accounting and Economics*, **30**(3), 245-278.

[1987] Newey, W. K., & Powell, J. L. (1987). Asymmetric least squares estimation and testing. : *Econometrica*, 819-847.

[2014] Peters, F. S., & Wagner, A. F. (2014). The executive turnover risk premium. The *Journal of Finance*, **69**(4), 1529-1563.

[1999] Prendergast, C. (1999). The provision of incentives in firms. *Journal of Economic Literature*, **37**(1), 7-63.

[2002] Prendergast, C. (2002). The tenuous trade-off between risk and incentives. *Journal of Political Economy*, **110**(5), 1071-1102.

[1985] Rogerson, W. P. (1985). The first-order approach to principal-agent problems. *Econometrica*: 1357-1367

[2010] Routledge, B., & S. Zin. 2010. Generalized disappointment aversion and asset prices. *Journal of Finance* **65**, 1303–1332.

[2005] Salanié, B. (2005). *The economics of contracts: a primer*. Cambridge, MA: MIT Press.

[1998] Safra, Z., & Segal, U. (1998). Constant risk aversion. *Journal of Economic Theory*, **83**(1), 19-42.

[1990] Segal, U., & Spivak, A. (1990). First order versus second order risk aversion. *Journal of Economic Theory*, **51**(1), 111-125.

[1979] Shavell, S. (1979). Risk sharing and incentives in the principal and agent relationship. *The Bell Journal of Economics*, 55-73.

[2018] United States Securities and Exchange Commission, Proxy Statement for Meeting of Stockholders of Tesla, Inc., March 21 2018., https://www.sec.gov

[1987] Yaari, M. E. (1987). The dual theory of choice under risk. *Econometrica* **55**(1), 95-115.