Optimal Delegation and Information Transmission under Limited Awareness*

Sarah Auster‡ and Nicola Pavoni§

January 2021

Abstract

We study the delegation problem between a principal and an agent, who not only has better information about the performance of the available actions but also superior awareness of the set of actions that are actually feasible. The agent decides which of the available actions to reveal and which ones to hide. We provide conditions under which the agent finds it optimal to leave the principal unaware of relevant options. By doing so, the agent increases the principal's cost of distorting the agent's choices and increases the principal's willingness to grant him higher information rents. We also consider communication between the principal and the agent after the contract is signed and the agent receives information. We show that limited awareness of actions improves communication in such signalling games: the principal makes a coarser inference from the recommendations of the privately informed agent and accepts a larger number of his proposals.

JEL Codes: D82, D83, D86

*The authors thank Sandeep Baliga, Pierpaolo Battigalli, Sylvain Chassang, Wouter Dessein, Antonio Guarino, Marina Halac, Johannes Hörner, Navin Kartik, Alessandro Pavan, and Yuval Salant for very helpful comments. The present draft benefited from several comments received by the seminar participants at the Toulouse School of Economics, Collegio Carlo Alberto, University College London, EIEF in Rome, Bonn University and at the Max Plank Institute, at the University of Bristol, the University of Cardiff, at the Bocconi and Cattolica universities in Milan, at the University of Verona, University of Malaga, Oxford University, Stockholm School of Economics and the Cowles Conference on Economic Theory. Nicola Pavoni acknowledges financial support from the MIUR-PRIN grant 20157NH5TP. Sarah Auster acknowledges funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy—EXC 2126/1—390838866 and CRC TR 224 (Project B02).

‡Department of Economics, University of Bonn: auster.sarah@gmail.com

§Department of Economics, Bocconi University: pavoni.nicola@gmail.com

1 Introduction

In many situations economic agents delegate decisions to experts whose preferences may not be perfectly aligned with their own. Headquarters rely on division managers who have superior information about the profitability of available projects but also a desire to attract additional resources to their own division, voters entrust decisions to politicians whose preferences may reflect a political bias or the interest of certain lobbies, financial investors seek advice from non-neutral financial professionals with a better understanding of the risks and returns of the available portfolios. The tension underlying these situations has been formalised in the delegation model—first introduced by Holmström (1977)—where an uninformed principal specifies a set of permissible actions to the informed agent and contingent transfers are infeasible.

In most of the described situations, the informed party not only has a better understanding of what the most suitable action is but also of the options that are actually available. For instance, corporate headquarters are more detached from the day-to-day business of the different divisions and may thus not be aware of all options the division managers could pursue. Similarly, voters tend to have a limited knowledge of available political instruments and legal constraints compared to politicians.[1] Also financial investors differ widely in their financial literacy. They not only face limits in their ability to assess the profitability of particular investments but also have limited awareness of the available investment opportunities.[2]

[1]For a recent application of the classical delegation model in political competition see Kartik et al. (2017). Somin (2016) and Carpini and Keeter (1996) present the results of a number of surveys on US voters over various decades and document the lack of knowledge of basic institutional rules and of the set of policies available to local governments. For example, Somin (2016) documents that 34% of US voters cannot name the three branches of the federal government, a similar percentage do not know which government officials are responsible for which issues; Carpini and Keeter (1996) document that less than 50% of US voters know whether the local governors have to approve the decisions made by their higher state court; 25% do not know whether states can pass a law prohibiting abortion.

[2]Following Baron and Holmström (1980), the standard delegation model has been adopted by several authors to study financial advice, investment banking and delegated portfolio management. The presence of partial awareness of financial products by investors has been recognised since at least Merton (1987). More recently, Guiso and Jappelli (2005) document the lack of awareness of financial assets among the 1995 and 1998 waves of the survey of Italian households (SHIW). Only 65% of potential investors were aware of stocks and only 30% of investment accounts; mutual funds and corporate bonds were known by only 50% of the sample. The share of wealth in the hand of unaware agents was also substantial. The share of wealth owned by households that were not aware of corporate bonds was approximately 20%, and so was the share owned by those unaware of mutual funds.

This paper studies the implications of such asymmetry by incorporating unawareness into the canonical delegation model. We consider the problem of a principal (she) who needs to take an action and delegates the task to an agent (he). The agent receives private information about the payoffs of each of the available actions and the principal's problem is to determine a set of actions from which the agent can choose (see for example Alonso and Matouschek, 2008). We depart from the traditional framework of optimal delegation by considering a situation where the principal is *unaware of some feasible actions*. Our key assumption is that the principal's unawareness limits her language to write a contract: the principal can only permit actions in the delegation set if she can name these actions explicitly, hence, if she is aware of them. Before the delegation stage the agent can expand the principal's awareness by revealing additional actions and thereby enrich the set of feasible contracts for the principal.

We are interested in the question if and how the agent distorts the principal's awareness in order to increase his own rents. We address this question in an environment with a continuum of payoff states, a continuum of feasible actions and an agent who in each state prefers a higher action than the principal. Given her awareness, the principal's optimal delegation set solves the usual tradeoff between minimising distortions and limiting the agent's information rent. Since the agent has an upward bias, optimal delegation entails that the principal limits the agent's choice from above. An optimal delegation set thus has a cap above which no action is permitted. How high this cap is depends on the principal's awareness set. We show under minimal restrictions that the agent optimally leaves the principal unaware of an interval of actions around the optimal upper cap under full awareness. By choosing the bounds of the interval appropriately, the agent makes it optimal for the principal—who still cares about the agent's information—to permit an action above the full awareness cap and, hence, an action that would be precluded if the principal was fully aware. We derive the agent's optimal disclosure policy and the resulting delegation set explicitly for the case of quadratic utility functions and a uniform bias.

The baseline model assumes that the disclosure of feasible actions takes place before the agent receives private information. We relax this assumption in the second part of the paper and allow the agent to propose additional actions after the contract is signed and the agent observes the state. Principal and agent thus play a signalling game in the last stage. The study of this problem reveals interesting implications of asymmetric awareness on strategic information transmission. We characterise the outcome of the agent's best equilibrium and show that the agent can fill all potential gaps of the delegation set below its upper cap. From a modeller's viewpoint the agent's strategy in this equilibrium is fully revealing. The unaware principal, however, cannot compare the agent's proposal at a given state to the actions which the agent would have proposed in a different state. Due to this asymmetry, each of the agent's equilibrium proposals is perceived to be consistent with an interval of states and these intervals overlap. Thus, in contrast to the case of full awareness, the principal's information cannot be represented by a partition of the state space into pairwise disjoint sets. Translated to a canonical cheap talk problem with a single communication phase (Crawford and Sobel, 1982), the result shows that asymmetric awareness allows for substantially finer communication and improved outcomes for both parties compared to the case of full awareness.

We discuss some extensions of the baseline model (without renegotiation). In particular, we consider the case where the set of feasible actions is an arbitrary subset of the reals, where actions are multidimensional, and where the agent does not know the principal's initial awareness. Finally, we discuss the role of the principal's sophistication.

After the literature review, the paper is organised as follows. Section 2 presents the delegation model with limited awareness. In Section 3 we analyse the agent's optimal disclosure and the resulting delegation set. Section 4 analyses the game with renegotiation after the agent receives private information. Section 5 discusses extensions and Section 6 concludes.

1.1 Related Literature

The paper makes both applied and theoretical contributions. It introduces unawareness to the canonical delegation problem and shows that a biased agent has incentives to hide moderate options from the principal in order to implement more extreme ones. The identified distortion may have significant effects on economic outcomes in a range of situations, as argued above. Our analysis builds on the literature on optimal delegation and on applications of unawareness to games and contracting problems. Holmström (1977) first defines the delegation problem and provides conditions for the existence of its solution. Following the seminal paper, the literature was further developed by Melumad and Shibano (1991), Szalay (2005), Martimort and Semenov (2006), Alonso and Matouschek (2008), Kováč and Mylovanov (2009), Armstrong and Vickers (2010), Amador and Bagwell (2013) and Halac and Yared (2020), among others. None of them consider limited awareness in this framework.

There are only few papers that apply unawareness to games in general and contracting problems in particular. In contrast to our setting, most of the existing work considers contracting problems where contingent transfers are feasible and where *the agent* has limited awareness, while the principal is fully unaware (Von Thadden and Zhao 2012; Zhao 2011; Filiz-Ozbay 2012; Auster 2013). One exception is Francetich and Schipper (2020) who consider a screening model where the principal is unaware of certain cost types (but has full awareness over actions) and the agent decides which types to disclose. Lei and Zhao (2020) consider a particular case of our model (quadratic utility, uniform bias, uniform distribution) to study unawareness of contingencies (nature's moves) rather than players' actions.[3]

On the theoretical side, the study of the disclosure problem reveals how the agent's rents depend on the set of feasible actions—or the principal's perception thereof—in delegation settings. This question is related to a recent literature looking at the determinants of agency rents in models with full awareness, e.g. Roesler and Szentes (2017), Garrett et al. (2020).

Finally, the paper introduces a new class of communication games with non-verifiable information about the payoff state and a receiver whose awareness of the possible signals depends on the realised state. We thereby contribute to the literature on signalling and strategic information transmission (Crawford and Sobel, 1982). We show that the

[3]Conceptually, there is also a connection to the literature on incomplete contracts and unforeseen contingencies (Grossman and Hart, 1986). In contrast to that literature, our model entails that the principal is limited in her choice of contracts but correctly foresees the outcome of any contract she can specify.

discrepancy between the sender's signalling strategy and the receiver's perception of it can substantially change the outcomes in such games. The equilibrium we describe also provides a neat illustration for how unawareness differs from the standard model and, in particular, 'zero probability' beliefs. To the best of our knowledge, we are the first to study the equilibrium implications of limited awareness on strategic information transmission in a setting where information about the payoff state is not verifiable.[4] In a different setting, Heifetz et al. (2020) study the strategic disclosure of hard information and find that if the information about the payoff state is multidimensional and the receiver is unaware of some dimension, unraveling is not a necessary outcome of the game.

In a companion paper, Auster and Pavoni (2020), we apply our delegation model (without signalling) to financial intermediation, considering a market with multiple fully aware brokers (agents) and a continuum of partially aware investors (principals). Self-reported data from customers in the Italian retail investment sector support the key predictions of the model: the menus offered to less knowledgable investors contain fewer products, which are perceived to be more extreme.

2 Environment

There is a principal and an agent. The agent has access to an interval of actions $Y^A = [y_{min}, y_{max}]$.[5] The principal's and agent's payoffs depend on the action that is chosen and an unknown payoff parameter θ, which can be privately observed by the agent. Let $\Theta = [0,1]$ be the set of payoff states and let $F(\theta)$ denote the cumulative distribution function on Θ, assumed to be twice differentiable on the support. The principal and the agent have expected utility functions with *continuous* Bernoulli components given by[6]

$$U^P(y,\theta), \quad U^A(y,\theta).$$

[4] In our model the disclosure of feasible actions is verifiable, but the relevant inference only occurs on the payoff state.

[5] We discuss in Section 5 the extension to general subsets of $\mathbb{R}$, for instance, finite collections of points, as well as the case of multi-dimensional actions.

[6] Note that the principal does not have full access to her payoff function U^P but just to a payoff function restricted to the domain of actions of which she is aware.

Fixing θ, $U^i, i = P, A$ is *assumed to be strictly concave* in y with an interior maximum on Y^A. The principal's and agent's conditionally preferred actions are described by the functions

$$y^P(\theta) := \arg\max_{y \in Y^A} U^P(\theta, y), \quad y^A(\theta) := \arg\max_{y \in Y^A} U^A(\theta, y).$$

We assume $U^i_{y\theta} > 0$, which implies that $y^P(\cdot), y^A(\cdot)$ are strictly increasing functions. Furthermore, we assume that conditional on the payoff parameter θ, the agent prefers a higher action than the principal: for all θ, $y^P(\theta) < y^A(\theta)$.[7]

Awareness. Let $\mathcal{Y}$ denote the set of closed subsets of $[y_{min}, y_{max}]$. The principal is aware of a subset of available actions, denoted by $Y^P \in \mathcal{Y}$. Hence, unawareness in our framework does not take the form of unforeseen contingencies but concerns the set of available actions.[8] Apart from the assumption that Y^P is closed, we impose no further structure on the principal's initial awareness set. Before the principal contracts with the agent and the agent observes θ, the agent can make the principal aware of additional actions by revealing a closed set $X \in \mathcal{Y}$. The principal fully understands the options that are revealed to her and accordingly updates her awareness to the union of whatever she knew initially and what the agent reveals.

Delegation. Given her updated awareness, the principal offers a contract to the agent. We rule out monetary transfers and assume that the agent's participation constraint is always satisfied. The contracting problem of the principal then reduces to the decision over the set of actions from which the agent can choose once he observes the payoff parameter θ.[9] Our substantial assumption is that the principal's unawareness restricts the language with which she can write a contract. In particular, we assume that the principal can only refer to actions in the contract which she can name explicitly. The

[7] To ease exposition we assume that the bias is always positive. However, it would be easy to extend the analysis where the bias is negative and, slightly adjusting our assumptions, to the case where the bias changes the sign.

[8] See Karni and Vierø (2013, 2017) for a decision theoretic model capturing this type of unawareness.

[9] The standard delegation problem is equivalent to a mechanism design problem when the principal restricts herself to deterministic allocations (see Alonso and Matouschek (2008) and Kováč and Mylovanov (2009)). Formally, the principal commits to a mechanism that specifies an action as a function of the agent's message.

larger the principal's awareness set is, the richer is the set of contracts she can write. Given the principal's updated awareness set, she then has two natural options: the principal can either name the actions she allows the agent to take or she can name the actions she explicitly forbids. Under full awareness, these two options are clearly equivalent. With unawareness, on the other hand, specifying only the forbidden actions leaves the principal vulnerable to the agent taking actions which the principal does not anticipate. We will discuss this case in Remark 1 in the following section and concentrate now on the case where the principal specifies the actions which she permits. Since the principal cannot specify actions of which she is unaware, the principal's delegation set is then a subset of her awareness set. The timing of the game can be summarised as follows:

1. The principal's initial awareness Y^P is realised and observed by all parties.
2. The agent reveals a closed set of actions $X \subseteq Y^A$ and the principal updates her awareness to $Y = Y^P \cup X$.
3. Given awareness set Y, the principal chooses a delegation set $D \subseteq Y$.
4. The agent observes θ and chooses an action from set D.
5. Payoffs are realised.

Notice that we have not made any explicit assumption on whether or not the principal is aware of her unawareness. The principal might take the world at face value or she might understand that there exist actions outside her awareness. Since she cannot include such actions in the delegation set, awareness of their possible existence neither affects her expected payoff nor optimisation problem.[10] Furthermore, within the constraints of her awareness, the principal is perfectly rational: she anticipates correctly the expected payoff associated to each feasible delegation set and will not be surprised ex-post.

The game between the principal and the agent can be formally represented by a family of partially ordered subjective game trees. Such family includes the modeller's view of the objectively feasible paths of play but also the feasible paths of play as subjectively viewed by some players, or as the frame of mind attributed to a player by other players or by the same player at a later stage of the game. In the online appendix, we provide a

[10]We discuss the principal's sophistication and awareness of unawareness further at the end of Section 5.

more extensive description of the family of game trees representing the generalised game with unawareness associated to our delegation model according to the approach proposed by Heifetz et al. (2013). Figure 3 in the online appendix reports a graphical example.

As a solution concept, we use a strong version of Perfect Bayesian Nash Equilibrium (PBE) which implies subgame perfection, adapted to generalised extensive-form games with unawareness (e.g., see Halpern and Rêgo (2014) and Feinberg (2020)). In the online appendix, we also describe the set of outcomes that satisfy a prudent version of extensive-form rationalizability and we show that whenever we restrict to pure strategies and assume the tie-breaking rules we adopt below to be commonly known, the PBE outcome we obtain is also the sole rationalizable outcome of the generalised game. Rationalizability assumes that players have common knowledge of their rationality and their preferences but it does not assume, for example, that a player is automatically certain of a ready-made convention of play upon becoming aware.[11] Despite the clear appeal of this notion for games with unawareness, our focus on Bayesian Nash equilibrium as a solution concept in the main body of the paper facilitates considerably the comparison to existing results in the literature, especially in Section 4, where signalling arises as part of the game.

3 Equilibrium Analysis

We will work backwards and start the analysis by considering the last stage of the game. Given a delegation set D and observed payoff state θ, the agent's best response for the last stage of the game is defined by

$$BR^A(\theta, D) := \arg\max_{y \in D} U^A(y, \theta). \tag{1}$$

When the agent is indifferent between two actions, let $y^*(\theta, D) := \min BR^A(\theta, D)$ be the selection that takes the smallest value (indifference is broken in favour of the principal).[12]

[11] See Guarino (2020) and Heifetz et al. (2020).

[12] Such selection is well defined as it is easy to show - from the joint continuity of U^A - that the BR^A correspondence is upper hemicontinuous.

Delegation stage. Turning to the principal's delegation choice, we first define the principal's value of delegation set $D \in \mathcal{Y}$ given y^*:

$$V^P(D) := \int_0^1 U^P(\theta, y^*(\theta, D))\mathrm{d}F(\theta). \tag{2}$$

There are typically actions that the principal could permit but the agent will not implement. W.l.o.g. we will restrict attention to delegation sets D such that for any $y \in D$, there is some state $\theta \in [0,1]$ such that $y^*(\theta, D) = y$.[13] Let $\mathcal{D}(Y)$ be the set of delegation sets in $\{D \in \mathcal{Y} : D \subseteq Y\}$ that satisfy this requirement. For each awareness set $Y \in \mathcal{Y}$, the principal's optimal delegation set solves the problem

$$\max_{D \in \mathcal{D}(Y)} V^P(D). \tag{3}$$

Again, if problem (3) has multiple solutions, we assume that the principal choses the agent-preferred set and denote by $D^*(\cdot)$ such selection from the set of maximisers for each Y. Furthermore, we assume that in the case where the principal is fully aware, delegation is valuable. A sufficient condition for valuable delegation is $y_0^* > y^A(0)$, where $y_0^* \in \arg\max_y V^P(\{y\})$. This requires that the bias is not too large and implies that the principal prefers the delegation set $[y^A(0), y_0^*]$ to the singleton $\{y_0^*\}$ (see also Alonso and Matouschek (2008), Corollary 2).

Disclosure stage. In the first stage of the game, the agent chooses an awareness set $Y \in \mathcal{Y}$. Since the agent cannot make the principal unaware of actions which the principal already knows, the induced awareness set must contain the principal's initial awareness set Y^P. The smaller Y^P is, the larger is the collection of awareness sets from which the agent can choose. An optimal awareness set Y^* solves the problem

$$\max_{Y \in \mathcal{Y}} \int_0^1 U^A(\theta, y^*(\theta, D^*(Y)))\mathrm{d}F(\theta) \quad \text{s.t.} \quad Y^P \subseteq Y. \tag{4}$$

Since different awareness sets might induce the same delegation set, the solution to problem (4) is typically not unique. Of course, this type of multiplicity does not affect the

[13] This restriction reduces multiplicities by eliminating 'redundant' delegations sets that contain actions that will not be chosen under any contingency.

outcome. We assume that when two solutions of problem (4) are nested, the agent discloses the larger set. This assumption allows us to distinguish the actions that remain undisclosed for strategic reasons from those that are redundant. Let $\mathcal{Y}^*$ denote the set of all solutions of (4) satisfying this requirement.

Equilibrium disclosure. The central question of this paper is whether the agent distorts the principal's delegation choice in his favour by leaving the principal unaware of some feasible actions. Due to the conflict of interest between the principal and the agent, a fully aware principal will not find it optimal to permit the agent his preferred action in every payoff state. Indeed, since the agent is upward biased, the principal can always improve on full delegation by excluding an interval of high actions, forcing the agent for high realisations of θ to take an action closer to the principal's conditionally preferred action.

Let $\hat{y} := \max D^*(Y^A) < y^A(1)$ denote the *upper cap* which the principal imposes under the optimal delegation set in the full awareness benchmark. The following proposition provides conditions under which unawareness of $\hat{y}$ is sufficient to ensure that the agent benefits from the principal's limited awareness.

Proposition 1 (Optimality of Limited Awareness). *Assume that problem (3) has a unique maximiser $D^*(Y^A)$ and that the upper cap $\hat{y}$ is not an isolated point of $D^*(Y^A)$. If $\hat{y} \notin Y^P$, then generically the agent strictly prefers not to disclose all actions in Y^A.*

Proposition 1 shows that if the principal is initially unaware of the highest action in the optimal delegation set under full awareness, then the agent finds it profitable to hide some of the feasible actions from the principal. To prove the result, we consider a simple perturbation of the full awareness set. The perturbation entails that the principal remains unaware of an interval (y^-, y^+) of actions around the upper cap $\hat{y}$. In the first step, we show that the bounds of the interval, y^- and y^+, can be chosen in a way such that the principal finds it optimal to include both y^- and y^+ in the delegation set. Hence, by leaving the principal unaware of actions around $\hat{y}$, the agent can implement an action $y^+ > \hat{y}$ that is not permitted under full awareness.

The agent's gain in flexibility comes at the cost of losing the option to take an action in the interval $(y^-, \hat{y}]$. In the second step of the proof, we show that the perturbation is profitable for the agent despite this cost. To see this, define $s(\cdot)$ as the inverse of $y^A(\cdot)$ and consider the state $s(\hat{y})$, where the agent's preferred action is $\hat{y}$. By the assumption that $\hat{y}$ is a limit point of $D^*(Y^A)$, there is an interval of actions to the left of $\hat{y}$ that are permitted under $D^*(Y^A)$. This implies that there is an interval of states to the left of $s(\hat{y})$ such that for all states belonging to the interval, the agent gets to take his preferred action under $D^*(Y^A)$. The perturbation forces the agent to move away from his bliss point in these states. However, since the marginal cost of moving away from the bliss point at the bliss point is zero, the effect of losing these actions is second order and thus dominated by the agent's benefit of increasing the implemented action in states to the right of $s(\hat{y})$.

We should emphasise that when the optimal delegation set under full awareness is not an interval, the agent profits from perturbations around other pooling points as well. For example, if $D^*(Y^A)$ has an intermediate gap $(\underline{y}, \overline{y})$, the agent benefits from moving up the lower bound $\underline{y}$ at the cost of losing some flexibility below $\underline{y}$.[14] The main complication here is that such perturbation may affect the principal's optimal choice of $\overline{y}$. If the optimal value for $\overline{y}$ decreases as a result of the perturbation, the agent strictly gains. If, on the other hand, it increases, then there are two first-order effects which need to be compared. To guarantee the profitability of the perturbation in this case, more stringent assumptions on the principal's initial awareness set are needed in order to give the agent the necessary tools to deter the principal from undesired movements of adjacent pooling actions. In the described situation, for instance, it might be necessary to keep the principal unaware of some actions to the right of $\overline{y}$.

Proposition 1 is a consequence of a more general principle. Revealing an action y to the principal typically has a benefit and a cost. Conditional on the principal permitting y, the benefit of revelation is the utility gain in the states where y is preferred by the agent. The downside is that the action may crowd out other actions which the principal would permit if she remains unaware. In regions of θ where the principal gives full discretion, crowding out is not an issue, so the agent optimally discloses the relevant options. In

[14]Similarly, the agent benefits from lowering $\overline{y}$ at the cost of losing some actions above $\overline{y}$.

regions where the conflict of interest is instead severe, the principal optimally restricts the agent's choice and full revelation can be detrimental to the agent. In the case of Proposition 1, the action $\hat{y}$ crowds out all actions higher than $\hat{y}$. Since the agent benefits from being permitted such actions, he optimally leaves the principal unaware of $\hat{y}$ (and some actions around it).

The same principle applies to other screening problems. When deciding on the contract, a principal facing a privately informed agent solves a tradeoff between eliciting the agent's information to take a suitable action and limiting the agent's information rent. *By making specific actions unavailable, the agent can increase the cost for the principal not to use the agent's information and thereby increase her willingness to grant the agent higher information rents.*

Interval delegation. The literature on optimal delegation establishes sufficient conditions under which the optimal delegation set under full awareness is an interval. These conditions assure that any delegation set that has gaps can be improved upon by adding intermediate actions to the set. Assumption 1 makes this requirement explicit.

Assumption 1. *Consider a delegation set $D \in \mathcal{Y}$ and its convex hull $Conv(D)$. Then, for all $A \subseteq Conv(D)$:*

$$\int_0^1 U^P(\theta, y^*(\theta, D))dF(\theta) \leq \int_0^1 U^P(\theta, y^*(\theta, D \cup A))dF(\theta).$$

Consider a convex delegation set and suppose the principal removes an interval of actions in the interior of the set. Let this interval be denoted by $(\underline{y}, \overline{y})$. The removal of actions in $(\underline{y}, \overline{y})$ means that there is an interval of states where the agent switches to the lower action $\underline{y}$ and an interval of states where the agent switches to the higher action $\overline{y}$ with respect to the original delegation set. Since the principal is downward biased with respect to the agent, the switch to the lower action benefits her, whereas the switch to the higher one does not. Concavity of U^P means that the principal is risk-averse and therefore has incentives to hedge against these two possibilities. Hence, unless the principal views the scenario of the beneficial switch considerably more likely, she favours intermediate actions. The literature on optimal delegation provides conditions on the state distribution with respect to the utility functions that guarantee this property. We provide a set of sufficient assumptions below. For more general conditions we refer the reader to Alonso and Matouschek (2008, Proposition 5 and Amador and Bagwell (2013, Propositions 1 and 2).

Assuming that interval delegation is optimal, the optimal delegation set under full awareness is described by an upper cap below which the agent is free to choose his preferred action. The optimal delegation set under full awareness thus takes the form $[y^A(0), y]$ for some $y < y^A(1)$.[15] The optimal cap is the value of y that maximises $W : Y \to \mathbb{R}$, where

$$W(y) := \int_0^{s(y)} U^P(\theta, y^A(\theta))\,\mathrm{d}F(\theta) + \int_{s(y)}^1 U^P(\theta, y)\,\mathrm{d}F(\theta). \tag{5}$$

We now show that unawareness of the optimal cap, which we denote again by $\hat{y}$, is not only a sufficient condition for less-than-full revelation to be strictly optimal but also a necessary one. We further show that the resulting delegation set has a single gap around $\hat{y}$.

Proposition 2 (Optimal Delegation under Limited Awareness)**.** *Let Assumptions 1 be satisfied. (i) The agent optimally reveals all feasible actions to the principal if and only if the principal is aware of the action $\hat{y}$. (ii) If in addition W in (5) is single peaked,*[16] *there exist two parameters $\Delta_1, \Delta_2 \geq 0$ such that:*

$$\begin{aligned} Y^* &= (y_{min}, \hat{y} - \Delta_1] \cup [\hat{y} + \Delta_2, y_{max}), \\ D^*(Y^*) &= [y^A(0), \hat{y} - \Delta_1] \cup \{\hat{y} + \Delta_2\}, \end{aligned}$$

with $\Delta_1, \Delta_2 > 0$ if and only if $\hat{y} \notin Y^P$.

Proposition 2 shows that when the principal is initially aware of $\hat{y}$ the agent optimally reveals everything. The principal will not allow the agent to take any action higher than $\hat{y}$, so the agent maximises his discretion by revealing all actions below $\hat{y}$. Hence, the agent cannot improve on the full awareness delegation set $[y^A(0), \hat{y}]$. On the other hand, if the principal is initially unaware of $\hat{y}$, it is optimal for the agent to leave the principal

[15] Recall, $y^A(\cdot)$ is the unrestricted choice function for the agent.

[16] Over the unidimensional space, single-peakedness is equivalent to strict quasi-concavity.

unaware of an interval around $\hat{y}$. The resulting delegation set includes all relevant actions below the interval and one action above it.

Remark: *Our model captures situations where the principal is unaware of unidimensional actions, such as economic policies, financial and non-financial products, procurement tasks, etc. Heifetz et al. (2020) consider a situation where actions are multidimensional and a receiver is unaware of certain dimensions (interpreted as attributes of the action). We could consider this type of unawareness in the context of our problem. If the language of the contract is such that the agent is free to choose on dimensions which are left unspecified, it is clear that the agent would have no incentives to reveal additional attributes of actions to the principal. Since disclosing new attributes can only reduce the agent's flexibility, disclosure cannot be profitable. The same is true in our one-dimensional setting if instead of specifying the actions that are permitted, the principal would specify those actions that are not permitted. Hence, in contrast to the benchmark case of full awareness where both types of specifications lead to the same outcomes, the language in which a contract is written matters here.*[17]

3.1 Quadratic Utility and Uniform Bias

For a concrete illustration of the main results and an explicit solution of the agent's optimal disclosure policy, consider the specification

$$U^P(y,\theta) = -(y-(\theta-\beta))^2, \quad U^A = -(y-\theta)^2. \tag{6}$$

The agent's conditional preferred action is $y^A(\theta) = \theta$, while the principal's preferred action is $y^P(\theta) = \theta - \beta$. The agent thus has a constant upward bias equal to β. In this environment, a condition implying Assumption 1 and hence guarantying interval delegation to be optimal is the following regularity condition on the distribution function (see Martimort and Semenov, 2006):

$$f'(\theta)\beta + f(\theta) > 0 \text{ for all } \theta \in (0,1). \tag{7}$$

[17] One could generalise our model allowing for a more complex relationship between principal awareness and her available actions. The key trade off faced by the agent in our model will be preserved as long as a more aware principal has both a non-trivial possibility of enlarging the set of actions permitted to the agent and the possibility of excluding some actions.

Delegation is valuable for the principal if $\mathbb{E}[\theta - \beta] > 0$. When these two conditions are satisfied, the optimal delegation set under full awareness is an interval $[0, \hat{y}]$, where $\hat{y}$ solves the following equality (Martimort and Semenov, 2006, and Alonso and Matouschek, 2008):[18]

$$\hat{y} = \mathbb{E}[\theta - \beta | \theta \geq \hat{y}]. \tag{8}$$

The following proposition extends the characterisation of the optimal delegation set to arbitrary awareness sets $Y \in \mathcal{Y}$. The proofs for the results in this section are reported in the Online Appendix.

Proposition 3 (Generalised Interval Delegation)**.** *Let $Y \in \mathcal{Y}$ and define $\hat{y}_Y := \arg\min_{y \in Y} |y - \hat{y}|$. If condition (7) is satisfied, the optimal delegation set with respect to awareness set Y is*

$$D^*(Y) = \{y \geq 0 : y \in Y \text{ and } y \leq \hat{y}_Y\}.$$

Proposition 3 shows that the optimal delegation set for awareness set Y contains all actions belonging to Y that are weakly smaller than the element of Y which is closest to $\hat{y}$. The optimal delegation set under partial awareness can thus be seen as the closest approximation of the optimal delegation interval under full awareness, $[0, \hat{y}]$, which is available to the principal given her restricted awareness. This approximation includes an element $y > \hat{y}$ if and only if y is closer to $\hat{y}$ than any element of Y smaller than $\hat{y}$, as illustrated in Figure 1.

Turning the attention to the agent's optimal disclosure, we know from Proposition 2 that if the principal is aware of the threshold action $\hat{y}$, the agent optimally reveals all other actions. Indeed, since there is no action closer to $\hat{y}$ than $\hat{y}$ itself, the upper bound of the optimal delegation set will always be $\hat{y}$. Disclosing actions above $\hat{y}$ is thereby irrelevant; the principal will never allow the agent to implement any of them. On the other hand, revealing actions below the threshold $\hat{y}$ is strictly optimal since they will be included in the optimal delegation set, therefore expanding the agent's choice.

Starting from an arbitrary set Y^P, the above argument implies that the optimal

[18] If instead $\mathbb{E}[\theta - \beta] < 0$, the optimal delegation set is $\{\mathbb{E}[\theta - \beta]\}$.

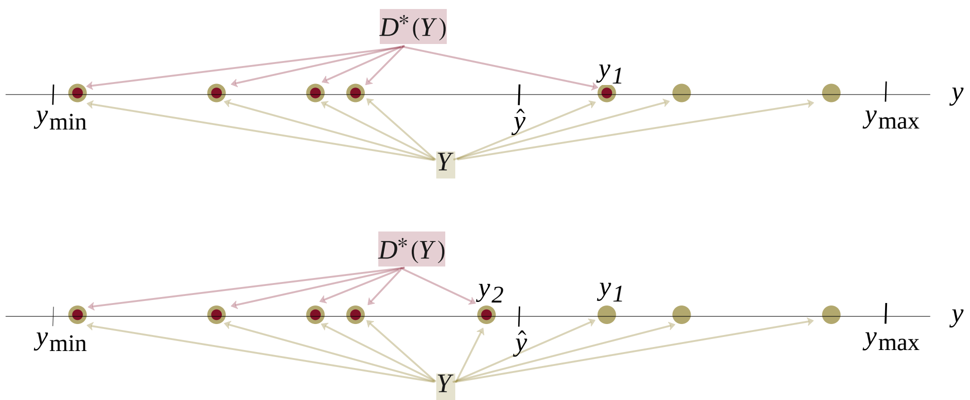

Figure 1: Optimal delegation set $D^*(Y)$. The figures represent two examples of the principal's awareness set Y. In both figures, the yellow bullets represent the set Y while the red bullets represent the resulting optimal delegation set $D^*(Y)$. In the upper figure, the principal includes action y_1 in the delegation set, as it is the closest action to $\hat{y}$. In the lower figure, the principal is aware of action y_2 as well and, for this reason, she excludes action y_1 from $D^*(Y)$.

awareness set Y^* is such that the upper bound of the corresponding delegation set $D^*(Y^*)$ is at least $\hat{y}$. Moreover, the only reason for the agent to leave the principal unaware of certain actions is to induce the principal to permit some action strictly greater than $\hat{y}$. By Proposition 3 this is optimal for the principal if and only if the principal is not aware of any action closer to $\hat{y}$. Hence, the gap around $\hat{y}$ is symmetric and the optimal awareness gap described in Proposition 2 is determined by a single parameter $\Delta_1 = \Delta_2 = \Delta$. The corresponding delegation set is then given by:

$$D^*(Y^*) = [0, \hat{y} - \Delta] \cup \{\hat{y} + \Delta\}.$$

Given such delegation set, the agent's optimal policy is as follows. In states below $\hat{y} - \Delta$ the agent uses his flexibility and implements his preferred action $y = \theta$. In states above $\hat{y} - \Delta$ the preferred action is not available, so the agent chooses the one closest to his bliss point. For states in the interval $(\hat{y} - \Delta, \hat{y}]$ the closest action is $\{\hat{y} - \Delta\}$, for the remaining states it is $\{\hat{y} + \Delta\}$.

Using this policy, we can write the agent's expected payoff as a function of Δ. The agent chooses the value of Δ that maximises his expected payoff. This choice is restricted by the principal's initial awareness set. Letting $\bar{\Delta}(Y^P) := \arg\min_{y \in Y^P} |y - \hat{y}|$ indicate

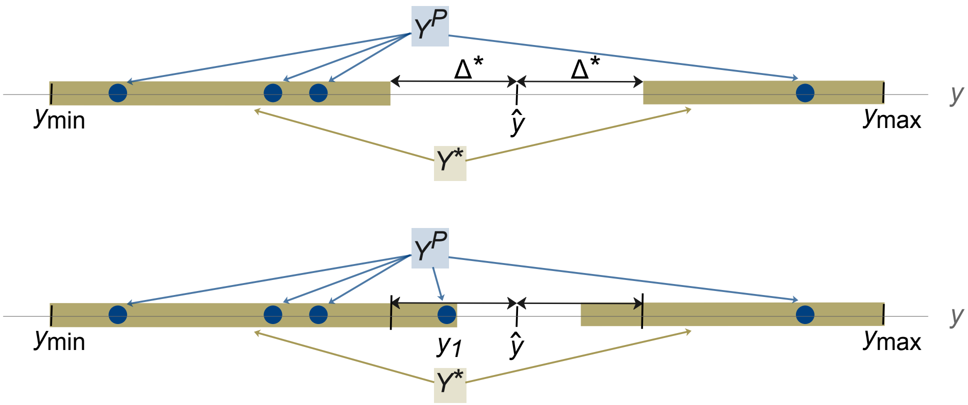

Figure 2: Optimal awareness set Y^*. The figures represent two examples of the principal's initial awareness set Y^P and associated awareness sets $Y^* = Y^P \cup X^*$ after including disclosed actions X^*. In both figures, the blue bullets represent the set Y^P, while the yellow set represents the resulting optimal awareness set Y^*. In the upper figure, the agent keeps the principal unaware of the interval $(\hat{y} - \Delta^*, \hat{y} + \Delta^*)$. In the lower figure, the principal is also aware of action y_1, making the unconstrained solution Δ^* infeasible.

the maximum feasible awareness gap, the agent's optimisation problem amounts to:

$$\max_{\Delta \geq 0} \quad -\int_{\hat{y}-\Delta}^{\hat{y}} (\hat{y} - \Delta - \theta)^2 dF(\theta) - \int_{\hat{y}}^{1} (\hat{y} + \Delta - \theta)^2 dF(\theta) \quad \text{s.t.} \quad \Delta \leq \bar{\Delta}(Y^P). \tag{9}$$

The solution to problem (9) is given by $\min\{\bar{\Delta}(Y^P), \Delta^*\}$, where Δ^* solves the first-order condition

$$\int_{\hat{y}-\Delta^*}^{\hat{y}} [\theta - (y - \Delta^*)]\, dF(\theta) = \int_{\hat{y}}^{1} [\theta - (y + \Delta^*)]\, dF(\theta). \tag{10}$$

A gap parametrised by Δ^* is implemented whenever the principal's initial awareness does not constrain the agent in his choice. If, however, the principal is aware of some action in the interval $(\hat{y} - \Delta^*, \hat{y} + \Delta^*)$, the agent's optimal strategy is to choose the largest feasible gap, as we illustrate in Figure 2.

Proposition 4 (Comparative Statics)**.** *Let $\Delta^*(\beta)$ be the unconstrained solution to problem* (9) *when the principal's preferences parameter is $\beta \in (0, \mathbb{E}[\theta])$ and condition (7) is satisfied. Then $\Delta^*(\cdot)$ is an increasing function.*

Proposition 4 shows an intuitive result: the larger is the divergence between the principal's and the agent's preferred action, the more the agent wants to distort the principal's delegation choice by hiding actions from the principal. For a simple illustration, consider the case where F is uniform.[19] The larger β is, the lower is the cap $\hat{y}(\beta)$ of the optimal delegation set under full awareness, as can be seen from condition (8). Considering the agent's tradeoff when choosing Δ, notice that when F is uniform, the cost associated to the loss of flexibility for a given gap Δ around $\hat{y}(\beta)$ is the same for all β. The desired consequence of generating a gap is an increase of the highest permitted action—from $\hat{y}(\beta)$ to $\hat{y}(\beta) + \Delta$—and hence an increase of the agent's information rent in all states above $\hat{y}(\beta) + \Delta$. The lower the original cap $\hat{y}(\beta)$ is, the larger is the range of values for θ above $\hat{y}(\beta) + \Delta$ and hence the set of types to whom this rent accrues.

4 Renegotiation

Thus far, we have assumed that the agent can only reveal actions to the principal before he learns the payoff state θ. This entails that even in states where the agent knows of an action that makes both parties better off, additional communication is not possible. An interesting question is how the outcome changes if after learning the payoff parameter θ, the agent can reveal additional actions to the principal, who then decides whether to permit a new action or to maintain the original contract. We thus consider a model where the agent can renegotiate with the principal. If the agent proposes to replace the original delegation set with a new action, the principal understands that the agent's choice signals something about the state. In particular, the principal can infer that the agent only reveals an action if that benefits him. However—due to the principal's limited awareness—she cannot conceive of alternative actions the agent could have disclosed. This implies that the principal cannot learn from particular actions not being proposed, a key difference to the case of full awareness.

Modified game. The possibility of renegotiation considerably changes the nature of the game. While the probability distribution over θ is common knowledge, renegotiation occurs under asymmetric information, implying that the principal and the agent play a signalling game. We will distinguish two phases of the game: the contracting phase and

[19] It turns out that our argument holds true for all distributions satisfying condition (7).

the renegotiation phase. To simplify the analysis, we assume that the principal is initially unaware of the possibility of renegotiation. The contracting phase is then the same as before: the agent reveals a set of actions to the principal and the principal chooses the optimal delegation set with respect to her updated awareness set. Note, however, that the equilibrium we derive continues to exist when this restriction is removed. According to the principal's perception, the allocation of any pure-strategy equilibrium involving renegotiation can be replicated by directly adding the actions that are permitted on path to the delegation set. Since the principal cannot foresee the possibility of renegotiating over actions of which she is unaware in the contracting phase, she sees no value in renegotiating the optimal contract.

In the renegotiation phase, the agent learns the realised value of θ and can propose additional actions to the principal. The principal updates her awareness and her beliefs about the payoff state. We restrict attention to a game where the agent proposes single actions to the principal in the renegotiation phase.[20] The principal's decision is then between maintaining the original delegation set and permitting the proposed action. A detailed description of the strategies of the two players and the definition of Perfect Bayesian Equilibrium (generalised to allow for differential awareness) is reported in Appendix A.3. In the online appendix, we describe the generalised extensive-form game à la Heifetz et al. (2013).

Disclosure in the renegotiation phase. As usual, we move backwards and start with the analysis of the renegotiation stage. The following proposition provides a key property of our model: in any equilibrium, the principal permits an agent's proposal only if she whould have included the action in the delegation set at the contracting phase.

Proposition 5 (Ex-ante like Reasoning)**.** *Let Y denote the principal's awareness set at the beginning of the renegotiation phase. (i) Fix an equilibrium, and let A be the (measurable) set of proposals which the principal accepts in the renegotiation phase. Then, for all $x \in A$,*

$$V^P(D^*(Y) \cup \{x\}) \geq V^P(D^*(Y)). \tag{11}$$

[20]It can be easily shown that the agent cannot improve on the agent-best equilibrium we describe in Proposition 6 by revealing more than one action at a time.

(ii) Conversely, for any set A constituted of points satisfying condition (11)*, there is an equilibrium such that A is the set of actions accepted in the renegotiation phase.*

Recall that V^P is defined as the expected utility of the principal under 'pure' delegation, that is, the case where the principal is not expecting to have any renegotiation (that is why the payoff only depends on the delegation set). Proposition 5 characterises the set of 'implementable' proposals: for each possible awareness level Y, it defines the set of proposals that can be accepted by the principal in equilibrium. Condition (11) vacuously holds for $x \in D^*(Y)$. Accepting the proposal or keeping the original delegation set results in the same action. On the other hand, for actions not belonging to $D^*(Y)$, the inequality can be satisfied only if x does not belong to Y. Hence, the agent can only gain from renegotiation if he discloses new actions with respect to the contracting phase.

To prove Proposition 5, we first show that there is no equilibrium in which the principal permits an action in Y which does not belong to $D^*(Y)$. Building on this result, we argue that in an equilibrium where proposal $x \notin Y$ is accepted in the renegotiation phase, the principal's beliefs after proposal x are described by the set

$$\{\theta \in [0,1] : U^A(\theta, x) \geq \max_{y \in D^*(Y)} U^A(\theta, y)\}. \tag{12}$$

According to the principal's awareness, x is the only new action which the agent can propose. The principal thus believes that the agent proposes x whenever he prefers it over his best alternative in $D^*(Y)$. The question is then whether conditional on the agent preferring x over the actions belonging to $D^*(Y)$, the principal prefers x as well. The answer to this question is yes if and only if the principal would have preferred to add x to the delegation set $D^*(Y)$, i.e., if and only if (11) is satisfied. Indeed, adding an action to the delegation set changes the outcome only in those states where the agent prefers the action to the alternatives in the delegation set. In the renegotiation phase, the same consideration applies.

Disclosure in the contracting phase. The disclosure of actions in the contracting phase determines a delegation set, which in turn determines a set of actions that can be implemented through further revelation in the renegotiation phase. For each initial disclosure, the set of actions that are ultimately implementable is characterised by (11).

Due to the signalling nature of the game, there are typically multiple equilibria. For instance, there is always a trivial equilibrium, where the agent believes that no proposal will be accepted and hence does not renegotiate. We are interested in the question of *what is possible* to achieve through strategic disclosure of actions and focus on the outcome of the best equilibrium for the agent.

Letting $Z : \mathcal{Y} \to \mathcal{Y}$ describe the mapping from awareness sets to the *maximal set* of implementable actions, the agent solves the problem

$$\max_{Y \supseteq Y^P} \quad \int_0^1 \max_{y \in Z(Y)} U^A(\theta, y) dF(\theta),$$

$$\text{where} \qquad Z(Y) := \{z \in Y^A : V^P(D^*(Y) \cup \{z\}) \geq V^P(D^*(Y))\}.$$

Without further restrictions, the analysis of this problem can be intricate. For an illustration, consider the following example with three actions $y_1 < y_2 < y_3$. Suppose the principal's initial awareness set is $Y^P = \{y_2\}$ and she has the following preferences:

$$V^P(\{y_1, y_3\}) > V^P(\{y_1\}) > V^P(\{y_1, y_2\}) > V^P(\{y_1, y_2, y_3\}).$$

Conditional on action y_2 being in the delegation set, the principal does not want to include action y_3. In order for the agent to take action y_3, he must make the principal aware of y_1 in the contracting stage. Action y_1 crowds out y_2 and thereby opens up the possibility for the principal to permit action y_3. The agent could also reveal y_1 and y_3 in the contracting phase so that the principal delegates the set $\{y_1, y_3\}$. To see why it can be better to hold back with action y_3, suppose there is a fourth action y_4 such that

$$V^P(\{y_1, y_3\}) > V^P(\{y_1, y_4\}), V^P(\{y_1, y_3, y_4\}) > V^P(\{y_1\}).$$

Revealing y_3 in the contracting phase implies that the agent will not be allowed to take action y_4. If instead the agent initially discloses only y_1, he can implement both y_3 and y_4, depending on which action he prefers after observing the realisation of θ.

In order to simplify the analysis, we impose some more structure on the principal's preference over delegation sets and require Assumption 1 to be satisfied. To recall, Assumption 1 assures that the principal always benefits from closing gaps in the delegation set. In order to state the result, we define the highest action that the principal is willing to delegate in the contracting phase for some awareness set that is consistent with the principal's initial awareness:

$$y^{max} := \max\{y \in Y^A : y \in D^*(Y) \text{ for some } Y \supseteq Y^P\}$$

For concreteness, consider the example of quadratic utility functions and a uniform bias, as specified in Section 3.1. Under this specification, the highest implementable action is given by $\hat{y}+\bar{\Delta}(Y^P)$, where $\bar{\Delta}(Y^P)$ was defined as the minimal distance between an action belonging to Y^P and the upper bound of the optimal delegation set under full awareness, $\hat{y}$.

The last-stage outcome of a pure-strategy equilibrium is described by a function $y^E : \Theta \to Y^A$, which maps each state θ to the implemented action.

Proposition 6 (Agent Best Equilibrium)**.** *Let Assumption 1 be satisfied. In the agent's best equilibrium of the renegotiation game, the outcome is described by*

$$y^E(\theta) = \begin{cases} y^A(\theta) & \text{if} \quad \theta \le s(y^{max}) \\ y^{max} & \text{if} \quad \theta > s(y^{max}). \end{cases}$$

Recall that $s(y) = (y^A)^{-1}(y)$ is the state in which y is the agent's preferred action. Proposition 6 shows that when the agent can renegotiate after observing the realisation of θ, he is able to implement all actions below y^{max}. Hence, by disclosing actions in sequence the agent cannot only increase the highest permitted action (as in the case without renegotiation), but he can do it without any loss of flexibility below the cap.

In Appendix A.5, we provide an example of equilibrium strategies and beliefs delivering the equilibrium outcome described by y^E. For a simple illustration, let us consider the case where the principal's optimisation problem under full awareness has a unique local maximum (W in (5) is single peaked). Under this assumption, we can restrict attention to initial disclosures that induce an awareness set with a single gap.[21] Suppose that in the contracting phase, the agent discloses a set $Y_{\bar{\Delta}}$ to generate the largest possible gap in the delegation set:

[21]See the proof of Proposition 2.

$$D^*\left(Y_{\bar{\Delta}}\right) = [y^A(0), \hat{y} - \bar{\Delta}] \cup \{y^{max}\}.$$

Moving to the renegotiation phase, we distinguish three different regions. If $\theta \leq s(\hat{y}-\Delta)$, there is no need for the agent to renegotiate, since the agent's preferred action already belongs to the delegation set. In states belonging the interval $(s(\hat{y}-\Delta), s(y^{max}))$, the agent proposes his preferred action $y^A(\theta)$. By Assumption 1, the principal would have preferred to add this action to the initial delegation set. As we have shown in Proposition 11, this implies that the principal is willing to accept $y^A(\theta)$ in the renegotiation phase. Hence, the principal permits all actions in the interval $(\hat{y} - \bar{\Delta}, y^{max})$. Finally, if $\theta > s(y^{max})$, the agent would like to take an action that is higher than y^{max}. However, by definition of y^{max},

$$V^P(D^*(Y_{\bar{\Delta}})) > V^P(D^*(Y_{\bar{\Delta}}) \cup \{x\})$$

holds for all $x > y^{max}$, which, by Proposition 5, implies that the principal rejects all such proposals. The best alternative for the agent is thus to take y^{max}. The argument further implies that there is no equilibrium of the renegotiation game where the principal permits a higher action than y^{max}. Hence, the outcome described in Proposition 6 maximises the agent's expected payoff. Notice also that y^{max} is equal to $\hat{y}$ (the optimal cap under full awareness) if and only if $\hat{y} \in Y^P$.

Dynamic awareness. The model with renegotiation highlights two important aspects of games with limited awareness. The first concerns the dynamics of unawareness. Much like information, awareness is not reversible. This means that if a player becomes aware of an action today, he remains aware of that action in the future (similarly for outcomes, events, etc.). Hence, the more a player reveals at an early stage of the game, the smaller is the collection of the opponent's awareness sets from which he can choose later on. When there is uncertainty about the future, this creates incentives to hide feasible actions from the other player until later stages of the game. In our environment, this principle is reflected in the fact that the agent reveals fewer actions in the contracting phase when renegotiation is possible than when it is not. In the case of quadratic utility functions, the optimal value of Δ parametrising the delegation set in the contracting phase is maximal when renegotiation is possible. Notice that even without renegotiation, the agent could implement *any single action below* y^{max} by revealing the 'right' set of actions. However, he cannot not implement *all actions below* y^{max} because some actions crowd out others. The agent has to make a choice based on the expected value of the feasible awareness sets and the resulting delegation sets. When renegotiation is possible, on the other hand, the agent can condition the principal's awareness on the realisation of θ.

Information transmission. In the renegotiation stage, the principal and the agent play a signalling game. Unless $y^{max} = \hat{y}$, the delegation set from the contracting phase has a gap below y^{max}. One striking feature of the equilibrium described in Proposition 6 is that the agent's implemented action is strictly increasing in θ for all θ such that $y^A(\theta) \leq y^{max}$. In other words, there is no pooling of types below y^{max}. This would not be possible under full awareness: in any candidate equilibrium where types separate themselves through their announcement, the fully aware principal learns the payoff state and has incentives to deviate to a strictly lower action at least in some states. In the case of limited awareness, the principal cannot contemplate moves of the agent of which she is unaware and this limits the extent to which she infers information from the agent's recommendation. In particular, if the realised value is θ and the agent proposes $y^A(\theta) \notin Y$, the subjective game tree that represents the principal's frame of mind after updating does not include moves of the agent involving an action just below or above $y^A(\theta)$. As a consequence, the principal cannot conceive of the fact that she would have permitted such actions if the agent had proposed them instead. In the principal's subjective game, there is an equilibrium where the agent reveals $y^A(\theta)$ in all states where the agent prefers $y^A(\theta)$ over the actions in the initial delegation set. Conditional on that information, the principal indeed prefers action $y^A(\theta)$ to the initial delegation set.

4.1 Communication without commitment.

A general takeaway from this analysis is the insight that *reduced information inference due to limited awareness can foster communication in situations with a conflict of interest.* The effect is not restricted to our renegotiation game but also arises in other communication settings. To see this, consider Crawford and Sobel's (1982) canonical cheap talk model. Their famous result shows that in the absence of commitment, any conflict of interest severely limits the information transmission that can be sustained in equilibrium. Communication is necessarily coarse. Our renegotiation game corresponds to the Crawford and Sobel (1982) setting if we remove the commitment assumption of the principal. Indeed, without commitment the initial delegation set plays no role in the renegotiation stage. Hence, for a fully aware principal the renegotiation game is identical to the standard cheap talk game.

What are the consequences of removing the principal's commitment and how does the set of equilibria differ from that of the standard cheap talk game? As under full awareness, removing the principal's commitment changes profoundly the nature of the game and Proposition 5 no longer applies. First of all, when becoming aware of a new action, the principal's equilibrium inference is no longer based on the initial delegation set as in (12) but on an analogous condition, where the 'reference set' $D^*(Y)$ is replaced by the set S of actions proposed by the agent and accepted by the principal in all contingencies θ where the principal's awareness does not increase.[22] Still, after correcting for the principal's inference, one might hope to characterise the set of implementable actions by the analogue of (11) with S replacing $D^*(Y)$. This is however not possible, since in the absence of commitment, the principal (*not the agent*) gets to pick her (interim) most preferred action in $Y \cup x$.

Intuitively, the best case scenario in terms of implementable actions and the agent's payoff is when S is small—so that inference is coarse—and Y is small—so that the principal's choices at the interim stage are limited. Consider an extreme case where the principal is initially aware of a single action y. If the agent reveals no further actions to the principal, the principal is forced to take y. The situation is thus analogous to one with a contracting phase and an initial delegation set $\{y\}$. Following our arguments above, it is then easy to see that the communication game without contracting has an equilibrium where the principal accepts all recommendations x such that $V^P(\{y, x\}) \geq V^P(\{y\})$. Since the agent is upward biased, this condition is always satisfied for $x < y$. It will also holds for certain $x > y$, as long as y is sufficiently low. Hence, in stark contrast to the case of full awareness, communication without commitment can result in full disclosure

[22] Since the cheap talk game with constant awareness has typically multiple equilibria, S is not uniquely determined by Y but depends on the equilibrium.

of information (from a modeller's point of view) on an interval of states.

Corollary 7. *(Cheap Talk with Unaware Receiver) Suppose the principal cannot commit and communication takes place after the agent observes θ. If $Y = \{y\}$, then there is a PBE with equilibrium outcome*

$$\forall\, \theta \in [0,1], \qquad y^E(\theta) = \max_{x \in \hat{Z}(\{y\})} U^A(\theta, x),$$

$$\textit{where} \qquad \hat{Z}(\{y\}) = \left\{ z \in Y^A : V^P(\{y, z\}) \geq V^P(\{y\}) \right\}.$$

Finally, it is interesting to point out that improvements in the information transmission due to limited awareness not only benefit the agent but typically also the principal gains. To see this, let Assumption 1 be satisfied and consider an initial awareness set Y and a communication equilibrium with A as the set of actions the principal takes on path. If Y is sufficiently rich, then equilibrium information transmission is necessarily coarse, which means that A is not an interval. By shrinking the set Y, e.g. to a singleton, gaps between actions belonging to A can be filled. Under Assumption 1, the principal benefits from closing such gaps ex-post. Hence, in a situation where the principal has no commitment, the principal may actually benefit from having limited awareness.

5 Discussion

Although we hope that our framework is able to bring important insights about general principal-agent problems where the agent has superior awareness of actions, we focus on a relatively simple delegation framework to identify the main effects. In this section, we discuss a few modifications of our baseline model (without renegotiation) to which the results directly extend.

Set of feasible actions. So far we assumed that the set of available actions is an interval. A possible concern is that the principal, despite being unaware of certain elements of the set of feasible actions, might understand that this set is an interval and try to implement the optimal delegation set under full awareness by describing it indirectly. In particular, the principal could attempt to include actions outside her awareness, maybe through a description of the properties of such actions. Although Proposition 1 relies on a variational argument and hence uses the fact that Y^A is a continuum, the general point we are making also applies to the case with discrete Y^A, so that *a priori* there is no specific structure of the set of available actions—or simply the awareness set of the agent—that might be commonly known. In both cases, the agent has incentives to leave the principal unaware of actions around the optimal cap under full awareness.

For an illustration consider the specification with quadratic utilities and a uniform bias as in Section 3.1 and assume that Y^A is an arbitrary closed subset of $\mathbb{R}$. The analysis of the optimal delegation set for a given awareness set $Y \subseteq Y^A$ remains valid, so we have $D^*(Y) = \{y \in Y : y \leq \hat{y}_Y\}$. With regard to the optimal awareness set, notice that if the agent reveals some $y \in Y^A$, he might as well reveal all actions that have a greater distance to $\hat{y}$ than y: their inclusion will weakly expand the agent's choice set. This implies that the optimal awareness set can again be described by a gap Δ.

Whether or not the agent reveals all feasible actions to the principal then depends on the particular form of Y^A and the principal's initial awareness Y^P. A sufficient condition for full awareness is $\hat{y}_{Y^P} = \hat{y}_{Y^A}$, i.e. the principal is aware of the action that is closest to $\hat{y}$. When this is not the case, the agent leaves the principal unaware of intermediate actions, provided they are close enough to $\hat{y}$ and there exists a greater action than $\hat{y}_{Y^A}$ that is implementable given the principal's initial awareness.

Multidimensional Actions One might further wonder how our analysis extends to situations where choices are multidimensional. First of all, in several applications, although choice sets are multidimensional, the choice is often restricted either by budget/resource or other feasibility constraints so that the relevant choice becomes unidimensional (e.g., Amador et al. (2006) and Amador and Bagwell (2013)).

In addition, the results of this paper directly apply to the case where y represents an index obtained by aggregating the different actions. For $n = 1, 2, \ldots N$, let $x_n \in X_n^A$, with $X^A := \times_{n=1}^N X_n^A$ an arbitrary subset of $\mathbb{R}^N$. Then we might assume there is a function $g : X \to \mathbb{R}$ such that $y := g(x_1, \ldots, x_N)$, with $Y^A = G(X^A)$ and $U^i(\cdot, \theta)$ for $i = A, P$ as before. Obviously, not knowing y amounts to not knowing any $x \in X^A$ such that $g(x) = y$. In a finance application, x might be a vector of attributes of the asset (return, volatility, liquidity, maturity, etc.) and the index y might be interpreted as 'the asset', which must be confronted to 'the state of the market' θ.

Another multidimensional case where our results apply is when θ is unidimensional and the objective function is additively separable and identical across the different choice dimensions. In this case, the optimal contract under full awareness replicates the same delegation set in each dimension (Koessler and Martimort, 2012).

Allowing for more general frameworks, such as those with a multidimensional state θ and multidimensional and/or asymmetric aggregators significantly complicates the model. Such multi-dimensional setups, among other things, roughly imply the possibility of imperfect transfers between the principal and the agent. This reduces the possibility of pooling and induces agent's choices that are generically different from the agent's bliss points for a given dimension (see Koessler and Martimort 2012). Our results use the pooling point of the unidimensional case, so an extension to this framework would in general require a separate analysis. The main principle, however, should apply to this case as well. Take for example the bi-dimensional version of the quadratic case considered above, and assume that $\beta_1 = -\beta_2 > 0$. This is a particularly simple case, since under full awareness, the principal is able to obtain her first-best allocation: $y_i(\theta) = \theta - \beta_i$, $i = 1, 2$. It is immediate to see that the agent can increase his information rents by hiding actions $[-\beta_1 - \Delta_1, 0]$ in the first dimension and actions $[1, \beta_2 + \Delta_2]$ in the second dimension for $\Delta_i \geq 0$ sufficiently large (of course, assuming these actions are outside the awareness set of the principal).

Private awareness. There are many situation where the principal's initial awareness may not be known to the agent. For instance, a financial expert may be uncertain about the investment options an investor has encountered before their meeting, a division manager may not know which of the feasible projects are known to the headquarter, etc. The following proposition shows that uncertainty over the principal's initial awareness does not fundamentally change the solution of the problem. For simplicity, we restrict again attention to the specification of Section 3.1.

Proposition 8. *Let utility functions be specified by (6) and condition (7) be satisfied. Furthermore, let the agent's belief about the principal's awareness set be described by a* *probability distribution over the set of closed subsets of Y^A with finite support. The set of actions that the agent optimally reveals takes the form*

$$Y^* = [y^A(0), \hat{y} - \Delta] \cup \{\hat{y} + \Delta\}$$

for some $\Delta \geq 0$. If the agent assigns a positive probability to the event that the principal's awareness set does not include $\hat{y}$, then $\Delta > 0$.

Proof. See the Online Appendix. □

In the proof of Proposition 8, we show that the agent can improve on an arbitrary awareness set by disclosing all actions that have a weakly greater distance to $\hat{y}$ than the closest action in the set. No matter what the realised awareness set of the principal is, actions that are further away from $\hat{y}$ than the closest one do not crowd out any additional actions. The optimal size of the awareness gap for the agent is determined by his beliefs about the principal's initial awareness. Proposition 8 shows that the agent leaves the principal unaware of some actions whenever he assigns a strictly positive probability to the event that the principal's awareness is bounded away from $\hat{y}$. The agent's cost of not disclosing actions around $\hat{y}$ when facing a principal who is aware of $\hat{y}$ is the (potential) loss of flexibility below $\hat{y}$. However, as we argue in Section 3, at $\Delta = 0$ the agent's marginal utility loss associated to the reduced flexibility equals zero, since in states below $\hat{y}$ the agent is at his bliss point. This implies that as long as the agent assigns a positive probability to the event that the principal is not aware of $\hat{y}$ and, hence, that there is a strict gain of introducing a gap, the net effect of marginally increasing Δ at $\Delta = 0$ is positive.

Principal's sophistication and awareness of unawareness We assumed that when additional options get revealed to the principal, she updates her awareness to the union of what she knew initially and what the agent reveals. This might suggest that the principal is naive, as she does not contemplate the possibility of other actions of which she is not aware. With regard to this point, we should note that the agent's equilibrium announcement is justifiable for the principal in the sense that it is consistent with the principal believing that the agent acts rationally. In our game, the requirement of justifiability, introduced by Ozbay (2007), states that given the principal's updated awareness Y, the principal cannot conceive of any announcement strategy which yields a higher expected payoff for the agent than Y. In other words, in equilibrium the principal should not believe that if the agent had revealed fewer actions, he would have been better off. Since the delegation set resulting from the optimal awareness set Y^* yields a weakly greater payoff for the agent than the payoff associated to any other subset $Y \subseteq Y^*$, this requirement is always satisfied in our setting.

It would however be interesting to consider a dynamic environment—much richer then ours—where the principal has ways to expand her awareness set (for example, by using a costly technology or by sampling other agents). In this case, the initial awareness of being partially aware and, perhaps more importantly, the assessment of the value of discovering new options would be key (see, for instance, Karni and Vierø (2017)). Moreover, it would be natural to assume that becoming aware can change the principal's perception about the possibility of being unaware of further actions. We leave this interesting issue for future research.

6 Conclusion

This paper formulates a flexible model of delegation with limited awareness and derives a number of properties of the optimal solution. The solution shows that by leaving the principal unaware of moderate options, the agent makes it optimal for the principal to permit actions closer to his own preferences. As argued in the Introduction, our framework allows for a number of useful applications that span from financial intermediation, human resources, and political economy. We however believe that a key component of the contribution is to provide at least two general insights that apply to games with principal-agent or sender-receiver structure where the agent/sender has superior awareness over feasible actions.

First, the paper illustrates that limited awareness can impose natural constraints on the language of contracts and that such limits may be exploited by contracting parties with superior awareness. This principle is not restricted to problems of delegation but applies to other contracting problems as well. A principal facing a privately informed agent must resolve a tradeoff between exploiting the agent's private information and limiting the agent's information rents. The distortions solving this tradeoff are optimal for the principal but not for the agent. By manipulating the principal's awareness set and hence the set of feasible contracts, the agent can increase the principal's cost of such distortions and thereby increase the principal's willingness to grant the agent higher information rents. The unconstrained solution to the agent's disclosure problem determines the maximal information rents he can get by modifying the set of feasible actions, which may provide an interesting new angle to look at principal-agent relationships.[23]

Second, our modification of the game with renegotiation exemplifies how unawareness changes the ways in which agents infer information. If a player is unaware of the set of possible signals and only becomes aware of the signal he observes, the player cannot infer information from the fact that a different signal did not realise. This asymmetry gives rise to non-standard information structures and hence to rather different equilibrium outcomes with respect to the full awareness benchmark. In the context of our game, we show that limited awareness can foster communication considerably and improve equilibrium outcomes in situations with conflicts of interest.

References

R. Alonso and N. Matouschek. Optimal delegation. *The Review of Economic Studies*, 75 (1):259–293, 2008.

M. Amador and K. Bagwell. The theory of optimal delegation with an application to tariff caps. *Econometrica*, 81(4):1541–1599, 2013.

M. Amador, I. Werning, and G.-M. Angeletos. Commitment vs. flexibility. *Econometrica*, 74(2):365–396, 2006.

M. Armstrong and J. Vickers. A model of delegated project choice. *Econometrica*, 78(1): 213–244, 2010.

S. Auster. Asymmetric awareness and moral hazard. *Games and Economic Behavior*, 82: 503–521, 2013.

S. Auster and N. Pavoni. Limited awareness and financial intermediation. mimeo, 2020.

[23]This is related to a recent literature looking at determinants of agency rents, which we mention in the Introduction. The focus of this literature is on the information structure as the modifiable primitive.

D. P. Baron and B. Holmström. The investment banking contract for new issues under asymmetric information: Delegation and the incentive problem. *The Journal of Finance*, 35(5):1115–1138, 1980.

M. X. D. Carpini and S. Keeter. *What Americans know about politics and why it matters.* Yale University Press, 1996.

V. P. Crawford and J. Sobel. Strategic information transmission. *Econometrica: Journal of the Econometric Society*, pages 1431–1451, 1982.

Y. Feinberg. Games with unawareness. *The BE Journal of Theoretical Economics*, 1 (ahead-of-print), 2020.

E. Filiz-Ozbay. Incorporating unawareness into contract theory. *Games and Economic Behavior*, 76(1):181–194, 2012.

A. Francetich and B. C. Schipper. Rationalizable screening and disclosure under unawareness. *University of California, Davis*, 2020.

D. Garrett, G. Georgiadis, A. Smolin, and B. Szentes. Optimal project design. mimeo, 2020.

S. J. Grossman and O. D. Hart. The costs and benefits of ownership: A theory of vertical and lateral integration. *Journal of Political Economy*, 94(4):691–719, 1986.

P. Guarino. An epistemic analysis of dynamic games with unawareness. *Games and Economic Behavior*, 120:257–288, 2020.

L. Guiso and T. Jappelli. Awareness and stock market participation. *Review of Finance*, 9(4):537–567, 2005.

M. Halac and P. Yared. Commitment vs. flexibility with costly verification. *Journal of Political Economy*, 128(12), 2020.

J. Y. Halpern and L. C. Rêgo. Extensive games with possibly unaware players. *Mathematical Social Sciences*, 70:42–58, 2014.

A. Heifetz, M. Meier, and B. C. Schipper. Dynamic unawareness and rationalizable behavior. *Games and Economic Behavior*, 81:50–68, 2013.

A. Heifetz, M. Meier, and B. C. Schipper. Prudent rationalizability in generalized extensive-form games with unawareness. *The BE Journal of Theoretical Economics*, 1 (ahead-of-print), 2020.

B. Holmström. On incentives and control in organisations, chapter ii: Delegation. *unpublished PhD dissertation*, 1977.

E. Karni and M.-L. Vierø. "reverse bayesianism": A choice-based theory of growing awareness. *American Economic Review*, 103(7):2790–2810, 2013.

E. Karni and M.-L. Vierø. Awareness of unawareness: A theory of decision making in the face of ignorance. *Journal of Economic Theory*, 168:301–328, 2017.

N. Kartik, R. Van Weelden, and S. Wolton. Electoral ambiguity and political representation. *American Journal of Political Science*, 61(4):958–970, 2017.

F. Koessler and D. Martimort. Optimal delegation with multi-dimensional decisions. *Journal of Economic Theory*, 147(5):1850–1881, 2012.

E. Kováč and T. Mylovanov. Stochastic mechanisms in settings without monetary transfers: The regular case. *Journal of Economic Theory*, 144(4):1373–1395, 2009.

A. Lei and X. Zhao. Delegation and unawareness. mimeo, 2020.

D. Martimort and A. Semenov. Continuity in mechanism design without transfers. *Economics Letters*, 93(2):182–189, 2006.

M. Meier and B. C. Schipper. Bayesian games with unawareness and unawareness perfection. *Economic Theory*, 56(2):219–249, 2014.

N. D. Melumad and T. Shibano. Communication in settings with no transfers. *The RAND Journal of Economics*, pages 173–198, 1991.

R. C. Merton. A simple model of capital market equilibrium with incomplete information. *The Journal of Finance*, XLII(3):483–510, 1987.

E. Y. Ozbay. Unawareness and strategic announcements in games with uncertainty. In *Proceedings of the 11th conference on Theoretical aspects of rationality and knowledge*, pages 231–238, 2007.

A.-K. Roesler and B. Szentes. Buyer-optimal learning and monopoly pricing. *American Economic Review*, 107(7):2072–80, 2017.

I. Somin. *Democracy and political ignorance: Why smaller government is smarter*. Stanford University Press, 2016.

D. Szalay. The economics of clear advice and extreme options. *The Review of Economic Studies*, 72(4):1173–1198, 2005.

E.-L. Von Thadden and X. Zhao. Incentives for unaware agents. *The Review of Economic Studies*, 79(3):1151–1174, 2012.

X. Zhao. Framing contingencies in contracts. *Mathematical Social Sciences*, 61(1):31–40, 2011.

Appendix

A Proofs

A.1 Proof of Proposition 1

Proof. We start by introducing two functions. First, define $s := (y^A)^{-1}$ as the inverse of the agent's bliss point function y^A; that is, $s(y)$ *represents the state in which* y *is the most preferred action for the agent.* Second, let $t : (Y^A)^2 \to [0,1]$ be a symmetric function, indicating *the state at which the agent is indifferent between any two action* y *and* y'. It is specified as follows. For $y = y'$, set $t(y, y') = s(y)$. For $y < y'$, $t(y, y')$ is defined by:

- if $U^A(\theta, y) < U^A(\theta, y')$ for all $\theta \in [0,1]$, then $t(y, y') = 0$;
- if $U^A(\theta, y) > U^A(\theta, y')$ for all $\theta \in [0,1]$, then $t(y, y') = 1$;
- otherwise $t(y, y')$ is such that

$$U^A(t(y, y'), y) = U^A(t(y, y'), y'). \tag{13}$$

Due to the single-crossing condition, the solution of (13) is unique. For $y > y'$, $t(y, y')$ is pinned down by the symmetry condition $t(y, y') = t(y', y)$. The following lemma links the slope of s with a partial derivative of t.

Lemma 9. *Consider* y_0 *such that* $s(y_0) \in (0,1)$*, then*

$$\lim_{y \to y_0} \frac{dt(y, y_0)}{dy} = \frac{1}{2} s'(y_0)$$

Proof. For the case where t is determined by (13), we apply the implicit function theorem to derive

$$\frac{\mathrm{d}t(y, y_0)}{\mathrm{d}y} = \frac{U_y^A(t(y, y_0), y)}{U_\theta^A(t(y, y_0), y_0) - U_\theta^A(t(y, y_0), y)}$$

Taking the limit, we have

$$\begin{aligned}
\lim_{y \to y_0} \frac{\mathrm{d}t(y, y_0)}{\mathrm{d}y} &= \lim_{y \to y_0} \frac{U_y^A(t(y, y_0), y)}{U_\theta^A(t(y, y_0), y_0) - U_\theta^A(t(y, y_0), y)} \\
&= \lim_{y \to y_0} \frac{U_{\theta y}^A(t(y, y_0), y) \frac{\mathrm{d}t(y, y_0)}{\mathrm{d}y} + U_{yy}^A(t(y, y_0), y)}{\left(U_{\theta\theta}^A(t(y, y_0), y_0) - U_{\theta\theta}^A(t(y, y_0), y)\right) \frac{\mathrm{d}t(y, y_0)}{\mathrm{d}y} - U_{\theta y}^A(t(y, y_0), y)} \\
&= \frac{U_{\theta y}^A(s(y_0), y_0) \lim_{y \to y_0} \frac{\mathrm{d}t(y, y_0)}{\mathrm{d}y} + U_{yy}^A(s(y_0), y_0)}{-U_{\theta y}^A(s(y_0), y_0)}
\end{aligned}$$

where the second equality follows from L'Hôspital's rule. Also recall $t(y_0, y_0) = s(y_0)$. We can solve the above equality for $\lim_{y \to y_0} \frac{\mathrm{d}t(y,y_0)}{\mathrm{d}y}$ and obtain

$$\lim_{y \to y_0} \frac{\mathrm{d}t(y, y_0)}{\mathrm{d}y} = -\frac{1}{2} \cdot \frac{U^A_{yy}(s(y_0), y_0)}{U^A_{y_0}(s(y_0), y_0)}$$

From $U^A_y(s(y), y) = 0$, we derive via the implicit function theorem:

$$s'(y) = -\frac{U^A_{yy}(s(y), y)}{U^A_{\theta y}(s(y), y)}$$

The two results together establish the claim. $\square$

Define $\bar{D}(y) := D^*(Y^A) \cap [y^A(0), y]$ as the set obtained by capping the optimal delegation set under full awareness at y.

Lemma 10. *Generically, there exists some* $\underline{y} < \hat{y}$ *such that for all* $y \in (\underline{y}, \hat{y})$,

$$V^P\left(\bar{D}(y)\right) < V^P\left(\bar{D}(y) \cup \{\hat{y}\}\right)$$

Proof. Under full awareness, the upper bound of the principal's optimal delegation is $\hat{y}$. Under the assumption that $\hat{y}$ is a limit point $D^*(Y^A)$, $\hat{y}$ must maximise

$$\int_{\underline{y}}^{s(y)} U^P(\theta, y^A(\theta)) dF(\theta) + \int_{s(y)}^{1} U^P(\theta, y) dF(\theta) \tag{14}$$

over y for some $\underline{y} < \hat{y}$. The first-order condition is

$$\int_{s(\hat{y})}^{1} U^P_y(\theta, \hat{y}) dF(\theta) = 0 \tag{15}$$

and the second-order condition is

$$\int_{s(\hat{y})}^{1} U^P_{yy}(\theta, \hat{y}) dF(\theta) - U^P_y(s(\hat{y}), \hat{y}) f(s(\hat{y})) s'(\hat{y}) \leq 0. \tag{16}$$

Since s is fully determined by U^A and thus exogenous and since $W'(y) := \int_{s(y)}^1 U^P_y(\theta, y) dF(\theta)$ is strictly decreasing on a neighbourhood around $\hat{y}$ (recall that $V^P(D^*(Y^A))$ is a strict maximum), *generically* this condition holds as a strict inequality.

Define the difference between the principal's expected payoffs under delegation sets $\bar{D}(y) \cup \{\hat{y}\}$ and $\bar{D}(y)$:

$$\Delta W(y) := V^P\left(\bar{D}(y) \cup \{\hat{y}\}\right) - V^P\left(\bar{D}(y)\right) = \int_{t(y,\hat{y})}^{1} U^P(\theta, \hat{y}) dF(\theta) - \int_{t(y,\hat{y})}^{1} U^P(\theta, y) dF(\theta).$$

We calculate the first derivative and evaluate it at $\hat{y}$:

$$\begin{aligned}\Delta W'(y) &= -\int_{t(y,\hat{y})}^{1} U_y^P(\theta,y)dF(\theta) - \left(U^P(t(y,\hat{y}),\hat{y}) - U^P(t(y,\hat{y}),y)\right) f(t(y,\hat{y}))\frac{d\,t(y,\hat{y})}{d\,y} \\ \Delta W'(\hat{y}) &= -\int_{s(\hat{y})}^{1} U_y^P(\theta,\hat{y})dF(\theta).\end{aligned}$$

By (15), the above term is equal to zero. We must therefore consider the second derivative:

$$\begin{aligned}\Delta W''(y) &= -\int_{t(y,\hat{y})}^{1} U_{yy}^P(\theta,y)dF(\theta) + 2U_y^P(t(y,\hat{y}),y)f(t(y,\hat{y}))\frac{d\,t(y,\hat{y})}{d\,y} \\ &\quad - \left(U_\theta^P(t(y,\hat{y}),\hat{y}) - U_\theta^P(t(y,\hat{y}),y)\right) f(t(y,\hat{y}))\left(\frac{d\,t(y,\hat{y})}{d\,y}\right)^2 \\ &\quad - \left(U^P(t(y,\hat{y}),\hat{y}) - U^P(t(y,\hat{y}),y)\right)\left(f'(t(y,\hat{y}))\frac{d\,t(y,\hat{y})}{d\,y} + f(t(y,\hat{y}))\frac{d^2\,t(y,\hat{y})}{d\,y^2}\right) \\ \Delta W''(\hat{y}) &= -\int_{s(\hat{y})}^{1} U_{yy}^P(\theta,\hat{y})dF(\theta) + 2U_y^P(s(\hat{y}),\hat{y})f(s(\hat{y}))\left.\frac{d\,t(y,\hat{y})}{d\,y}\right|_{y=\hat{y}}\end{aligned}$$

Since, by Lemma 9, we have $-\left.\frac{d\,t(y,\hat{y})}{d\,y}\right|_{\varepsilon=0} = \frac{1}{2}s'(\hat{y})$, condition (16) holding as a strict inequality implies $\Delta W''(\hat{y}) > 0$. Remembering $\Delta W'(\hat{y}) = 0$, there is then an interval for y to the left of $\hat{y}$, where $\Delta W'(y) < 0$. With $\Delta W(\hat{y}) = 0$, this property implies, in turn, that there is an interval for y to the left of $\hat{y}$ such that $\Delta W(y) > 0$. Hence, for y sufficiently close to $\hat{y}$, the inequality $V^P\left(\bar{D}(y)\right) < V^P\left(\bar{D}(y) \cup \{\hat{y}\}\right)$ holds. □

Perturbation. Lemma 10 shows that, generically, there exists some $\underline{y} < \hat{y}$, such that for all $y \in (\underline{y}, \hat{y}]$, $V^P(\bar{D}(y)) < V^P(\bar{D}(\bar{y}) \cup \{\hat{y}\})$ is satisfied. Hence, for each $y \in (\underline{y}, \hat{y}]$ there exists some $\delta_y > 0$ such that for all $\delta \in [0, \delta_y)$, $V^P(\bar{D}(y)) \leq V^P(\bar{D}(y) \cup \{\hat{y} + \delta\})$ holds.[24] Due to this property, we can find a continuous, strictly decreasing function which maps each $y \in (\underline{y}, \hat{y}]$ to a value of δ satisfying this condition. Letting $y(\cdot)$ denote the inverse of this function, we define the associated awareness set by

$$Y(\delta) := Y^A \backslash (y(\delta), \hat{y} + \delta).$$

Principal Optimality. In the last stage of the game, the agent's best response is described by $y^*(\theta, D)$. Given that the agent chooses according to y^*, the principal with awareness $Y \in \mathcal{Y}$ optimally selects a delegation set $D \subseteq Y$ to solve (3). Holmström (1977, Theorem 1) provides conditions under which the principal's value function V^P is upper semi-continuous. These conditions are satisfied in our framework. In particular:

- The set $Y^A = [y_{min}, y_{max}]$ is a compact subset of $\mathbb{R}$, a complete, separable metric space.

[24] Since the agent simply chooses the better of the two actions closest to him, it is immediate to see that $V^P(\bar{D}(\bar{y}) \cup \{\hat{y}\})$ is continuous in $\hat{y}$ (see expression (14)).

- $\mathcal{D}(Y^A)$ is a closed subset of $2^{[Y^A]}$ with respect to the Hausdorff-metric

$$d_H(D, D') = \max\left\{\sup_{y\in D}\inf_{y'\in D'} d(y, y'), \sup_{y'\in D'}\inf_{y\in D} d(y, y')\right\}.$$

- U^A and U^P are uniformly bounded on their domains $[0, 1] \times [y_{min}, y_{max}]$.

Let $y_1(\delta) = \max\{\tilde{y} \in D^*(Y(\delta)) : \tilde{y} \leq \hat{y}\}$ with $y_1(0) = \hat{y}$. We want to show that $y_1(\cdot)$ is continuous in δ on a right neighbourhood of zero. Suppose this is not true. Then, since $y_1(\cdot)$ is bounded above by $\hat{y}$, there exists a (sub)sequence $\{\delta_n\}$ with $\lim_{n\to+\infty} \delta_n = 0$ such that $\lim_{n\to+\infty} y_1(\delta_n) = y_0 \leq \hat{y}$, with y_0 possibly depending on the sequence. To violate continuity it must be that one of them satisfies $y_0 < \hat{y}$. Denote this sequence by $\{\bar{\delta}_n\}$. Then,

$$d_H(D\backslash(y_0, \hat{y}), D^*(Y^A)) \geq \frac{\hat{y} - y_0}{2}.$$

By upper semicontinuity and uniqueness of the solution of (3), for all D and n sufficiently large, $V^P(\bar{D}(y(\bar{\delta}_n))) > V^P(D\backslash(y_0, \hat{y}))$ is satisfied, which is a contradiction to $y_1(\bar{\delta}_n) \in D^*(Y(\bar{\delta}_n))$. Hence, $y_1(\cdot)$ is continuous on a right neighbourhood of 0.

The principal's optimisation regarding the inclusion of actions below $y_1(\delta)$ is equivalent to that under full awareness, as their potential inclusion only affects the agent's choice in states below $s(y_1(\delta))$. Indeed, given delegation set D and state θ, the agent has to consider at most two actions, which are the points in D on the left and right from his preferred action $y^A(\theta)$. Conditional on $y_1(\delta)$ belonging to the delegation set, the principal's design problem for actions below $y_1(\delta)$ can thus be separated from that for actions above. Hence, the optimal delegation set under awareness $Y(\delta)$ satsifies $\bar{D}(y_1(\delta)) \subseteq D^*(Y(\delta))$.

Next, we show that the principal permits at least one action above $\hat{y}$. To see this notice that for all D with $\max D \leq y(\delta)$, the following inequalities hold:

$$V^P(D) \leq V(\bar{D}(y(\delta))) \leq V^P(\bar{D}(y(\delta)) \cup \{\hat{y} + \delta\}).$$

The first inequality follows from the facts that 1) generating a payoff close to $V^P(D^*(Y^A))$ requires that $\max D$ is close to $\hat{y}$, 2) conditional on $\max D$ being close to $\hat{y}$, we have $V^P(D) \leq V^P(\bar{D}(\max D))$, and 3) $V^P(\bar{D}(y))$ is strictly increasing on a left neighbourhood of $\hat{y}$. We thus established $\max D^*(Y(\delta)) > \hat{y}$.

With this observation, we can define $y_2(\delta) := \max\{\tilde{y} \in D^*(Y(\delta)) : \tilde{y} > \hat{y}\}$. By an analogous argument to the one above y_2 is continuous in δ on a right neighbourhood of zero. Since the function y_2 is bounded below by $\hat{y} + \delta$ and satisfies $y_2(0) = \hat{y}$, it must be increasing on a right neighbourhood of 0. $D^*(Y(\delta))$ thus satisfies

$$\bar{D}(y_1(\delta)) \cup \{y_2(\delta)\} \subseteq D^*(Y(\delta)). \tag{17}$$

Agent optimality. Let V^A denote the agent's value as a function of the delegation set. Property (17) implies that $V^A(\bar{D}(y_1(\delta)) \cup \{y_2(\delta)\})$ constitutes a lower bound for the payoff the agent obtains when the principal's awareness set is $Y(\delta)$: additional actions in $D^*(Y(\delta))$ can only benefit the agent. For ease of notation, we change variables and write $y^+ = y_2(\delta)$ and $y^-(y^+) = y_1(y_2^{-1}(y^+))$. The agent's expected payoff for the delegation set $D(y^+) := \bar{D}(y^-(y^+)) \cup \{y^+\}$ can then be written as:

$$\begin{aligned} V^A(D(y^+)) &= \int_0^{s(y^-(y^+))} U^A(\theta, y^*(\theta, D^*(Y^A)))\mathrm{d}F(\theta) + \int_{s(y^-(y^+))}^{t(y^-(y^+),y^+)} U^A(\theta, y^-(y^+))\mathrm{d}F(\theta) \\ &\quad + \int_{t(y^-(y^+),y^+)}^1 U^A(\theta, y^+)\mathrm{d}F(\theta). \end{aligned}$$

The first derivative of this payoff with respect to y^+ is

$$\frac{\mathrm{d}\, V^A(D(y^+))}{\mathrm{d}\, y^+} = \int_{s(y^-(y^+))}^{t(y^-(y^+),y^+)} U_y^A(\theta, y^-(y^+))y'(\varepsilon)\mathrm{d}F(\theta) + \int_{t(y^-(y^+),y^+)}^1 U_y^A(\theta, y^+)\mathrm{d}F(\theta).$$

Evaluated at $y^+ = \hat{y}$, this derivative is equal to:

$$\left.\frac{\mathrm{d}\, V^A(D(y^+))}{\mathrm{d}\, y^+}\right|_{y^+=\hat{y}} = \int_{s(\hat{y})}^1 U_y^A(\theta, \hat{y})\mathrm{d}F(\theta).$$

Since $U_y^A(s(\hat{y}), \hat{y}) = 0$ and $U_{\theta y}^A > 0$, we have $U^A(\theta, y) > 0$ for all $\theta > s(\hat{y})$. The derivative of the agent's value at $y^+ = \hat{y}$ is thus strictly positive.

Taken toegther, there exists an $y^+ > \hat{y}$ and an associated awareness set that yields a delegation set which the agent strictly prefers to $D^*(Y^A)$. Hence, revealing all actions in Y^A is strictly dominated for the agent. □

A.2 Proof of Proposition 2

Proof. First, we prove statement (i). Let $\hat{y} \notin Y^P$. Since Assumption 1 is satisfied, the optimal delegation set under full awareness takes the form $D^*(Y^A) = [y^A(0), \hat{y}]$. Hence, $\hat{y}$ is a limit point of $D^*(Y^A)$ and, by Proposition 1, the agent does not disclose all actions. Next, let $\hat{y} \in Y^P$. We want to show that full disclosure is optimal for the agent. Towards a contradiction, suppose this is not true. Then there exists an awareness set Y with $\hat{y} \in Y$ such that the agent strictly prefers $D^*(Y)$ over $[y^A(0), \hat{y}]$. This implies that $D^*(Y)$ contains a non-empty set of actions $\tilde{X}$ such that $x > \hat{y}$ for all $x \in \tilde{X}$. Since $D^*(Y^A)$ is the largest optimal delegation set in $\mathcal{D}(Y^A)$,

$$V^P(D^*(Y^A)) > V^P(D^*(Y^A) \cup \tilde{X})$$

holds. Monotonicity of the agent's policy then implies that conditional on permitting action $\hat{y}$, the principal is strictly better off by removing all actions in $\tilde{X}$. Hence, $\hat{y}$ cannot belong to

$D^*(Y)$. Since restricting the agent's choice from below is never optimal, we have $\hat{y} > \min D^*(Y)$ and hence $\hat{y} \in Conv(D^*(Y))$. Assumption 1 then implies $V^P(D^*(Y) \cup \{\hat{y}\}) \geq V^P(D^*(Y))$, a contradiction. Disclosing all actions is thus optimal for the agent.

To prove statement (ii), notice first that if $\hat{y} \in Y^P$, then $\Delta_1 = \Delta_2 = 0$. Consider then the case $\hat{y} \notin Y^P$. We start by showing that the principal permits at most one action weakly greater than $\hat{y}$. Suppose instead there is an awareness set Y given which the principal optimally delegates a set $D^*(Y)$ which contains two distinct actions weakly greater than $\hat{y}$ and let $\bar{y}$ be the largest action in $D^*(Y)$. Given that (5) is single-peaked, we know that for any $y \in (\hat{y}, \bar{y})$ we have $V^P([y^A(0), y]) > V^P([y^A(0), \bar{y}])$ and hence:

$$\int_{s(y)}^{1} U^P(\theta, y)\,\mathrm{d}F(\theta) > \int_{s(y)}^{s(\bar{y})} U^P(\theta, y^A(\theta))\,\mathrm{d}F(\theta) + \int_{s(\bar{y})}^{1} U^P(\theta, \bar{y})\,\mathrm{d}F(\theta). \tag{18}$$

This inequality, together with the assumption that permitting $\bar{y}$ is optimal for the principal, implies that $\bar{y}$ cannot be a limit point of $D^*(Y)$. Consider then action $y > \hat{y}$ such that $y < \bar{y}$ and $[y, \bar{y}] \cap D^*(Y) = \{y, \bar{y}\}$. We can show the following:

$$\begin{aligned}
& V^P(D^*(Y)) \\
= & \int_0^{s(y)} U^P(\theta, y^*(\theta, D^*(Y)))\,\mathrm{d}F(\theta) + \int_{s(y)}^{t(y,\bar{y})} U^P(\theta, y)\,\mathrm{d}F(\theta) + \int_{t(y,\bar{y})}^{1} U^P(\theta, \bar{y})\,\mathrm{d}F(\theta) \\
\leq & \int_0^{s(y)} U^P(\theta, y^*(\theta, D^*(Y)))\,\mathrm{d}F(\theta) + \int_{s(y)}^{s(\bar{y})} U^P(\theta, y^A(\theta))\,\mathrm{d}F(\theta) + \int_{s(\bar{y})}^{1} U^P(\theta, \bar{y})\,\mathrm{d}F(\theta) \\
< & \int_0^{s(y)} U^P(\theta, y^*(\theta, D^*(Y)))\,\mathrm{d}F(\theta) + \int_{s(y)}^{1} U^P(\theta, y)\,\mathrm{d}F(\theta) \\
= & V^P(D^*(Y) \backslash \{\bar{y}\}),
\end{aligned}$$

where the weak inequality follows from Assumption 1 and the strict inequality follows from (18). Taken together, these inequalities imply that the principal can strictly improve her payoff by removing action $\bar{y}$ from the delegation set, which yields the contradiction.

Consider now an optimal awareness set Y^* with $\bar{y} = \max D^*(Y^*)$. The set Y^* must clearly satisfy $\max D^*(Y^*) > \hat{y}$, as any delegation set with an upper bound weakly smaller than $\hat{y}$ is dominated by the full awareness delegation set and $D^*(Y^*) \neq D^*(Y^A)$. We set $\Delta_2 = \max D^*(Y^*) - \hat{y} > 0$. We know $[\hat{y}, \hat{y} + \Delta_2) \cap D^*(Y^*) = \emptyset$ because we have shown that the principal allows at most one action above $\hat{y}$. Assumption 1 and the fact that the agent is upward biased imply that the principal permits all actions in Y^* that are weakly smaller than

$\hat{y} + \Delta_2$. Define $\Delta_1 \geq 0$ as the smallest value for Δ that satisfies the inequality

$$\int_{s(\hat{y}-\Delta)}^{1} U^P(\theta, \hat{y} - \Delta)\, \mathrm{d}F(\theta)$$
$$\leq \int_{s(\hat{y}-\Delta)}^{t(\hat{y}-\Delta,\hat{y}+\Delta_2)} U^P(\theta, \hat{y} - \Delta)\, \mathrm{d}F(\theta) + \int_{t(\hat{y}-\Delta,\hat{y}+\Delta_2)}^{1} U^P(\theta, \hat{y} + \Delta_2)\, \mathrm{d}F(\theta).$$

By continuity of U^P, t and s (due to the continuity of U^A and recalling our tie-breaking rule), Δ_1 is well-defined. For a fixed Δ_2, this inequality is satisfied whenever, conditional on permitting action $\hat{y} - \Delta$, the principal prefers also permitting $\hat{y} + \Delta_2$ over capping the agent's choice at $\hat{y} - \Delta$. Since $\Delta_2 > 0$, also Δ_1 is positive.

By definition of Δ_1, Assumption 1 implies that revealing actions in the interval $(\hat{y} - \Delta_1, \hat{y})$, would make it strictly optimal for the principal to exclude ('crowd out') the action $\hat{y}+\Delta_2$; hence $Y^* \cap (\hat{y} - \Delta_1, \hat{y}) = \emptyset$. Assumption 1 together with the fact that the agent is upward biased, implies that, conditional on $\hat{y} - \Delta_1 \in D^*(Y^*)$, revealing any action below $\hat{y} - \Delta_1$ and weakly above $y^A(0)$ results in the inclusion of that action in the delegation set and thus strictly benefits the agent. Optimality of Y^* thus requires $Y^* \cap [y^A(0), \hat{y}+\Delta_2] = [y^A(0), \hat{y}-\Delta_1] \cup \{\hat{y}+\Delta_2\}$. Given such awareness set, the principal optimally chooses $D^*(Y^*) = [y^A(0), \hat{y} - \Delta_1] \cup \{\hat{y} + \Delta_2\}$. $\square$

A.3 The Game with Renegotiation

An equilibrium for the new extensive form game will be defined as an awareness choice Y at the root for the agent, a delegation set $D \subseteq Y$ in the contracting phase for the principal and a PBE at the renegotiation stage given (Y, D).

Suppose that after the contracting phase the principal's awareness is Y and the optimal delegation set is $D \subseteq Y$. The game played at the renegotiation phase is a signalling game. We now define the strategies of the principal and the agent at this stage. We will first of all concentrate on agent's moves constituted by either 'no new proposal' (let us call it N) or proposal of singletons y. Let $X := N \cup Y^A$ be the set of possible proposals of the agent. The N proposal will be interpreted as if the agent does not propose any action outside the original delegation set D, and hence, in particular, he does not increase the awareness of the principal. When the agent proposes a new action $x \in X$ the principal can decide whether to replace the original delegation set D with x.

It is clear that the agent will never benefit from proposals $x \in D$ as these are weakly dominated by the 'null' or 'no proposal' $x = N$. It will also be clear below that the agent will not be able to increase his payoff in the best equilibrium by augmenting the proposal/signal y with an additional payoff-irrelevant message. We hence - to simplify notation - do not allow for any further message at the renegotiation stage.

In Appendix C.3 we show that, given initial history (Y, D), that the component of agent's strategy at this stage can be described by a map from each realisation of $\theta \in [0, 1]$ into a recommendation $x \in X$ and then a map from the new delegation set D' into an action $y \in D'$.

At this stage, the agent's strategy is thus a pair of functions $x(Y,D,\cdot), y(Y,D,\cdot,\cdot)$, where $x(Y,D,\theta)$ is the proposal given (Y,D,θ), and $y(Y,D,\theta,D')$ is the final choice given (Y,D,θ,D'). The principal's strategy is a mapping $\rho(Y,D,\cdot) \in \{0,1\}$, where $\rho(Y,D,x)=0$ implies $D'=D$ while $\rho(Y,D,x)=1$ implies $D'=\{x\}$.[25] With her action, the principal assigns to each proposal a delegation set which can be either the original D or the singleton $\{x\}$. If the agent proposes $x(Y,D,\cdot)=N$, the principal can only choose $\rho(Y,D,N)=0$.

We also concentrate on outcomes in pure strategies. We can hence define a beliefs' support function Θ for the principal mapping each triplet (Y,D,x) to a subset of values for θ. The set $\Theta(Y,D,x) \subseteq [0,1]$ describes the states which the principal considers possible when the agent proposes x when the principal has awareness level Y and decided D. For completeness, we set $\Theta(Y,D,N)=[0,1]$.

Definition 1. *Given, (Y,D), with $D \subseteq Y$, the strategy profile (x^*,y^*,ρ^*) and beliefs' support function Θ^* constitute a PBE of the renegotiation game if and only if the following conditions hold:*[26]

- *Principal optimality: for all $x \in X \setminus N$,*

$$\rho^*(Y,D,x) \in \arg\max_{\rho\in\{0,1\}} \rho\,\mathbb{E}[U^P(\theta,x)\,|\theta \in \Theta^*(Y,D,x)]+(1-\rho)\,\mathbb{E}[U^P(\theta,y^*(Y,D,\theta,D))\,|\theta\in\Theta^*(Y,D,x)];$$

- *Agent optimality: for all $\theta \in [0,1]$,*

$$\begin{aligned} x^*(Y,D,\theta) &\in \arg\max_x \rho^*(Y,D,x)U^A(\theta,x) + (1-\rho^*(Y,D,x))U^A(\theta,y^*(Y,D,\theta,D))\,; \\ y^*(Y,D,\theta,D) &\in \arg\max_{y\in D} U^A(\theta,y); \end{aligned}$$

- *Consistency of beliefs: for all (Y,D,x) that are allowed on path,*

$$\Theta^*(Y,D,x)=\{\theta\in[0,1]: U^A(\theta,x) \geq \max_{x'\in Y} \rho^*(Y,D,x')U^A\left(\theta,x'\right) + (1-\rho^*(Y,D,x'))U^A(\theta,y^*(Y,D,\theta,D))\}.$$

The consistency condition assures that the principal's beliefs are coherent with the agent's strategy, as perceived through the principal's awareness (see that the max is taken only over Y). According to the condition, the principal evaluates payoffs for lower levels of awareness using her own strategy. Hence, with respect to lower levels of awareness that are reached on path, the principal's beliefs are correct.

Finally, note that in the last stage, for each (Y,D), the function $y^*(Y,D,\cdot,\cdot)$ coincides with the function y^* we defined after condition (1), hence, its use in the main text (abusing notation). Recall we assumed that the principal is unaware of the renegotiation possibility.

A PBE equilibrium for the extensive-form game is constituted by an initial awareness choice Y^*, the delegation function $D^*(\cdot)$ as described in the 'pure' delegation game of the previous

[25] Of course, we could have equivalently defined the feasibility set for ρ as the pair $\{D,x\}$.
[26] All expectations are taken with respect to F.

section together with a PBE equilibrium as defined in Definition 1 for each (Y, D)*, with* $D \subseteq Y$*.* This notion is strong in the sense that it imposes PBE also for subtrees not reached along the equilibrium path.

Proposition 15 in the online appendix shows the substantial equivalence between the equilibrium outcomes according to this definition and a more complete definition of equilibrium for extensive form games with unawareness we provide there.

A.4 Proof of Proposition 5

Proof. Fix an equilibrium and let Y denote the principal's awareness set after the contracting phase with associated delegation set $D^*(Y)$ (for a detailed description of strategies and beliefs see Section A.3). Consider the renegotiation phase and let S denote the set of actions that the agent will be able to take in the last stage for some contingency and that will actually be taken by the agent under some contingency. Since the agent always has the option not to renegotiate, we have $D^*(Y) \subseteq S$. We write $\bar{X} := S \setminus D^*(Y)$ and assume that $\bar{X}$ measurable.

We first want to show $\bar{X} \cap Y = \emptyset$. Towards a contradiction, suppose this intersection is nonempty and call the non-empty set $\hat{X} = \bar{X} \cap Y = (S \cap Y) \setminus D^*(Y)$. Consider the principal's subjective game after receiving a proposal $x' \in \hat{X}$. The principal is now aware that the agent can renegotiate and therefore of all proposals in Y. By definition of S, since $\hat{X}$ is disjoint from $D^*(Y)$, we have $\rho^*(Y, D^*(Y), x) = 1$ for all $x \in \hat{X}$ and $\rho^*(Y, D^*(Y), x) = 0$ for all $x \in Y \setminus S$. Consistency of beliefs implies

$$\Theta^*(Y, D^*(Y), x) = \{\theta \in [0,1] : U^A(\theta, x) \geq \sup_{x' \in S \cap Y} U^A(\theta, x')\} \text{ for all } x \in \hat{X}.$$

Given Θ^*, principal optimality requires for all $x \in \hat{X}$[27]

$$\int_{\Theta^*(Y, D^*(Y), x)} U^P(\theta, x)\, \mathrm{d}F(\theta) \geq \int_{\Theta^*(Y, D^*(Y), x)} U^P(\theta, y^*(\theta, D^*(Y)))\mathrm{d}F(\theta) \tag{19}$$

Summing (19) over all $x \in \hat{X}$, we obtain:

$$\int_{\hat{X}} \int_{\Theta^*(Y, D^*(Y), x)} U^P(\theta, x)\, \mathrm{d}F(\theta) dx \geq \int_{\hat{X}} \int_{\Theta^*(Y, D^*(y), x)} U^P(\theta, y^*(\theta, D^*(Y)))\mathrm{d}F(\theta) dx \tag{20}$$

It is easy to show that the properties of U^A imply $\{\Theta^*(Y, D^*(Y), x)\}_{x \in \hat{X}}$ is a collection sets with measure zero intersection. Letting $\Theta^C(Y, D^*(Y), \bar{X}) := [0,1] \setminus \left(\bigcup_{x \in \hat{X}} \Theta^*(Y, D^*(Y), x)\right)$,

[27] Note that the set $\Theta^*(Y, D^*(Y), x)$ is measurable for all $x \in \hat{X}$, as both S and Y are measurable and U^A is continuous.

(20) can be written as

$$\begin{aligned}
&\int_{\Theta^C(Y,D^*(Y),\hat{X})} U^P(\theta, y^*(\theta, D^*(Y)))\mathrm{d}F(\theta) + \int_{\hat{X}}\int_{\Theta^*(Y,D^*(Y),x)} U^P(\theta,x)\,\mathrm{d}F(\theta)dx \\
\geq\ &\int_{\Theta^C(Y,D^*(Y),\hat{X})} U^P(\theta, y^*(\theta, D^*(Y)))\mathrm{d}F(\theta) + \int_{\hat{X}}\int_{\Theta^*(Y,D^*(y),x)} U^P(\theta, y^*(\theta, D^*(Y)))\mathrm{d}F(\theta)dx
\end{aligned}$$

or equivalently

$$V^P(D^*(Y) \cup \hat{X}) \geq V^P(D^*(Y)).$$

Since $\hat{X}$ is a subset of Y and D^* is the largest optimal awareness set with respect to Y that includes actions that will actually be taken by the agent under some contingency, this inequality yields a contradiction. Hence, $S \cap Y = D^*(Y)$.

Having shown that the principal only accepts proposals in the renegotiation phase if they do not belong to Y, consider proposal $x \in S \setminus Y$. Consistency of beliefs requires

$$\Theta^*(Y, D^*(Y), x) = \{\theta \in [0,1] : U^A(\theta, x) \geq \max_{y \in D^*(Y)} U^A(\theta, y)\}. \tag{21}$$

Principal optimality is then satisfied if and only if (19) holds (with Θ^* defined by (21)), or equivalently:

$$V^P(D^*(Y) \cup \{x\}) \geq V^P(D^*(Y)).$$

Finally, for each set A constituted of actions x satisfying (11) we can construct an equilibrium of the renegotiation game where proposal x is accepted by the principal whenever $x \in A$. To this end, we set $\rho^*(Y, D^*(Y), x) = 1$ for all $x \in A$ and $\rho^*(Y, D^*(Y), x) = 0$ otherwise, $x^*(Y, D^*(Y), \theta) = \arg\max_{x \in A} U^A(\theta, x)$ for all θ, and we define $\Theta^*(Y, D^*(Y), x)$ by (21) for all $x \in A$. Off-path beliefs can be set arbitrarily such that the principal rejects proposals outside A. This strategy and belief profile clearly satisfies principal optimality, agent optimality and consistency of beliefs and thus constitutes a PBE of the renegotiation game, as specified in Definition 1 of Section A.3. □

A.5 Proof of Proposition 6

Proof. We want to show that in the agent's best equilibrium the set of actions which the agent can implement with and without renegotiation is given by $S^* := [y^A(0), y^{max}]$. By definition of y^{max} and Proposition 5, there is no equilibrium where the principal accepts a proposal $y > y^{max}$. To see this, let S be the set of implementable actions with $\bar{y} = \max S$ and assume $\bar{y} > y^{max}$. By definition of y^{max}, there is no awareness set Y such that $Y^P \subseteq Y$ and $\bar{y} \in D^*(Y)$. Hence, $\bar{y}$ must be proposed in the renegotiation phase. By Proposition 5, $\bar{y} \in S$ implies that the awareness set

Y after the contracting phase is such that

$$V^P(D^*(Y) \cup \{\bar{y}\}) \geq V^P(D^*(Y)).$$

This inequality implies $\bar{y} \in D^*(Y \cup \{\tilde{y}\})$ and hence $Y^P \not\subseteq Y$, a contradiction. All we need to show is then that there exists an equilibrium where the set of the agent's implementable actions is S^*.

Recall the strategies we defined above and consider the following candidate equilibrium. In the contracting phase (the root of the game) the agent reveals a set Y such that $\max D^*(Y) = y^{max}$. Let this set be denoted by Y^{max}. The principal is unaware of the renegotiation stage and thus offers a delegation set that is optimal without renegotiation, $D^*(Y^{max})$. By Proposition 5, there is then an equilibrium where each x satisfying

$$V^P(D^*(Y^{max}), x) \geq V^P(D^*(Y^{max})) \tag{22}$$

is accepted in the renegotiation stage. For instance, we can specify, for all (Y, D), the agent's strategy by $x^*(Y, D, \theta) = y^A(\theta), \forall\theta$, the principal's strategy by $\rho^*(Y, D, x) = 1$ if $V^P(D \cup x) \geq V^P(D)$ and $\rho^*(Y, D, x) = 0$ otherwise, and the beliefs by

$$\Theta^*(Y, D, x) = \left\{\theta \in [0, 1] : U^A(\theta, x) \geq \max_{y \in D} U^A(\theta, D)\right\}$$

for all x. We then want to show that (22) is satisfied if and only if $x \in S^*$. We distinguish three cases:

- If $x \in S^*$ and $x \leq \min D^*(Y^{max})$, then (22) follows from the fact that the agent is upward biased—conditional on the agent preferring x over $\min D^*(Y^{max})$, the principal prefers x as well.
- If $x \in S^*$ and $\min D^*(Y^{max}) < x \leq \max D^*(Y^{max})$, inequality (22) follows from Assumption 1.
- If $x \notin S$ and hence $x > y^{max}$, then (22) is violated by definition of y^{max}.

Taken together, the agent can implement any action belonging to S^*, hence the equilibrium outcome y^E is as specified in Proposition 6. $\square$