DISCUSSION PAPER SERIES

DP16896

Contracting over Persistent Information

Wei ZHAO, Claudio Mezzetti, Ludovic Renou and Tristan Tomala

ORGANIZATIONAL ECONOMICS

CEPR

ISSN 0265-8003

Contracting over Persistent Information

Wei ZHAO, Claudio Mezzetti, Ludovic Renou and Tristan Tomala

Discussion Paper DP16896
Published 13 January 2022
Submitted 08 December 2021

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

Contracting over Persistent Information

Abstract

We consider a dynamic principal-agent problem, where the sole instrument the principal has to incentivize the agent is the disclosure of infor- mation. The principal aims at maximizing the (discounted) number of times the agent chooses the principal's preferred action. We show that there exists an optimal contract, where the principal stops disclosing information as soon as its most preferred action is a static best reply for the agent, or else con- tinues disclosing information until the agent perfectly learns the principal's private information. If the agent perfectly learns the state, he learns it in fi- nite time with probability one; the more patient the agent, the later he learns it.

JEL Classification: N/A

Keywords: N/A

Wei ZHAO - wei.zhao1@hec.edu
HEC

Claudio Mezzetti - c.mezzetti@uq.edu.au
University of Queensland

Ludovic Renou - lrenou.econ@gmail.com
Queen Mary University of London and CEPR

Tristan Tomala - tomala@hec.fr
HEC

Acknowledgements
Wei Zhao gratefully acknowledges the support of the HEC Foundation. Claudio Mezzetti thankfully acknowledges financial support from the Australian Research Council Discovery grant DP190102904. Ludovic Renou gratefully acknowledges the support of the Agence Na- tionale pour la Recherche under grant ANR CIGNE (ANR-15-CE38-0007-01) and through the ORA Project "Ambiguity in Dynamic Environments" (ANR-18-ORAR-0005). Tristan Tomala gratefully acknowledges the support of the HEC Foundation and ANR/Investissements d'Avenir under grant ANR-11-IDEX-0003/Labex Ecodec/ANR-11-LABX-0047.

CONTRACTING OVER PERSISTENT INFORMATION

WEI ZHAO, CLAUDIO MEZZETTI, LUDOVIC RENOU, AND TRISTAN TOMALA

ABSTRACT. We consider a dynamic principal-agent problem, where the *sole* instrument the principal has to incentivize the agent is the disclosure of information. The principal aims at maximizing the (discounted) number of times the agent chooses the principal's preferred action. We show that there exists an optimal contract, where the principal stops disclosing information as soon as its most preferred action is a static best reply for the agent, or else continues disclosing information until the agent *perfectly* learns the principal's private information. If the agent perfectly learns the state, he learns it in finite time with probability one; the more patient the agent, the later he learns it.

KEYWORDS: Dynamic, contract, information, revelation, disclosure, sender, receiver, persuasion.

JEL CLASSIFICATION: C73, D82.

Date: October 5, 2021.

Wei Zhao gratefully acknowledges the support of the HEC Foundation. Claudio Mezzetti thankfully acknowledges financial support from the Australian Research Council Discovery grant DP190102904. Ludovic Renou gratefully acknowledges the support of the Agence Nationale pour la Recherche under grant ANR CIGNE (ANR-15-CE38-0007-01) and through the ORA Project "Ambiguity in Dynamic Environments" (ANR-18-ORAR-0005). Tristan Tomala gratefully acknowledges the support of the HEC Foundation and ANR/Investissements d'Avenir under grant ANR-11-IDEX-0003/Labex Ecodec/ANR-11-LABX-0047.

1. Introduction

We consider a dynamic "principal-agent" model, where the sole instrument the principal has is information.[1] Principal and agent are engaged in a long-term relationship. The principal aims at inducing the agent to choose an action – the principal's most preferred action – as long as possible, and can only do so by disclosing information about an unknown state. To give examples, the agent is a politician and the principal a lobbyist who aims to maintain the status-quo course of action. Or the principal is: (i) an external consultant with a clear agenda about what a company (the agent) should do, (ii) a department in a corporation aiming to maintain a central role while advising the CEO, (iii) a technology leading, multinational firm in a joint venture with another firm located in a less developed country, aiming for the joint venture to continue, which may require not fully disclosing all information about the multinational's technology. We assume that the principal commits to a disclosure policy, which we refer to as the offer of a "contract." The dynamic contracting problem we study is, therefore, a *dynamic persuasion problem*.

The standard approach in the study of dynamic contracting models (e.g., Spear and Srivastava (1987)) is to use the agent's continuation value as a state variable. The principal's Bellman equation is then the fixed point of an operator, which satisfies a promised keeping constraint in addition to incentive constraints. In dynamic persuasion models, there is an additional complication, however. The information the principal commits to disclose to the agent generates a *martingale* of beliefs: the posterior beliefs of the agent must be equal in expectation to his prior beliefs. We thus need to incorporate the agent's beliefs as an additional state variable and to impose the constraint that the belief process is a martingale. In spite of the increased dimensionality of the principal's problem, we are able to provide a complete characterization of an

[1]That is, the principal cannot make transfers, terminate the relationship, choose allocations or constrain the agent's choices.

optimal contract by simultaneously solving for the evolution of the agent's beliefs and promised utility. To the best of our knowledge, we are the first to tackle this difficulty.

We illustrate the main properties of our optimal policy – particularly how beliefs evolve over time – with the help of Figure 1. Figure 1 plots four representative evolutions of the agent's belief about the "high opportunity cost" state – the state where the cost to incentivize the agent relative to the benefit is the highest. In each panel, the grey region "OPT" indicates the region at which choosing the principal's most preferred action is (statically) optimal for the agent. An arrow pointing from one belief to another indicates how the agent revised his belief within the period following a signal's realization. Multiple arrows originating from the same point thus represent the information disclosed by the policy. Within a period, the agent takes a decision after having revised his beliefs. Arrows have different colors/patterns. At all beliefs at the end of continuous black arrows, the agent chooses the principal's most preferred action. At all beliefs at the end of dotted magenta arrows, he chooses what is best given his current belief.

Here are the general properties of our optimal policy. First, the agent updates his belief until either he perfectly learns the state, or choosing the principal's most preferred action becomes (statically) optimal. Moreover, if the agent learns the state, he learns it in finite time. After the agent has learned the state, he will take his optimal action in that state, while as long as he keeps getting pieces of information from the principal he will take the principal's preferred action. By trickling down bits of information the principal is able to induce the agent to delay moving away from his favorite course of action. In some instances, the principal will promise eventual full disclosure of the state with probability one. In other instances, the principal will be able to stir the agent's beliefs so that with positive probability the agent takes the principal's favorite action forever. We provide a characterization of when this occurs.

Second, an important novel ingredient of our model is that the agent's opportunity cost relative to the principal's benefit of taking the principal's preferred

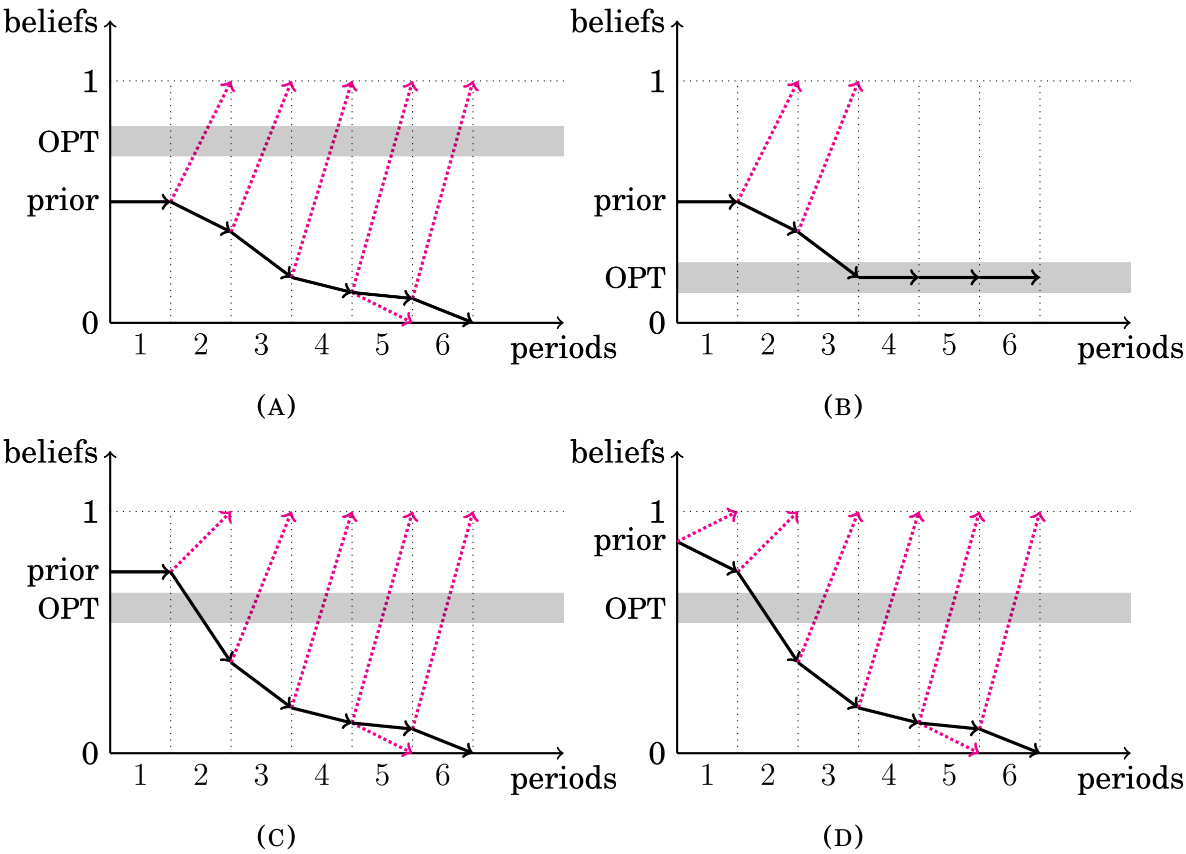

FIGURE 1. Evolution of actions and beliefs over time

action is different in different states. We show that along the paths at which the agent plays the principal's most preferred action, his beliefs about the "high opportunity cost" state are decreasing. Intuitively, the optimal contract exploits the asymmetry in opportunity costs and lowers the perceived opportunity cost – hence making it easier to incentivize the agent – by biasing information disclosure in the direction of informing him when the opportunity cost is high.[2] For instance, suppose the agent is a company's CEO and the principal an advisor to the company who prefers the status quo. If the company's relatives benefit from moving away from the status quo to undertake a new project is higher in state ω_1, then the advisor would want to slowly trickle down information about whether the state is indeed ω_1. Such information is more valuable to the company, hence can be packaged in smaller bits, inducing the status quo to last longer.

[2]To be precise, under our policy, upon receiving the signal "the opportunity cost is high," the agent learns that this is indeed true. However, the signal is not sent with probability one. This corresponds to the (magenta/dotted) arrows ending at 1.

Third, with the exception of panel (D), the policy does not disclose information to the agent at the first period. Thus, defining persuasion as the act of changing the agent's beliefs prior to his making decision, information disclosure rewards the agent for following the recommendation, but does not *persuade* him in panels (A), (B) and (C). Yet, as panel (D) illustrates, the policy sometimes needs to persuade the agent. For instance, if the promise of full information disclosure at the next period wouldn't incentivize the agent, then persuading the agent is necessary. That is, the policy must generate a strictly positive value of information for the agent. There are other circumstances at which persuading the agent may be necessary. Persuasion may reduce the agent's opportunity cost of following the principal's recommendation sufficiently enough to compensate the loss to the principal, due to providing information at the start of the relationship.

Finally, with the exception of panel (B), the policy does not induce the agent to believe that playing the principal's most preferred action is optimal. This is markedly different from what we would expect from the static analysis of Kamenica and Gentzkow (2011). Intuitively, the "static" persuasion policy is sub-optimal because it does not extract all the information surplus it creates. Even in panel (B), the beliefs do not jump immediately to the "OPT" region. In fact, the belief process may approach the "OPT" region only asymptotically.

Related literature. The paper is part of the literature on Bayesian persuasion, pioneered by Kamenica and Gentzkow (2011), and recently surveyed by Kamenica (2019). The three most closely related papers are Ball (2019), Ely and Szydlowski (2020), and Orlov et al. (2019). In common with our paper, these papers study the optimal disclosure of information in dynamic games and show how the disclosure of information can be used as an incentive tool. The observation that information can be used to incentivize agents is not new and dates back to the literature on repeated games with incomplete information, e.g., Aumann et al. (1995). See Garicano and Rayo (2017) and Fudenberg and Rayo (2019) for some more recent papers exploring the role of information provision as an incentive tool.

The classes of dynamic games studied differ considerably from one paper to another, which makes comparisons difficult. In Ely and Szydlowski (2020), the agent has to repeatedly decide whether to continue working on a project or to quit (i.e., unlike our paper, there are only two actions); quitting ends the game. The principal aims at maximizing the number of periods the agent works on the project and can only do so by disclosing information about its complexity, modeled as the number of periods required to complete it. Thus, their dynamic game is a quitting game, while ours is a repeated game. When the project is either easy or difficult (i.e., when there are two states), the optimal disclosure policy initially persuades the agent that the task is easy, so that he starts working. (Naturally, if the agent is sufficiently convinced that the project is easy, there is no need to persuade him initially.) If the project is in fact difficult, the policy then discloses it at a later date, when completing the project is now within reach. A main difference with our optimal disclosure policy is that information comes in lumps in Ely and Szydlowski (2020), i.e., information is disclosed only at the initial period and at a later period, while information is repeatedly disclosed in our model.[3] Another main difference is as follows. In Ely and Szydlowski, only when the promise of full information disclosure at a later date is not enough to incentivize the agent to start working does the principal persuade the agent initially. This is not so with our policy: the principal persuades the agent in a larger set of circumstances. This initial persuasion reduces the cost of incentivizing the agent in future periods.

Orlov et al. (2019) also consider a quitting game, where the principal aims at delaying the quitting time as far as possible. The quitting time is the time at which the agent decides to exercise an option, which has different values to the principal and the agent. The principal chooses a disclosure policy informing the agent about the option's value. When, as in this paper, the principal commits to a long-run policy, the optimal policy is to fully reveal the state

[3]When there are more than two states, the optimal policy discloses information more frequently in Ely and Szydlowski (2020). The frequency of disclosure is thus a consequence of the dimensionality of the state space in their model, while it is not so in our model.

with some delay. (Note that the principal is referred to as the agent in their work.) This policy is not optimal in Ely and Szydlowski (2020), or in our paper. See Au (2015), Bizzotto et al. (Forthcoming), Che et al. (2020), Henry and Ottaviani (2019) and Smolin (2018) for other papers on information disclosure in quitting games, where the agent either waits and obtains additional information, or takes an irreversible action and stops the game.

Ball (2019) studies a continuous time model of information provision, where the state changes over time and payoffs are the ones of the quadratic example of Crawford and Sobel (1982). Ball shows that the optimal disclosure policy requires the sender to disclose the current state at a later date, with the delay shrinking over time. The main difference between his work and ours is the persistence of the state (also, we consider two different classes of games). When the state is fully persistent, as in Ely and Szydlowski (2020) and our model, full information disclosure with delay is not optimal in general. (See the discussion of Example 1 in Section 3.)

Finally, there are a few papers on dynamic persuasion, where the agent takes an action repeatedly. However, either the agent is myopic, e.g., Ely (2017) and Renault et al. (2017), or the principal cannot commit, e.g., Escude and Sinander (2020).

2. The problem

A principal and an agent interact over an infinite number of periods, indexed by $t \in \{1, 2, \dots\}$. At the first stage, before the interaction starts, the principal learns a payoff-relevant state $\omega \in \Omega$, while the agent remains uninformed. The prior probability of ω is $p_0(\omega) > 0$. At each period t, the principal sends a signal $s \in S$ and, upon observing the signal s, the agent takes decision $a \in A$. The sets A, S and Ω are finite. The cardinality of S is as large as necessary for the principal to be unconstrained in his information disclosure policy.[4]

[4]From Makris and Renou (2021), it is enough to have the cardinality of S as large as the cardinality of A.

We assume that there exists $a^* \in A$ such that the principal's payoff is strictly positive whenever a^* is chosen, and zero otherwise. We refer to a^* as the principal's preferred action. E.g., the principal wants the agent to purchase its products, to follow its advice, or to maintain the status-quo. We let $v : A \times \Omega \to \mathbb{R}$ be the principal's payoff function, with $v(a^*, \omega) > v(a, \omega) = 0$ for all $\omega \in \Omega$ and $a \in A \setminus \{a^*\}$. The agent's payoff function is $u : A \times \Omega \to \mathbb{R}$. The (common) discount factor is $\delta \in (0, 1)$.

We write A^{t-1} for $\underbrace{A \times \cdots \times A}_{t-1 \text{ times}}$ and S^{t-1} for $\underbrace{S \times \cdots \times S}_{t-1 \text{ times}}$, with generic elements a^t and s^t, respectively. A behavioral strategy for the principal is a collection of maps $\tau = (\tau_t)_{t=1}^{\infty}$, with $\tau_t : A^{t-1} \times S^{t-1} \times \Omega \to \Delta(S)$. Similarly, a behavioral strategy for the agent is a collection of maps $\sigma = (\sigma_t)_{t=1}^{\infty}$ with $\sigma_t : A^{t-1} \times S^{t-1} \times S \to \Delta(A)$.

We write $\mathbf{V}(\tau, \sigma)$ for the principal's payoff and $\mathbf{U}(\tau, \sigma)$ for the agent's payoff under the strategy profile (τ, σ). The objective is to characterize the maximal payoff the principal achieves if he commits to a strategy τ, that is,

$$\sup_{(\tau,\sigma)} \mathbf{V}(\tau, \sigma),$$

subject to

$$\mathbf{U}(\tau, \sigma) \geq \mathbf{U}(\tau, \sigma'),$$

for all σ'.

Several comments are worth making. First, we interpret the strategy the principal commits to as a *contract* specifying, as a function of the state, the information to be disclosed at each history of realized signals and actions. In other words, the contract specifies a statistical experiment at each history of realized signals and actions, which enables the principal to punish the agent for choosing the "wrong action." The principal chooses the contract prior to learning the state. An alternative interpretation is that neither the principal nor the agent know the state, but the principal has the ability to conduct statistical experiments contingent on past signals and actions. We can partially dispense with the commitment assumption. Indeed, since the choices of statistical experiments are observable, we can construct strategies that incentivize the principal to implement the specified statistical experiments.[5] Second, the only additional information the agent obtains each period is the outcome of the statistical experiment. Third, the state is fully persistent and the principal perfectly monitors the action of the agent. Finally, the only instrument available to the principal is information. The principal can neither remunerate the agent nor terminate the relationship nor allocate different tasks to the agent. We purposefully make all these assumptions to address our main question of interest: what is the optimal way to incentivize the agent with information only?

3. Optimal contracts

This section characterizes optimal contracts and discusses their most salient properties.

3.1. **A recursive formulation.** The first step towards characterizing optimal contracts is to reformulate the principal's problem as a recursive problem. To do so, we introduce two state variables. The first state variable is promised continuation payoff. It is well-known that classical dynamic contracting problems admit recursive formulations if we introduce promised continuation payoff as a state variable and impose promise-keeping constraints, e.g., Spear and Srivastava (1987). The second state variable we introduce is beliefs. We now turn to the formal reformulation of the problem.[6]

We first need some additional notation. Denote $p \in \Delta(\Omega)$ a generic belief about the state. We let $u(a,p) := \sum_{\omega} p(\omega)u(a,\omega)$ be the agent's expected stage

[5]The simplest such strategy is for the agent to play $a \neq a^*$ in all future periods after a deviation by the principal. This strategy may not be sequentially rational, however.

[6]A nearly identical reformulation already appeared in Ely (2015), one of the working versions of Ely (2017). We remind the reader that Ely (2017) analyzes the interaction between a long-run principal and a sequence of short-run agents. (See also Renault et al. (2017).) While discussing the extension of his model to the interaction between a long-run principal and a long-run agent, Ely (2015) derived a recursive reformulation nearly identical to ours. However, he didn't study further the reformulated problem. We start from the recursive formulation and use it to derive an optimal policy. See Section A.2 for a detailed comparison of the two formulations.

payoff of choosing a when his belief is p, $m(p) := \max_{a\in A} u(a,p)$ be the agent's optimal stage payoff when his belief is p, and $M(p) := \sum_{\omega} p(\omega) \max_{a\in A} u(a,\omega)$ be the agent's expected stage payoff if he learns the state prior to choosing an action. Note that m is a piecewise linear convex function, that M is linear and that $m(p) \leq M(p)$ for all p. Similarly, we let $v(a,p)$ be the principal's payoff when the agent chooses a and the principal's belief is p. Finally, let $P := \{p \in \Delta(\Omega) : m(p) = u(a^*,p)\}$ be the set of beliefs at which a^* is optimal. The set P is a closed convex set.

Let $\mathcal{W} \subseteq \Delta(\Omega) \times \mathbb{R}$ be such that $(p,w) \in \mathcal{W}$ if and only if $w \in [m(p), M(p)]$. Throughout, we consider functions $V : \mathcal{W} \to \mathbb{R}$, with the interpretation that $V(p,w)$ is the principal's payoff if he promises a continuation payoff of w to the agent when the agent's current belief is p.

The principal's maximal payoff is $V^*(p_0, m(p_0))$, where V^* is the unique fixed point of the contraction mapping T, defined by

$$T(V)(p,w) := \begin{cases} \max_{\left((\lambda_s,(p_s,w_s),a_s)\in\Delta(\Omega)\times\mathcal{W}\times A\right)_{s\in S}} \sum_{s\in S} \lambda_s[(1-\delta)v(a_s,p_s)+\delta V(p_s,w_s)], \\ \text{subject to:} \\ (1-\delta)u(a_s,p_s)+\delta w_s \geq m(p_s) \quad \text{for all } s \text{ such that } \lambda_s>0, \\ \sum_{s\in S}\lambda_s[(1-\delta)u(a_s,p_s)+\delta w_s] \geq w, \\ \sum_{s\in S}\lambda_s p_s = p, \sum_{s\in S}\lambda_s = 1. \end{cases}$$

We briefly comment on this maximization program. The program maximizes the principal's expected payoff over *policies*, i.e., maps from $\mathcal{W}$ to $(\Delta(\Omega)\times\mathcal{W}\times A)^{|S|}$. At each (p,w), a policy prescribes the probability λ_s that the realized signal is s and conditional on s, the belief p_s, the promised utility w_s, and the recommended action a_s. The first constraint is the incentive-compatibility condition that the agent prefers to obey the recommendation a_s when w_s is the promised continuation payoff and p_s is the agent's belief. To understand the right-hand side, observe that the agent can always play a static best reply to any belief, so that his expected payoff must be at least $m(p_s)$ when his current belief is p_s.[7] Conversely, if the contract recommends action a_s and the agent does not obey, the contract can specify no further information revelation, in which case the agent's payoff is at most $m(p_s)$. Therefore, $m(p_s)$ is the agent's min-max payoff. The second constraint is promise-keeping: if the principal promises the continuation payoff w at a period, the contract must honor that promise in subsequent periods. The third constraint states that the policy selects a splitting of p, i.e. a distribution over posteriors with expectation p.

Throughout, we slightly abuse notation and write τ for a policy. Note that to each contract corresponds a policy, and conversely. A policy is feasible if it specifies a feasible tuple $((\lambda_s, (p_s, w_s), a_s))_{s\in S}$ for each (p, w), i.e., a tuple satisfying the constraints of the maximization problem $T(V)(p, w)$.

Two important observations are worth making. First, for any function V, $T(V)$ is a concave function of the pair (p, w). Concavity reflects the fact that the more information the principal discloses, the harder it is to reward the agent in the future. Second, $T(V)$ is a (weakly) decreasing function of w, that is, the more the principal promises to the agent, the harder it is to incentivize the agent to play a^*.

Proposition 1. *The value function V^* is concave in both arguments and weakly decreasing in w.*

Proposition 1 together with the recursive formulation has a number of implications. First, if the principal induces the posterior p_s while recommending the action a_s and promising the continuation payoff w_s, then he should not have an incentive to disclose more information in that period, that is, we cannot have $V(p_s, (1-\delta)u(a_s, p_s) + \delta w_s) > (1-\delta)v(a_s, p_s) + \delta V(p_s, \omega_s)$. Indeed, if the inequality was satisfied at some signal s, the principal would strictly benefit from releasing further information at p_s so as to achieve $V(p_s, (1-\delta)u(a_s, p_s) + \delta w_s)$.

[7]More precisely, if the agent's belief at period t is p_t, he obtains the payoff $m(p_t)$ by playing a static best-reply. Since the function m is convex and beliefs follow a martingale, his expected payoff is therefore at least $(1-\delta)\sum_{t'\geq t}\delta^{t'-t}\mathbb{E}[m(\mathbf{p}_{t'})|\mathcal{F}_t] \geq m(p_t)$, where $\mathcal{F}_t$ is the agent's filtration at period t.

Second, if the principal does not recommend a^* at a period, then the principal never recommends a^* at a subsequent period, that is, the principal's continuation value is zero. In other words, as soon as an action other than a^* is played, the principal stops incentivizing the agent to play a^*. The intuition is simple. Suppose to the contrary that the principal were to recommend $a_s \neq a^*$ after the signal s at period t and a^* at the next period. Consider the policy change where the principal anticipates the disclosure of the information: what incentivizes the agent to play a^* at period $t+1$ is disclosed in period t. This policy change is feasible and increases the principal's payoff, a contradiction. This property justifies thinking of the principal's preferred action a^* as a status quo, which the principal tries to induce the agent to maintain as long as possible.

Third, there is at most one signal s^* at which the principal recommends the agent to play a^*. Moreover, whenever the principal recommends a^*, the agent is indifferent between obeying the recommendation or deviating. In other words, the promised continuation payoff does not leave rents to the agent. Intuitively, if two signals recommended a^*, the principal would not lose from merging them into one. If the agent were given a positive rent when signal s^* realizes, the principal would benefit by a change in policy that reduces the agent's promised utility associated with s^* (since the value function is decreasing in promised utility). For that change in policy to be feasible, the change must increase the promised utility when some other signal $s \neq s^*$ is realized. As we have already seen, this does not affect the principal's payoff (since the principal obtains a zero payoff in all periods which follow a recommendation different from a^*).

These observations are summarized in Proposition 2.

Proposition 2. *For all* (p, w)*, there exists a solution* $(\lambda_s, p_s, w_s, a_s)_{s\in S}$ *to* $T(V^*)(p, w)$ *such that*

(i): *For all* $s \in S$ *such that* $\lambda_s > 0$*, we have*

$$(1-\delta)\, v\left(a_s, p_s\right) + \delta V^*\left(p_s, w_s\right) = V^*\left(p_s, (1-\delta)\, u\left(a_s, p_s\right) + \delta w_s\right).$$

(ii): *For all* $s \in S$ *such that* $\lambda_s > 0$ *and* $a_s \neq a^*$*,* $V^*(p_s, w_s) = 0$.

(iii): *There exists at most one signal* $s^* \in S$ *such that* $\lambda_{s^*} > 0$ *and* $a_{s^*} = a^*$. *Moreover,*

$$(1-\delta)u(a_{s^*}, p_{s^*}) + \delta w_{s^*} = m(p_{s^*}).$$

Proposition 2 states key properties that an optimal policy possesses. We conclude this section with a partial converse, that is, we state properties that guarantee optimality of a policy. This partial converse is at the heart of our analysis, which consists in constructing a policy and proving that it satisfies all the properties required for optimality.

We need some additional notation. First, let Q^1 be the set of beliefs at which the agent has an incentive to play a^* if he is promised full information disclosure at the next period. That is,

$$Q^1 := \{p \in \Delta(\Omega) : (1-\delta)u(a^*, p) + \delta M(p) \geq m(p)\}.$$

If Q^1 is empty, then all policies are optimal as the principal can never incentivize the agent to play a^*. The set Q^1 is convex.

Second, for all $p \in Q^1$, we write $\mathbf{w}(p) \in [m(p), M(p)]$ for the continuation payoff that makes the agent indifferent between playing action a^* and receiving the continuation payoff $\mathbf{w}(p)$ in the future, and playing a best reply to the belief p forever. That is, $\mathbf{w}(p)$ solves:

$$(1-\delta)u(a^*, p) + \delta \mathbf{w}(p) = m(p).$$

Theorem 1. *Consider any feasible policy inducing the value function* $\tilde{V}$. *If* $\tilde{V}$ *is concave in both arguments, decreasing in* w *and satisfies*

$$\tilde{V}(p, m(p)) \geq (1-\delta)v(a^*, p) + \delta \tilde{V}(p, \mathbf{w}(p)),$$

for all $p \in Q^1$*, then the policy is optimal.*

Proof. We argue that $\tilde{V}$ is the fixed point of the operator T, hence $\tilde{V} = V^*$. Let $(\lambda_s, p_s, w_s, a_s)_{s \in S}$ be a solution to the maximization problem $T(\tilde{V})(p, w)$. We start by the following observation. Consider any s such that $a_s \neq a^*$. We have

$$(1-\delta)v(a_s,p_s)+\delta\tilde{V}(p_s,w_s)=\delta\tilde{V}(p_s,w_s)\leq\tilde{V}(p_s,w_s)\leq\tilde{V}(p_s,(1-\delta)u(a_s,p_s)+\delta w_s),$$

where the last inequality follows from the fact that $\tilde{V}$ is decreasing in w and $m(p_s)\leq(1-\delta)u(a_s,p_s)+\delta w_s\leq(1-\delta)m(p_s)+\delta w_s\leq w_s$.

Consider now any s such that $a_s = a^*$. Since $(\lambda_s,p_s,w_s,a_s)_{s\in S}$ is feasible, we have

$$(1-\delta)u(a^*,p_s)+\delta w_s\geq m(p_s),$$

hence $p_s\in Q^1$ and therefore,

$$\tilde{V}(p_s,m(p_s))\geq(1-\delta)v(a^*,p_s)+\delta\tilde{V}\Big(p_s,\underbrace{\frac{-(1-\delta)u(a^*,p_s)+m(p_s)}{\delta}}_{\mathbf{w}(p_s)}\Big).$$

The concavity of $\tilde{V}$ implies that

$$\tilde{V}(p_s,(1-\delta)u(a^*,p_s)+\delta w_s)-\tilde{V}(p_s,m(p_s))\geq\delta\Big[\tilde{V}(p_s,w_s)-\tilde{V}\Big(p_s,\mathbf{w}(p_s)\Big)\Big],$$

where we use the identity $(1-\delta)u(a^*,p_s)+\delta w_s-m(p_s)=\delta(w_s-\mathbf{w}(p_s))$ and observation (a) about concave functions in Section A.1.

Combining the above two inequalities implies,

$$\tilde{V}(p_s,(1-\delta)u(a^*,p_s)+\delta w_s)\geq(1-\delta)v(a^*,p_s)+\delta\tilde{V}(p_s,w_s).$$

It follows that

$$\begin{aligned}
T(\tilde{V})(p,w) &= \sum_{s\in S}\lambda_s\Big[(1-\delta)v(a_s,p_s)+\delta\tilde{V}(p_s,w_s)\Big]\\
&\leq \sum_{s\in S}\lambda_s\Big[\tilde{V}(p_s,(1-\delta)u(a_s,p_s)+\delta w_s)\Big]\\
&\leq \tilde{V}\left(\sum_{s\in S}\lambda_s p_s,\sum_{s\in S}\lambda_s((1-\delta)u(a_s,p_s)+\delta w_s)\right)\\
&\leq \tilde{V}(p,w),
\end{aligned}$$

where the second inequality follows from the concavity of $\tilde{V}$ and the third inequality from $\tilde{V}$ being decreasing in w.

Conversely, since the policy inducing $\tilde{V}$ is feasible, we must have that $T(\tilde{V})(p,w) \geq \tilde{V}(p,w)$ for all (p,w). This completes the proof. $\square$

We now turn to the construction of an optimal policy. We first build some intuition.

3.2. **Optimal policy: building intuition.** From Proposition 2, any optimal policy must satisfy three properties: (i) The principal should not have an incentive to disclose additional information at any period. (ii) If the principal does not recommend a^* at a period, then he never recommends it at any subsequent period. (iii) At the unique signal s^* at which the principal recommends the agent to play a^*, the promised continuation payoff w_{s^*} leaves no rents to the agent, i.e., $(1-\delta)u(a_{s^*}, p_{s^*}) + \delta w_{s^*} = m(p_{s^*})$.

This leaves important questions unanswered. What are the beliefs at which the agent plays a^*? How does the principal compensate the agent for playing a^*? Does the principal need to reveal information at the prior belief? Does the agent learn the state? If so, does he learn it in finite time?

In the remainder of the paper, we answer these questions in the binary case $\Omega = \{\omega_0, \omega_1\}$. Throughout, probabilities refer to the probability of ω_1 and abusing notation write p for $p(\omega_1)$. If non-empty, the set of beliefs P at which a^* is optimal is then a closed interval $[\underline{p}, \overline{p}]$. Similarly, if non-empty, the set of beliefs Q^1 is a closed interval $[\underline{q}^1, \overline{q}^1]$. (Recall that Q^1 is the set of beliefs at which the agent has an incentivize to play a^* if promised full information revelation at the next period.) Note that $\underline{q}^1 = 0$ if and only if a^* is optimal at $p = 0$ and $\overline{q}^1 = 1$ if and only if a^* is optimal at $p = 1$. For a graphical illustration, see Figure 2.

An important feature of our model is that the agent's opportunity cost of choosing a^* rather than his best action, relative to the principal's benefit, differs in different states. When the state is ω_0 (resp. ω_1) the opportunity cost relative to the benefit is $[m(0) - u(a^*, 0)]/v(a^*, 0)$ (resp. $[m(1) - u(a^*, 1)]/v(a^*, 1)$).

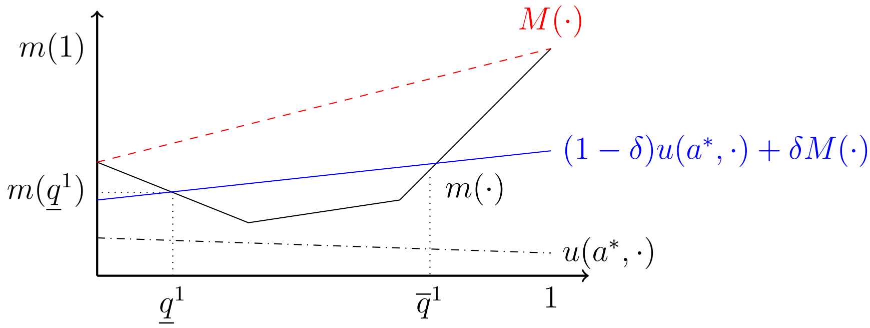

FIGURE 2. Construction of the set Q^1 when $\Omega = \{\omega_0, \omega_1\}$

Without loss of generality, we assume that the agent's opportunity cost relative to the principal's benefit of a^* is higher in state ω_1 than in state ω_0:

Assumption 1.

$$\frac{m(1) - u(a^*, 1)}{v(a^*, 1)} \geq \frac{m(0) - u(a^*, 0)}{v(a^*, 0)}.$$

Intuitively, Assumption 1 implies that it is more efficient to incentivize the agent to take action a^* at lower beliefs. As we shall see, the optimal policy heavily exploits this observation. We note that if a^* is optimal for the agent at $p = 1$, i.e., $m(1) = u(a^*, 1)$, then a^* is also optimal at $p = 0$. Consequently, a^* is optimal at all beliefs, i.e., $P = [0, 1]$. In what follows, we exclude this trivial case and assume that $1 \notin P$.

To strengthen our intuition on the construction of an optimal policy, we briefly return to the original description of the problem. Let (τ, σ) be a profile of strategies and denote by $\mathbb{P}_{\tau,\sigma}(\cdot|\omega)$ the distribution over signals and actions induced by (τ, σ) conditional on ω. We can write the principal's expected payoff $\mathbf{V}(\tau, \sigma)$ as:

$$(1-\delta) \sum_{\omega} \left(p_0(\omega) \left(\sum_t \sum_{s^t, a^{t-1}} \delta^{t-1} \mathbb{P}_{\sigma,\tau}(s^t, a^{t-1}|\omega) \tau_t(a^*|s^t, a^{t-1}) \right) v(a^*, \omega) \right) = \lambda^* v^*(a^*, p^*),$$

with

$$\lambda^* := (1-\delta) \sum_{\omega} p_0(\omega) \left(\sum_t \sum_{s^t, a^{t-1}} \delta^{t-1} \mathbb{P}_{\sigma,\tau}(s^t, a^{t-1}|\omega) \tau_t(a^*|s^t, a^{t-1}) \right)$$

the discounted probability of recommending action a^* and

$$p^* \quad := \quad \frac{(1-\delta)p_0(\omega_1)\left(\sum_t \sum_{s^t,a^{t-1}} \delta^{t-1}\mathbb{P}_{\sigma,\tau}(s^t,a^{t-1}|\omega_1)\tau_t(a^*|s^t,a^{t-1})\right)}{\lambda^*},$$

the average discounted probability of ω_1 when a^* is played.[8] As expected, the principal's payoff only depends on how often a^* is played and the average belief at which it is played.

We now make two observations, which will enable us to rewrite the principal's expected payoff and get important insights on optimal policies. First, if we let $p^\dagger$ be the average discounted probability of ω_1 when a^* is *not* recommended, we have that $\lambda^* p^* + (1-\lambda^*)p^\dagger = p_0$ since the belief process is a martingale. Second, since the agent's static payoff is bounded from above by $M(p)$ when his belief is p, his ex-ante expected payoff is bounded above by:

$$\lambda^* u(a^*, p^*) + (1-\lambda^*)M(p^\dagger) = \lambda^*[u(a^*, p^*) - M(p^*)] + M(p_0). \tag{1}$$

Since the agent's ex-ante payoff must be at least $m(p_0)$, we can define $c \geq 0$ as the maximum agent's rent:

$$c = \lambda^*[u(a^*, p^*) - M(p^*)] + M(p_0) - m(p_0). \tag{2}$$

With the help of these two observations, we can rewrite the principal's expected payoff as:

$$\frac{v(a^*, p^*)}{M(p^*) - u(a^*, p^*)} \times \big(M(p_0) - m(p_0) - c\big).$$

The first term captures the benefit of incentivizing the agent to play a^* relative to the cost. Since $\frac{v(a^*,0)}{v(a^*,1)} \geq \frac{m(0)-u(a^*,0)}{m(1)-u(a^*,1)}$, it is decreasing in p^*.[9] Ceteris paribus, the lower the average belief at which the agent plays a^*, the higher the principal's expected payoff.

[8] Note that p^* cannot be lower than $\underline{q}^1$ since the agent would never play a^* at beliefs lower than $\underline{q}^1$.

[9] This follows from the observation that $M(p^*) - u(a^*, p^*) = p^*[(m(1) - m(0)) - (u(a^*, 1) - u(a^*, 0))] + m(0) - u(a^*, 0)$, $v(a^*, p^*) = p^*(v(a^*, 1) - v(a^*, 0)) + v(a^*, 0)$, and simple algebra.

The second term captures how the principal rewards the agent for playing a^* with his only instrument: information. The term $M(p_0)-m(p_0)$ is the maximal value of information the principal can create. Ceteris paribus, the principal's payoff is decreasing in c, that is, the best is to leave no rents to the agent and to create as much information as necessary to repay the agent. Notice that $c = 0$ is only achieved by both leaving no rents to the agent and having the agent informed of the state when he does not play a^*.

The above discussion thus suggests three guiding principles in constructing an optimal policy. First, the policy perfectly informs the agent whenever the recommendation differs from a^*, that is, the agent's belief is either 0 or 1 when he does not play a^*. Second, the policy leaves as little rent as possible to the agent. Naturally, it is not always possible to leave no rents. E.g., when the prior belief $p_0 \notin Q^1$, the agent must be given some strictly positive rent if he is to ever play a^*. Third, the policy must recommend a^* at the lowest beliefs possible.

Now, recall that we can always find optimal policies where a^* is recommended after at most one signal s^*. The guiding principles thus further hint that all splittings must have at most three values p_{s^*}, 0, and 1 in their support: the belief p_{s^*} at which a^* is recommended and the degenerate beliefs 0 and 1 at which the agent is perfectly informed. Moreover, we want p_{s^*} as low as possible subject to satisfying all constraints. Finally, since the policy must leave as little rents as possible, the best is to promise the continuation payoff $\mathrm{w}(p_{s^*})$ that makes the agent indifferent between playing a^* at the current period and receiving $\mathrm{w}(p_{s^*})$ in the future and playing a best reply to the belief p_{s^*} forever. As we see next, we can indeed construct an optimal policy with all these features.

3.3. **Optimal policy: a formal description.** We define a family of policies $(\tau_q)_{q\in[\underline{q}^1,\overline{q}^1]}$ indexed by a belief q, and prove later the existence of $q^* \in [\underline{q}^1,\overline{q}^1]$ such that the policy τ_{q^*} is optimal. Recall that a policy prescribes a splitting $(\lambda_s,p_s)_{s\in S}$, a profile of recommendations $(a_s)_{s\in S}$ and a profile of continuation payoffs $(w_s)_{s\in S}$ for each $(p,w)\in\mathcal{W}$. Policy τ_q is qualitatively the same in each of the following four regions:

$$\begin{aligned}
\mathcal{W}^1_q &:= \Big\{(p,w): p\in[0,\underline{q}^1), w\le \frac{\underline{q}^1-p}{\underline{q}^1}m(0)+\frac{p}{\underline{q}^1}m(\underline{q}^1)\Big\},\\
\mathcal{W}^2_q &:= \Big\{(p,w): p\in(q,1], \frac{1-p}{1-q}m(q)+\frac{p-q}{1-q}m(1)<w\le\frac{1-p}{1-\underline{q}^1}m(\underline{q}^1)+\frac{p-\underline{q}^1}{1-\underline{q}^1}m(1)\Big\}\\
&\qquad \bigcup\Big\{(p,w): p\in[\underline{q}^1,q], w\le\frac{1-p}{1-\underline{q}^1}m(\underline{q}^1)+\frac{p-\underline{q}^1}{1-\underline{q}^1}m(1)\Big\},\\
\mathcal{W}^3_q &:= \Big\{(p,w): p\in(q,1], w\le\frac{1-p}{1-q}m(q)+\frac{p-q}{1-q}m(1)\Big\},\\
\mathcal{W}^4_q &:= \mathcal{W}\setminus(\mathcal{W}^1_q\cup\mathcal{W}^2_q\cup\mathcal{W}^3_q).
\end{aligned}$$

Figure 3 illustrates the four regions, with $\mathcal{W}^1_q$ the black region, $\mathcal{W}^2_q$ the region with vertical lines, $\mathcal{W}^3_q$ the gray region, and $\mathcal{W}^4_q$ the region with northwest lines. Observe that regions $\mathcal{W}^1_q$ and $\mathcal{W}^4_q$ do not depend on the parameter q, while the other two do.

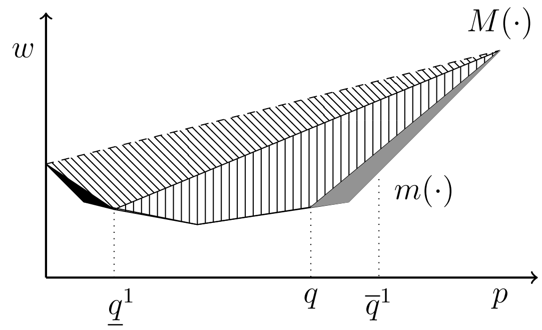

FIGURE 3. The regions $\mathcal{W}^1_q$, $\mathcal{W}^2_q$, $\mathcal{W}^3_q$ and $\mathcal{W}^4_q$.

We now describe the policy τ_q starting with region $\mathcal{W}^2_q$. Define functions $\lambda:\mathcal{W}\to[0,1]$ and $\varphi:\mathcal{W}\to[0,1]$ so that $(\lambda(p,w),\varphi(p,w))$ is the unique solution of

$$\begin{pmatrix}p\\ w\end{pmatrix}=\lambda(p,w)\begin{pmatrix}\varphi(p,w)\\ m(\varphi(p,w))\end{pmatrix}+(1-\lambda(p,w))\begin{pmatrix}1\\ m(1)\end{pmatrix} \tag{3}$$

for all $w>m(p)$, and $(\lambda(p,m(p)),\varphi(p,m(p)))=(1,p)$.

When (p,w) is in region $\mathcal{W}^2_q$, $\tau_q(p,w)$ induces belief and continuation payoff $(\varphi(p,w),m(\varphi(p,w)))$ with probability $\lambda(p,w)$, and belief and continuation payoff $(1,m(1))$ with the complementary probability. In the former case, the policy recommends action a^*. (In the latter case, the policy obviously recommends an action that is optimal at belief $p = 1$.)

The policy thus informs the agent when the state is ω_1. As we already suggested, the rationale for disclosing when the state is ω_1 is two-fold. First, the lower the agent's belief, the lower the cost of incentivizing the agent to play a^* relative to the principal's benefit. Second, to satisfy the promise-keeping constraint, the policy needs to compensate the agent for playing a^*. Since the principal's payoff is zero when the agent takes any action different from a^*, the best is to choose a compensation, which guarantees the highest probability of playing a^*. Putting these two observations together, at (p, w), policy $\tau_q(p, w)$ finds two beliefs (p', p'') such that (i) the agent is asked to play a^* at p', (ii) $p' < p$ since the agent should play a^* at the lowest belief, and (iii) the probability of p' is as high as possible. The best splitting is to have p' as close as possible to p and p'' as far as possible, i.e., equal to 1. Note that since $(1 - \lambda(p, w))m(1) + \lambda(p, w)m(\varphi(p, w)) = w$, the policy leaves no rents to the agent in region $\mathcal{W}_q^2$. See Figure 4 for an illustration.

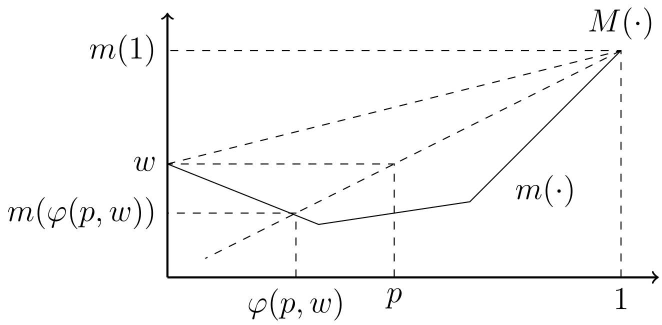

FIGURE 4. Construction of λ and φ

Observe that starting with $(p, w) \in \mathcal{W}_q^2$, the decreasing sequence of beliefs $(\varphi(p, w), \varphi^2(p, w), \dots)$ (and corresponding payoffs) reaches either region $\mathcal{W}_q^4$ – as in Panels (A) and (C) of Figure 1 – or a belief in P at which it is statically optimal for the agent to play a^* – as in panel (B) of Figure 1.[10] In the latter case, the policy recommends a^* and stops disclosing information (i.e., the belief stays constant).

[10]We write $\varphi^2(p, w)$ for $\varphi(\varphi(p, w), m(\varphi(p, w)))$.

When (p, w) is in region $\mathcal{W}_q^4$, the agent cannot be incentivized to play a^* at (p, w).[11] In that case, the policy splits p into posteriors 0, $\underline{q}^1$, and 1 with respective probabilities λ_0, $\lambda_{\underline{q}^1}$ and λ_1. Conditional on 0 (resp., 1), the policy recommends action the action optimal at 0, (resp., the action optimal at 1), and promises a continuation payoff of $m(0)$ (resp., $m(1)$). Conditional on $\underline{q}^1$, the policy recommends action a^* and promises a continuation payoff of $\text{w}(\underline{q}^1)$. Doing so, the principal ensures that the agent plays a^* one more time. The probabilities $(\lambda_0, \lambda_{\underline{q}^1}, \lambda_1) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+$ are the unique solution to:

$$\lambda_0 \begin{pmatrix} 0 \\ m(0) \\ 1 \end{pmatrix} + \lambda_{\underline{q}_1} \begin{pmatrix} \underline{q}^1 \\ m(\underline{q}^1) \\ 1 \end{pmatrix} + \lambda_1 \begin{pmatrix} 1 \\ m(1) \\ 1 \end{pmatrix} = \begin{pmatrix} p \\ w \\ 1 \end{pmatrix}.$$

A solution exists since $\mathcal{W}_q^4$ is the convex hull of $(0, m(0))$, $(\underline{q}^1, m(\underline{q}^1))$ and $(1, m(1))$. In this region, the policy leaves no rents to the agent either.

When (p, w) is in region $\mathcal{W}_q^1$, the policy splits p into 0 (i.e., discloses that the state is ω_0) and $\underline{q}^1$ with respective probabilities $\frac{\underline{q}^1 - p}{\underline{q}^1}$ and $\frac{p}{\underline{q}^1}$. If the realized belief is 0, the policy recommends the action optimal at 0 and promises a continuation payoff of $m(0)$. If the realized belief is $\underline{q}^1$, the policy recommends a^* and promises a continuation payoff of $\text{w}(\underline{q}^1)$. The agent is thus made indifferent between playing a^* and receiving $\text{w}(\underline{q}^1)$ in the future, and playing a best reply to the belief $\underline{q}^1$ forever. Intuitively, in region $\mathcal{W}_q^1$, the principal cannot incentivize the agent to take action a^* by promising future information disclosure (since $p < \underline{q}^1$). Hence, the principal must first persuade the agent by disclosing some information. Note that the policy leaves rents to the agent – since $\frac{\underline{q}^1 - p}{\underline{q}^1} m(0) + \frac{p}{\underline{q}^1} m(\underline{q}^1) > w$.

When (p, w) is in region $\mathcal{W}_q^3$, the policy splits p into q and 1 with respective probabilities $\frac{1-p}{1-q}$ and $\frac{p-q}{1-q}$. Conditional on 1, the policy recommends the action optimal at 1 and promises a continuation payoff of $m(1)$. Conditional on q, the policy recommends a^* and promises a continuation payoff of $\text{w}(q)$. The agent is thus made indifferent between playing a^* and receiving $\text{w}(q)$ in the

[11] Recall that $\underline{q}^1$ is the lowest belief at which the agent can be incentivized to play a^*.

future, and playing a best reply to the belief q forever. The policy in this region is analogous to the one in region $\mathcal{W}^1_q$– the policy starts by disclosing some information. When $q = \overline{q}^1$, the reason for the analogy is immediate, as $\overline{q}^1$ is the highest belief at which the agent is willing to take action a^* at the current period in exchange for full information at the next period. As we shall see later, the optimal policy τ_{q^*} may require $q^* < \overline{q}^1$, in order to guarantee that the principal's value function is concave, a necessary requirement to minimize the cost of incentivizing the agent relative to the benefit to the principal. This completes the description of the policy τ_q.

3.4. **A worked-out example.** We now illustrate our construction with the help of an example.

Example 1. The agent has three possible actions a_0, a_1 and a^*, with a_0 (resp., a_1) the agent's optimal action when the state is ω_0 (resp., ω_1) and a^* the principal's favorite action. The prior probability of ω_1 is $\frac{3}{20}$ and the discount factor is $\frac{1}{2}$. The per-period payoffs are in Table 1, with the first entry corresponding to the principal's payoff.

TABLE 1. Payoff table of Example 1

	a_0	a_1	a^*
ω_0	$0, 1$	$0, 0$	$1, 3/4$
ω_1	$0, 0$	$0, 2$	$1, 3/4$

An interpretation of this example is that the principal is a lobbyist, an expert advisor or a multi- national firm, that wants to maintain the status quo action a^* (e.g., an import tariff, on ongoing project, or a joint venture). There are two alternatives to the status quo, each being optimal for the agent in the matched state, but worse than the status quo if adopted in the wrong state. For example, if it knew the state (i.e., the technology), the local firm would break the joint venture with the multinational firm, but without knowing the technology it is unclear what the best course of action is.

In Example 1, $M(p) = 1+p$, $m(p) = \max(1-p, 3/4, 2p)$ and $\mathbf{w}(p) = 2\max(2p, 3/4, 1-p) - (3/4)$. The set Q^1 (where a^* is incentivized by the promise of full disclosure next period) is $[1/12, 7/12]$, and the set P (where a^* is statically optimal for the agent) is $[\frac{1}{4}, \frac{3}{8}]$. Since $\frac{3}{20} < \frac{1}{4}$, a^* is not optimal at the prior belief.

We now describe the policy τ_q for $q = 3/8$, the highest prior at which the agent finds it statically optimal to choose a^*. We argue later that the choice of $q = 3/8$ is optimal.

The initial pair of belief and (promised continuation) payoff at $t = 1$ is $(p_0, m(p_0)) = (3/20, 17/20)$. This pair is on the lower boundary of region $\mathcal{W}^2_{3/8}$, hence $\varphi(p_0, m(p_0)) = p_0$ and $\mathrm{w}(p_0) = 19/20$. In words, the policy does not provide any information to the agent, recommend action a^* and promise the continuation payoff $19/20$.

The belief-payoff pair at the start of $t = 2$ is therefore $(3/20, 19/20)$, which is in the interior of region $\mathcal{W}^2_{3/8}$. The policy splits $(3/20, 19/20)$ into

$$(\varphi(3/20, 19/20), m(\varphi(3/20, 19/20))$$

and $(1, m(1))$ with probability $\lambda(3/20, 19/20)$ and $1-\lambda(3/20, 19/20)$, respectively. Simple algebra gives $\varphi(3/20, 19/20) = 2/19$ and $\lambda(3/20, 19/20) = 19/20$.[12] The policy thus reveals that the state is ω_1 with probability $1/20$, in which case the agent takes action a_1 forever and obtains payoff 2. With the remaining probability, the agent's beliefs becomes $p = \frac{2}{19}$. In such a case, the agent is recommended to take action a^* and promised continuation payoff $\mathrm{w}(2/19) = 79/76$.

If the state has not been revealed, the belief-payoff pair at $t = 3$ is then $(2/19, 79/76)$, which is in region $\mathcal{W}^4_{3/8}$. In that region, the policy induces the belief 0, 1, and $\underline{q}^1 = \frac{1}{12}$, that is, either the state is revealed or the agent is made indifferent between playing a^* and being promised full information disclosure at the next period and playing an (statically) optimal action. In the latter case, at the start of $t = 4$ the belief-payoff pair is $(1/12, 13/12)$ which is on the upper boundary of region $\mathcal{W}^4_{3/8}$ (i.e., $13/12 = M(1/12)$), where the principal fully reveals the state.

[12]We stress that the policy describes the "equilibrium path." If the agent disobeys the recommendation, no further disclosure ever happens and the agent repeatedly plays a statically optimal action.

Figure 5 summarizes the evolution of beliefs under the policy $\tau_{3/8}$. The expected number of times the agent takes action a^* is $\frac{93}{40} = 2.325$ and the principal obtains a payoff of $251/320 \approx 0.78$.

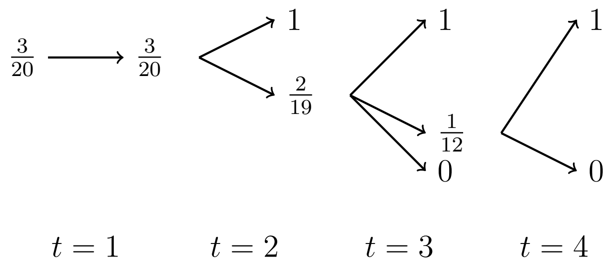

FIGURE 5. Evolution of the beliefs in Example 1

It remains to explain how to choose the parameter q^* to guarantee the optimality of τ_{q^*}.

3.5. **Construction of q^* and optimality.** For all $q \in [\underline{q}^1, \overline{q}^1]$, let $V_q : \mathcal{W} \to \mathbb{R}$ be the value function induced by the policy τ_q. For all q, note that $V_q(1, m(1)) = 0$ since a^* is not optimal at $p = 1$, and $V_q(0, m(0)) = 0$ if a^* is not optimal at $p = 0$ (resp., $= v(a^*, 0)$) if a^* is optimal at $p = 0$). Also, $V_q(\underline{q}^1, m(\underline{q}^1)) = (1 - \delta)v(a^*, \underline{q}^1)$ if $\underline{q}^1 > 0$ (resp., $V_q(0, m(0)) = v(a^*, 0)$ if $\underline{q}^1 = 0$, since a^* is then optimal at $p = 0$). Therefore, any two policies τ_q and $\tau_{q'}$ induce the same values at all $(p, w) \in \mathcal{W}_q^1 \cup \mathcal{W}_q^4 = \mathcal{W}_{q'}^1 \cup \mathcal{W}_{q'}^4$. (Remember that the regions $\mathcal{W}_q^1$ and $\mathcal{W}_q^4$ do not vary with q – see Figure 3.)

Similarly, any two policies τ_q and $\tau_{q'}$ induce the same values at all $(p, w) \in \mathcal{W}^2_{\min(q,q')}$. Thus, in particular, τ_q and $\tau_{\overline{q}^1}$ induce the same values at all $(p, w) \in \mathcal{W} \setminus \mathcal{W}_q^3$. Finally, at all $(p, w) \in \mathcal{W}_q^3$, $V_q(p, w) = \frac{1-p}{1-q} V_q(q, m(q)) = \frac{1-p}{1-q} V_{\overline{q}^1}(q, m(q))$. (See Section A.4 for more details.)

Recall that V^* is the unique solution to the fixed-point problem – to be optimal, a policy must therefore induce the value function V^*. Let

$$q^* = \sup \left\{ p \in [\underline{q}^1, \overline{q}^1] : V_{\overline{q}^1}(p, m(p)) \geq V_{\overline{q}^1}(p, w) \text{ for all } w \right\}.$$

We are now ready to state our main result.

Theorem 2. *The policy τ_{q^*} is optimal: $V_{q^*} = V^*$.*

To understand the role of q^*, recall that for all $p \in [q^*, 1]$, the policy leaves rents to the agent.[13] To minimize these rents, the principal therefore would like to have q^* as high as possible, i.e., equal to $\overline{q}^1$, the highest belief at which the agent is willing to play a^* in exchange for full information disclosure at the next period. However, $V_{\overline{q}^1}(\cdot, m(\cdot))$ is not guaranteed to be concave in p, a necessary condition for optimality. To see that $V^*(\cdot, m(\cdot))$ must be concave in p, consider any pair $(p, p') \in [0,1] \times [0,1]$ and $\alpha \in [0,1]$. We have

$$\begin{aligned}\alpha V^*(p, m(p)) + (1-\alpha)V^*(p', m(p')) &\leq V^*(\alpha p + (1-\alpha)p', \alpha m(p) + (1-\alpha)m(p')) \\ &\leq V^*(\alpha p + (1-\alpha)p', m(\alpha p + (1-\alpha)p')),\end{aligned}$$

where the first inequality follows from the concavity of V^* in both arguments and the second from V^* decreasing in w and the convexity of m. The optimal choice of q^* is thus the largest q, which guarantees $V_q(\cdot, m(\cdot))$ to be concave.

More precisely, as we show in Section A.5, the definition of q^* guarantees that V_{q^*} is concave in both arguments and decreasing in w, so that $V_{q^*}(\cdot, m(\cdot))$ is a concave function of p. We also prove that $V_{q^*}(p, m(p)) \geq V_{\overline{q}^1}(p, m(p))$ for all p. Since it is clearly the smallest such function, V_{q^*} is the concavification of $V_{\overline{q}^1}$. In particular, $q^* = \overline{q}^1$ if $V_{\overline{q}^1}(\cdot, m(\cdot))$ is already concave. Figure 6 illustrates the concavification in the context of Example 1. In black is the value function of policy $\tau_{\overline{q}^1}$; in red its concavification – the value function of policy τ_{q^*}, with $q^* = \frac{3}{8}$.

The policy τ_{q^*} leaves rents to the agent, that is, the (ex-ante) participation constraint does not bind, for all priors in $[0, \underline{q}^1) \cup (q^*, 1]$. This is quite natural for all priors in $[0,1] \ \setminus Q^1$ since the agent cannot be incentivized to play a^* even once. In the language of Ely and Szydlowski (2020), "the goalposts need to move," that is, one needs to disclose information at the ex-ante stage to persuade the agent to play a^*. However, our policy also leaves rents for all priors in $(q^*, \overline{q}^1]$. The intuitive reason is that the initial information disclosure reduces the cost of incentivizing the agent in subsequent periods sufficiently

[13]That is, the agent is promised a payoff of $\frac{1-p}{1-q^*}m(q^*) + \frac{p-q^*}{1-q^*}m(1) > m(p)$.

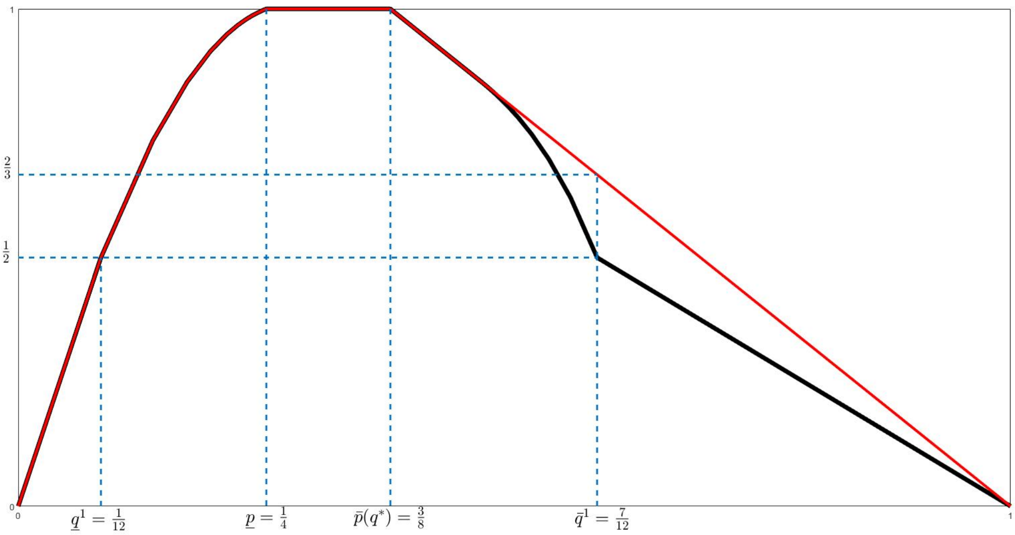

FIGURE 6. The concavification of $V_{\overline{q}^1}(\cdot, m(\cdot))$ in Example 1

enough to compensate for the initial loss. (When the realized posterior is 1, the agent never plays a^*, thus creating the loss.)

4. Optimal Policy: Properties and comparison with alternatives

4.1. **Evolution of the beliefs.** The optimal policy discloses information gradually over time, with beliefs evolving until either the agent learns the state or believes that a^* is (statically) optimal. We can be more specific. First, we consider the instances when the policy converges with positive probability to a belief $p \in P = [\underline{p}, \overline{p}]$, the set of beliefs at which a^* is statically optimal. Let $Q^\infty = [\underline{p}, \overline{q}^\infty]$, with $\overline{q}^\infty$ the solution to

$$m(\overline{q}^\infty) = (1-\delta)u(a^*, \overline{q}^\infty) + \delta\left(\frac{1-\overline{q}^\infty}{1-\underline{p}}m(\underline{p}) + \frac{\overline{q}^\infty - \underline{p}}{1-\underline{p}}m(1)\right),$$

if P is non-empty, and $Q^\infty = \emptyset$, otherwise. Note that $P \subseteq Q^\infty$. See Figure 7 for a graphical illustration.

Intuitively, the set Q^∞ has the "fixed-point property," that is, if one starts with a belief $p \in Q^\infty$ and promised utility $\mathrm{w}(p)$, then the belief $\varphi(p, \mathrm{w}(p)) \in Q^\infty$. To see this, note that the pair $(p, \mathrm{w}(p))$ is in region $\mathcal{W}^2_q$. Since $\varphi(p, \mathrm{w}(p)) \leq p$ (with a strict inequality if $p \notin P$), we then have a decreasing sequence of beliefs converging to an element in P. This is because, at all beliefs $p \in Q^\infty$, the

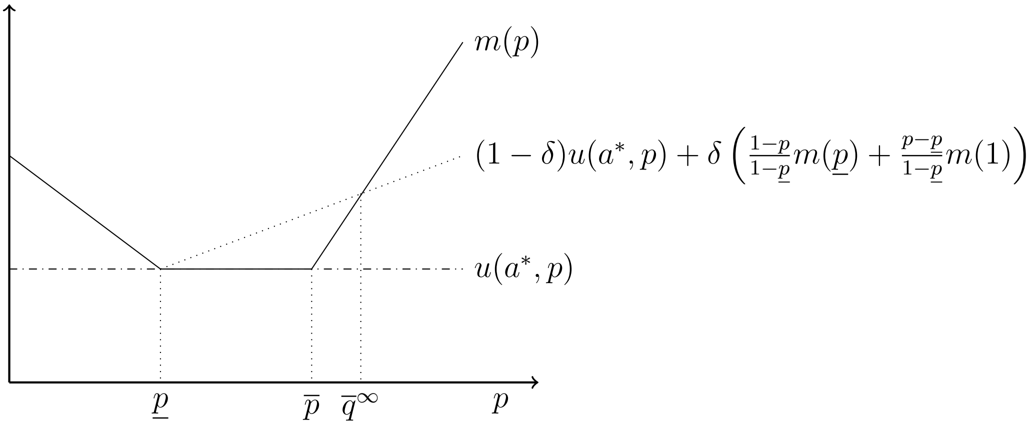

FIGURE 7. Construction of $\overline{q}^\infty$

policy splits p into $p' = \varphi(p, \mathrm{w}(p))$ and 1, then splits p' into $p'' = \varphi(p', \mathrm{w}(p'))$ and 1, etc. The decreasing sequence $(p, p', p'', \dots)$ converges, either in finite time or asymptotically, to a belief in P, at which no further splitting occurs and the agent plays a^* forever. See panel (B) of Figure 1 for an an illustration.

Recall that if the prior p_0 is larger than q^*, the policy first splits p_0 into q^* and 1. Hence, if $q^* \le \overline{q}^\infty$, the agent's belief enters the set Q^∞ with strictly positive probability.[14] Therefore, if the agent's prior beliefs are in the set $Q^\infty_{q^*}$, then the agent will choose action a^* forever with positive probability, where

$$Q^\infty_{q^*} := \begin{cases} Q^\infty & \text{if } q^* > \overline{q}^\infty, \\ [\underline{p}, 1) & \text{otherwise.} \end{cases}$$

Second, at all priors in $[0,1] \setminus Q^\infty$, there exists $T_\delta < \infty$ such that the belief process is absorbed in the degenerate beliefs 0 or 1 after at most T_δ periods. In other words, the agent learns the state for sure in finite time. The number of periods T_δ corresponds to the maximal number of periods the agent can be incentivized to play a^*. We provide an explicit computation in Section A.4, in Example 1, we have $T_\delta = 3$. Moreover, the number T_δ is increasing in δ and converges to $+\infty$ as δ converges to 1. (Note that the convergence is uniform in that it does not depend on $p_0 \in [0,1] \setminus Q^\infty$.) Thus, we have the following corollary:

[14]From the definition of q^*, we have that $q^* \ge \overline{p}$ since $V_{\overline{q}^1}(p, m(p)) = u(a^*, p)$ for all $p \in P$.

Corollary 1. *Under the optimal disclosure policy τ_{q^*}, the agent chooses action a^* forever with positive probability if, and only if, $p_0 \in Q^\infty_{q^*}$.*

The interval $Q^\infty_{q^*}$ includes the sub-interval $[\underline{p}, \bar{p}]$, where the agent takes action a^* with probability one. In the complementary set $Q^\infty_{q^*} \setminus [\underline{p}, \bar{p}]$, the probability that the agent takes action a^* forever is strictly less than 1. That is, the principal discloses the state with positive probability, and with the complementary probability he lowers the agent's belief so that it converges to the region where taking action a^* is statically optimal. Convergence may be asymptotic or may happen in finite time.

We now highlight the novelty of our optimal policy by comparing it with two alternatives commonly found in the literature. (See Appendix B for a formal discussion.)

4.2. **Comparison with the KG policy.** The KG policy aims at persuading the agent to choose a^* forever by disclosing information at the initial stage only (KG stands for Kamenica-Gentzkow). In Appendix B, we show that it is optimal in all problems with two actions. In problems with three or more actions, the KG policy may, however, be strictly sub-optimal. In Example 1, it gives the principal a payoff of $3/5$, which is strictly lower than the optimal payoff of about 0.78.[15] We now discuss several reasons why this is the case.

Recall that $P = [\underline{p}, \bar{p}]$ is the set of beliefs under which it's statically optimal for the agent to choose a^*. The KG policy splits the prior beliefs of the agent so as to move the posterior into the set P at the lowest possible "cost." Once beliefs are in P, the agent takes the principal's preferred action a^* forever. The first reason why the KG policy may be strictly dominated is that this splitting may generate information value and leave "too large" a rent to the agent.

Second, even if as in Example 1 the KG policy does not generate positive information value, it may be strictly dominated when the prior belief is $p_0 \in (0, \underline{p})$. In this case, the KG policy increases the (discounted) average belief at which a^* is recommended relative to our policy. As a consequence, it recommends

[15]In Example 1, the KG policy splits the prior $p_0 = 3/20$ into 0 and $1/4$ with probability $2/5$ and $3/5$, respectively.

a^* forever with lower probability. Our optimal policy provides incentives to choose a^* more often.

Third, the KG policy may be dominated since the set P may be empty. In such a case, the KG policy cannot incentivize the agent ever to take action a^* whereas our optimal policy can. This is because our optimal policy allows the principal to condition his information disclosure on the agent's actions, which is not permitted by the KG policy. Consider the following modification of Example 1 in which the agent's payoff from action a^* is $1/2$ rather than $3/4$.

TABLE 2. Payoff table of Example 1 modified

	a_0	a_1	a^*
ω_0	$0, 1$	$0, 0$	$1, 1/2$
ω_1	$0, 0$	$0, 2$	$1, 1/2$

Action a^* is not optimal, no matter the initial beliefs of the agent. However, there are beliefs under which the promise of full disclosure at the next period induces the agent to play a^*, i.e. $Q^1 = [\frac{1}{6}, \frac{1}{2}]$. Suppose that the prior belief is $p_0 = \frac{1}{3}$. Our optimal policy is $\tau_{1/3}$ and it splits beliefs as shown on Figure 8. At all beliefs other than 0 and 1, the agent is recommended to play a^*. The principal's expected payoff is $\frac{1285}{1536}$, i.e., about 0.83.

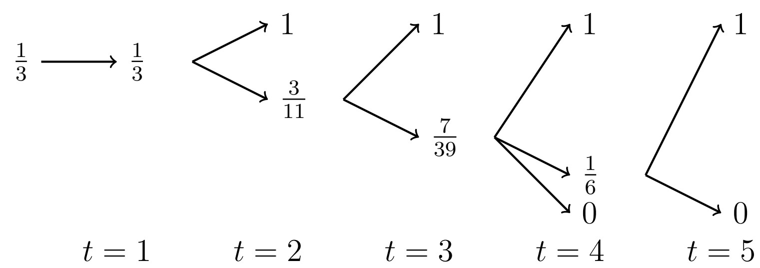

FIGURE 8. Evolution of the beliefs.

Figure 9 illustrates the value function of our optimal policy for this modification of Example 1; it is the concavification of the value function of policy $\tau_{\overline{q}^1}$.

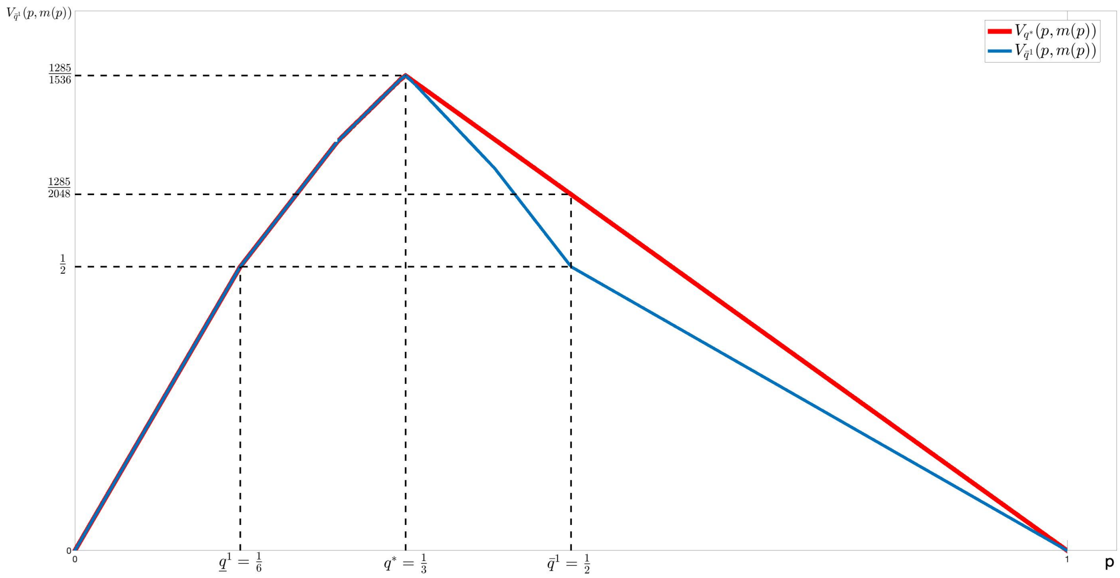

FIGURE 9. The concavification of $V_{\bar{q}^1}(\cdot, m(\cdot))$ in the modification of Example 1

4.3. **Comparison with the policy of fully disclosing the state with delay.** Another alternative policy, which plays a prominent role in the work of Ball (2019), Ely and Szydlowski (2020), and Orlov et al. (2019), is to incentivize the agent to play a^* with the promise of fully disclosing the state at a later period. The policy of fully disclosing the state with delay selects the largest integer T^* such that

$$(1-\delta)\left(u(a^*, p)(\delta^0 + \delta^1 + \cdots + \delta^{T^*-1}) + M(p)(\delta^{T^*} + \cdots)\right) \geq m(p).$$

Usually, this constraint does not bind and, therefore, the policy leaves strictly positive rent to the agent. This is the case in the modification of Example 1 presented in Table 2, where the constraint would bind for the non-integer value $\ln(5)/\ln(2)$ and hence $T^* = 2$. However, this rounding problem could be solved by adopting instead a recursive policy of *random full disclosure*, in which at each period t the principal fully discloses the state with some probability α if the agent plays a^* at period $t-1$ (and withholds all information with the complementary probability). In continuous time, the policies of random disclosure and disclosure with delay are equivalent. Intuitively, in the modification of Example 1, random disclosure performs better than disclosure with delay because it makes it possible to incentivize the agent to play a^* a discounted number of periods slightly larger than 2, namely $\ln(5)/\ln(2)$ by picking $\alpha = 1/4$.

This "integer problem" is only part of the reason for the sub-optimality of the policy of disclosure with delay, since in Example 1 the constraint binds at prior $3/20$ with $T^* = 2$. The remaining reason is that this policy (or the policy of full random disclosure) does not exploit the asymmetries in opportunity costs. In fact, we show in Appendix B that if there are no asymmetries, i.e. if $\frac{m(0)-u(a^*,0)}{v(a^*,0)} = \frac{m(1)-u(a^*,1)}{v(a^*,1)}$, then the random disclosure policy is also optimal. In Example 1, the policy of fully disclosing the state with delay does not alter the belief that the state is ω_1 when a^* is played; the belief stays fixed at the prior $p_0 = 3/20$. By contrast, at each period in which information is disclosed and a^* is played, our policy decreases the belief; the average discounted beliefs is $p^* = 133/1004 \approx 0.13 < 3/20$. Information disclosure plays two roles in our optimal policy. First, the generated information value is a carrot to motivate the agent to take action a^* in early periods. Second, information disclosure decreases the discounted average belief that the state is the high opportunity cost state ω_1 and, therefore, makes it easier to incentivize the agent to take action a^* for a longer expected time.

Appendix A. Proofs

A.1. **Mathematical preliminaries.** We collect without proofs some useful results about concave functions. Let $f : [a, b] \to \mathbb{R}$ be a concave function and $a \leq x < y < z \leq b$. The following properties hold:

(a) $\frac{f(y)-f(x)}{y-x} \geq \frac{f(z)-f(y)}{z-y}$,
(b) $\frac{f(y)-f(a)}{y-a} \geq \frac{f(z)-f(a)}{z-a}$,
(c) $\frac{f(b)-f(x)}{b-x} \geq \frac{f(b)-f(y)}{b-y}$.
(d) $\frac{f(y)-f(x)}{y-x} \geq \frac{f(y+\Delta)-f(x+\Delta)}{y-x}$ for all $\Delta \geq 0$ such that $y + \Delta \leq b$.

Note that property (a) implies (d) and is true irrespective of whether $x+\Delta \gtreqless y$. We will repeatedly use these properties in most of the following proofs.

To prove Lemma 6, we will use the following property: if $f:[a,b]\to\mathbb{R}$ satisfies $\frac{f(x)-f(a)}{x-a}\geq\frac{f(y)-f(a)}{y-a}$ for all $a<x\leq y\leq b$, then f is concave.

A.2. **Recursive formulation: Theorem 4 of Ely (2015, p. 44).** We first note that the operator T is monotone, i.e., for all $V\geq V'$, $T(V)\geq T(V')$. It also satisfies $T(V+c)\leq T(V)+\delta c$ for all positive constant $c\geq 0$, for all V. Hence, it is indeed a contraction by Blackwell's theorem.

Ely (2015) proves that the principal's maximal payoff is $\max_{w\in[m(p_0),M(p_0)]}\widehat{V}^*(p_0,w)$, with $\widehat{V}^*$ the unique fixed point of the contraction $\widehat{T}$, with $\widehat{T}$ differing from T in that the promise-keeping constraint is as an equality; all other constraints are the same. Note that the operator $\widehat{T}$ is also monotone.

We now argue that that $V^*(p_0,m(p_0))=\max_{w\in[m(p_0),M(p_0)]}\widehat{V}^*(p_0,w)$. (Note that we are not arguing that $T=\widehat{T}$.)

As a preliminary observation, note that $T(V)(p,w)\geq\widehat{T}(V)(p,w)$ for all $(p,w)\in\mathcal{W}$, for all V. Let $w_0\in\arg\max_{w\in[m(p_0),M(p_0)]}\widehat{V}^*(p_0,w)$. We have that

$$\begin{aligned}V^*(p_0,m(p_0))\geq V^*(p_0,w_0)=T(V^*)(p_0,w_0) &\geq \widehat{T}(V^*)(p_0,w_0)\geq\widehat{T}^2(V^*)(p_0,w_0)\geq\cdots\geq\\ &\geq \lim_{n\to\infty}\widehat{T}^n(V^*)(p_0,w_0)=\widehat{V}^*(p_0,w_0),\end{aligned}$$

where the first inequality follows from V^* being decreasing in w.

Conversely, let $(\lambda_s^*,p_s^*,w_s^*,a_s^*)_{s\in S}$ be a maximizer of $T(V^*)(p_0,m(p_0))$. We have that

$$M(p_0)=\sum_{s\in S}\lambda_s^*M(p_s^*)\geq\sum_{s\in S}\lambda_s^*[(1-\delta)u(a_s^*,p_s^*)+\delta w_s^*]:=\widehat{w}_0\geq\sum_{s\in S}\lambda_s^*m(p_s^*)\geq m(p_0),$$

hence $(\lambda_s^*,p_s^*,w_s^*,a_s^*)_{s\in S}$ is a maximizer for $T(\widehat{V}^*)(p_0,\widehat{w}_0)$ and, consequently,

$$V^*(p_0,m(p_0))=\widehat{V}^*(p_0,\widehat{w}_0)\leq\max_{w\in[m(p_0),M(p_0)]}\widehat{V}^*(p_0,w).$$

A.3. **Proposition 2.** We break Proposition 2 into several lemmata.

Lemma 1. *Let* $(\lambda_s, p_s, w_s, a_s)_{s\in S}$ *be a solution to the maximization program* $T(V^*)(p,w)$. *For all* $s \in S$ *such that* $\lambda_s > 0$*, we have*

$$(1-\delta)\, v\,(a_s, p_s) + \delta V^*\,(p_s, w_s) = V^*\,(p_s, (1-\delta)\, u\,(a_s, p_s) + \delta w_s)\,.$$

Proof. By contradiction, assume that there exists $s' \in S$ such that $\lambda_{s'} > 0$ and

$$(1-\delta)\, v\,(a_{s'}, p_{s'}) + \delta V^*\,(p_{s'}, w_{s'}) < V^*\,(p_{s'}, (1-\delta)\, u\,(a_{s'}, p_{s'}) + \delta w_{s'})\,.$$

Let $(\lambda^*_s, p^*_s, w^*_s, a^*_s)_{s\in S}$ be the policy, which achieves $V^*(p_{s'}, (1-\delta)u(a_{s'}, p_{s'}) + \delta w_{s'})$, and consider the new policy

$$((\lambda_s, p_s, w_s, a_s)_{s\in S\setminus\{s'\}}, (\lambda_{s'}\lambda^*_s, p^*_s, w^*_s, a^*_s)_{s\in S}).$$

By construction, the new policy is feasible. Moreover, we have that

$$\sum_{s\in S\setminus\{s'\}} \lambda_s[(1-\delta)v(a_s,p_s) + \delta V^*(p_s,w_s)] + \lambda_{s'} \sum_{s\in S} \lambda^*_s[(1-\delta)v(a^*_s,p^*_s) + \delta V^*(p^*_s,w^*_s)] =$$
$$\sum_{s\in S\setminus\{s'\}} \lambda_s[(1-\delta)v(a_s,p_s) + \delta V^*(p_s,w_s)] + \lambda_{s'} V^*(p_{s'}, (1-\delta)u(a_{s'},p_{s'}) + \delta w_{s'}) >$$
$$\sum_{s\in S} \lambda_s[(1-\delta)v(a_s,p_s) + \delta V^*(p_s,w_s)],$$

a contradiction with the optimality of $(\lambda_s, p_s, w_s, a_s)_{s\in S}$.

Since the fixed point satisfies $V^*(p_s, (1-\delta)u(a_s,p_s) + \delta w_s) \geq (1-\delta)v(a_s,p_s) + \delta V^*(p_s, w_s)$, we have the desired result. ☐

Lemma 2. *Let* $(\lambda_s, p_s, w_s, a_s)_{s\in S}$ *be a solution to the maximization program* $T(V^*)(p,w)$. *For all* $s \in S$ *such that* $\lambda_s > 0$, $V^*(p_s, w_s) = 0$ *if* $a_s \neq a^*$.

Proof. Let $s \in S$ such that $\lambda_s > 0$ and $a_s \neq a^*$. We have

$$(1-\delta)v(a_s,p_s) + \delta V^*(p_s,w_s) = \delta V^*(p_s,w_s) \geq V^*(p_s, (1-\delta)u(a_s,p_s) + \delta w_s) \geq V^*(p_s,w_s),$$

where the first inequality follows from Lemma 1 and the second follows from V^* decreasing in w and $w_s \geq u(a_s, p_s)$ for

$$(1-\delta)u(a_s,p_s) + \delta w_s \geq m(p_s),$$

to hold. It follows that $V^*(p_s, w_s) = 0$. ☐

Lemma 3. *If there exists a solution $(\lambda'_s, p'_s, w'_s, a'_s)_{s\in S'}$ to the maximization program $T(V^*)(p,w)$, then there exists a solution $(\lambda_s, p_s, w_s, a_s)_{s\in S}$ such that $a_s = a^*$ for at most one $s \in S$ with $\lambda_s > 0$.*

Proof. Let $(\lambda'_s, p'_s, w'_s, a'_s)_{s\in S'}$ be a solution to the maximization program $T(V^*)(p,w)$. Let $S^* \subseteq S'$ be the set of signals such that $a_s = a^*$ and $\lambda_s > 0$. If S^* is empty, there is nothing to prove. If S^* is non-empty, define p^* as

$$\sum_{s\in S^*}\Big(\frac{\lambda'_s}{\sum_{s\in S^*}\lambda'_s}\Big)p_s = p^*,$$

and $\sum_{s\in S^*}\lambda'_s = \lambda^*$. From the concavity of V^*, we have that

$$\begin{aligned}\sum_{s\in S^*}\lambda'_s(v(a^*,p'_s)(1-\delta)+\delta V^*(p'_s,w'_s)) &= \lambda^*\Big(v(a^*,p^*)(1-\delta)+\delta\sum_{s\in S^*}\Big(\frac{\lambda'_s}{\lambda^*}\Big)V(p'_s,w'_s)\Big)\\ &\leq \lambda^*\Big(v(a^*,p^*)(1-\delta)+\delta V(p^*,w^*)\Big),\end{aligned}$$

where

$$w^* = \sum_{s\in S^*}\Big(\frac{\lambda'_s}{\sum_{s\in S^*}\lambda'_s}\Big)w'_s.$$

Notice that $w^* \in [m(p^*), M(p^*)]$ since the convexity of m implies

$$M(p^*) = \sum_{s\in S^*}\Big(\frac{\lambda'_s}{\sum_{s\in S^*}\lambda'_s}\Big)M(p'_s) \geq \sum_{s\in S^*}\Big(\frac{\lambda'_s}{\sum_{s\in S^*}\lambda'_s}\Big)w_s \geq \sum_{s\in S^*}\Big(\frac{\lambda'_s}{\sum_{s\in S^*}\lambda'_s}\Big)m(p'_s) \geq m(p^*).$$

It is routine to verify that the new contract

$$((\lambda'_s, p'_s, w'_s, a'_s)_{s\in S'\setminus S^*}, (\lambda^*, p^*, a^*, w^*))$$

is feasible and, therefore, also optimal. □

Lemma 4. *For any $\omega \in \Omega$, denote by q_ω the degenerate belief which puts probability 1 on ω, and denote by $\hat{a}(q_\omega)$ the optimal action at belief q_ω for the agent. If there exists a solution $(\lambda'_s, p'_s, w'_s, a'_s)_{s\in S'}$ to the maximization program $T(V^*)(p,w)$, then there exists a solution $(\lambda_s, p_s, w_s, a_s)_{s\in S}$ such that for all $s \in S$ with $a_s \neq a^*$ it is $p_s = q_\omega$ and $a_s = \hat{a}(q_\omega)$ for some $\omega \in \Omega$.*

Proof. We must argue that it is without loss of generality to focus on a subset of policies such that, if a^* is not recommended at some signal s, then p_s is a degenerate belief. Consider an optimal policy $\tau = (\lambda_s, p_s, w_s, a_s)_{s\in S}$ as the solution to $T(V^*)(p,w)$. Recall that $p_s(\omega)$ is the agent's belief that the state is ω after observing signal s. By Lemma 3, we can restrict attention to a solution τ such that there is at most one signal $s^* \in S$ at which a^* is recommended. Define the alternative policy τ',

$$\tau' = \left(\left(\sum_{s \in S, s \neq s^*} \lambda_s \cdot p_s(\omega), q_\omega, m(q_\omega), \hat{a}(q_\omega) \right)_{\omega \in \Omega}, (\lambda_{s^*}, p_{s^*}, w_{s^*}, a^*) \right).$$

According to policy τ', with probability λ_{s^*}, the agent forms a posterior beliefs p_{s^*} while the principal recommends a^* and promises future payoff w_{s^*}. With probability $\sum_{s \in S, s \neq s^*} \lambda_s \cdot p_s(\omega)$, the agent knows that the state is ω for sure while the principal recommends $\hat{a}(q_\omega)$ and promises future payoff $m(q_\omega)$. It's a property of τ' that whenever a^* is not recommended at some signal, then the posterior at that signal is degenerate. We must show that τ' is an optimal policy. We begin by showing that τ' is a feasible policy. Note first that

$$\sum_{\omega \in \Omega} \sum_{s \in S, s \neq s^*} \lambda_s \cdot p_s(\omega) + \lambda_{s^*} = \sum_{s \in S, s \neq s^*} \lambda_s + \lambda_{s^*} = 1.$$

Second, for any $\omega' \in \Omega$, we have that

$$\sum_{\omega \in \Omega} \Big[\sum_{s \in S, s \neq s^*} \lambda_s \cdot p_s(\omega) \Big] q_\omega(\omega') + \lambda_{s^*} p_{s^*}(\omega') = \sum_{s \in S, s \neq s^*} \lambda_s \cdot p_s(\omega') + \lambda_{s^*} p_{s^*}(\omega') = p(\omega'),$$

where the equality holds since q_ω is a degenerate belief such that

$$q_\omega(\omega') = \begin{cases} 1 & \text{if } \omega' = \omega \\ 0 & \text{otherwise} \end{cases}.$$

Third, the following equation holds,

$$\begin{aligned} &\lambda_{s^*}[(1-\delta)u(a^*; p_{s^*}) + \delta w_{s^*}] + \sum_{\omega \in \Omega} \sum_{s \in S, s \neq s^*} \lambda_s \cdot p_s(\omega)[(1-\delta)u(\hat{a}(q_\omega); q_\omega) + \delta m(q_\omega)] \\ =&\lambda_{s^*}[(1-\delta)u(a^*; p_{s^*}) + \delta w_{s^*}] + \sum_{\omega \in \Omega} \sum_{s \in S, s \neq s^*} \lambda_s \cdot p_s(\omega) m(q_\omega) \\ =&\lambda_{s^*}[(1-\delta)u(a^*; p_{s^*}) + \delta w_{s^*}] + \sum_{s \in S, s \neq s^*} \lambda_s \cdot M(p_s) \\ \geq&\lambda_{s^*}[(1-\delta)u(a^*; p_{s^*}) + \delta w_{s^*}] + \sum_{s \in S, s \neq s^*} \lambda_s \cdot [(1-\delta)u(a_s; p_s) + \delta w_s \geq w. \end{aligned}$$

This shows that τ' is feasible, as it satisfies all the constraints in the optimization problem with value $T(V^*)(p,w)$. It remains to show that it is optimal. By Lemma 2, when a^* is not recommended at signal s, i.e., $a_s \neq a^*$, it is $V^*(p_s, w_s) = 0$. Therefore, we have that

$$\begin{aligned} V^*(p,w) &= \sum_{s\in S} \lambda_s[(1-\delta)v(a_s;p_s) + \delta V^*(p_s;w_s)] \\ &= \lambda_{s^*}[(1-\delta)v(a^*;p_{s^*}) + \delta V^*(p_{s^*};w_{s^*})], \end{aligned}$$

which is also the value that the principal achieves under policy τ', which is then optimal. □

Lemma 5. *If there exists a solution $(\lambda'_s, p'_s, w'_s, a'_s)_{s\in S'}$ to the maximization program $T(V^*)(p,w)$, then there exists a solution $(\lambda_s, p_s, w_s, a_s)_{s\in S}$ such that*

$$(1-\delta)u(a_s,p_s) + \delta w_s = m(p_s),$$

for all s such that $\lambda_s > 0$ and $a_s = a^*$.

Proof. By Lemma 3, there exists a solution $(\lambda_s, p_s, w_s, a_s)_{s\in S}$ such that there is at most one $s^* \in S$ such that $a_{s^*} = a^*$. If $\frac{m(p_{s^*})-(1-\delta)u(a^*;p_{s^*})}{\delta} = w_s^*$, then the lemma holds. Hence from now on we consider the case:

$$w_s^* > \frac{m(p_{s^*}) - (1-\delta)u(a^*;p_{s^*})}{\delta}. \tag{4}$$

By Lemma 1, we have that

$$V^*(p_{s^*}, (1-\delta)u(a^*,p_{s^*}) + \delta w_{s^*}) = (1-\delta)v(a^*,p_{s^*}) + \delta V^*(p_{s^*}, w_{s^*}).$$

Define $\bar{w} \geq w_{s^*}$ as follows:

$$\bar{w} := \sup\left\{ w \in [m(p_{s^*}), (1-\delta)u(a^*,p_{s^*}) + \delta M(p_{s^*})] : V^*(p_{s^*}, w) = (1-\delta)v(a^*,p_{s^*}) + \delta V^*\left(p_{s^*}, \frac{w-(1-\delta)u(a^*,p_{s^*})}{\delta}\right)\right\}.$$

We now show that, for all $\forall w' \in [m(p_{s^*}), \bar{w})$ it is

$$V^*(p_{s^*}, w') = (1-\delta)v(a^*,p_{s^*}) + \delta V^*\left(p_{s^*}, \frac{w'-(1-\delta)u(a^*,p_{s^*})}{\delta}\right). \tag{5}$$

By the definition of $\bar{w}$, for all $w' < \bar{w}$ there exists $w^o \in (w', \bar{w}]$ such that

$$V^*(p_{s^*}, w^o) = (1-\delta)v(a^*, p_{s^*}) + \delta V^* \left(p_{s^*}, \frac{w^o - (1-\delta)u(a^*, p_{s^*})}{\delta} \right). \quad \textbf{(6)}$$

Take any w', w'' such that $m(p_{s^*}) \leq w' \leq w'' \leq (1-\delta)u(a^*, p_{s^*}) + \delta M(p_{s^*})$. Note that $w'' \leq \frac{w''-(1-\delta)u(a^*;p_{s^*})}{\delta}$ and $w' \leq \frac{w'-(1-\delta)u(a^*;p_{s^*})}{\delta} \leq \frac{w''-(1-\delta)u(a^*;p_{s^*})}{\delta}$, since $u(a^*; p_{s^*}) \leq m(p_{s^*})$. Hence, we can find $\alpha, \beta \leq [0,1]$ such that

$$\frac{w' - (1-\delta)u(a^*; p_{s^*})}{\delta} = \alpha w' + (1-\alpha)\frac{w'' - (1-\delta)u(a^*; p_{s^*})}{\delta}$$

$$w'' = \beta w' + (1-\beta)\frac{w'' - (1-\delta)u(a^*; p_{s^*})}{\delta}$$

Moreover, since for any w', w'',

$$\begin{aligned} w' + \delta\frac{w'' - (1-\delta)u(a^*; p_{s^*})}{\delta} &= w'' + \delta\frac{w' - (1-\delta)u(a^*; p_{s^*})}{\delta} \\ &= \begin{matrix} \left[\beta w' + (1-\beta)\frac{w'' - (1-\delta)u(a^*; p_{s^*})}{\delta}\right] \\ +\delta\left[\alpha w' + (1-\alpha)\frac{w'' - (1-\delta)u(a^*; p_{s^*})}{\delta}\right] \end{matrix} \\ &= (\beta + \delta\alpha)w' + [(1-\beta) + \delta(1-\alpha)]\frac{w'' - (1-\delta)u(a^*; p_{s^*})}{\delta} \end{aligned}$$

we have that

$$\begin{cases} (\beta + \delta\alpha) = 1 \\ [(1-\beta) + \delta(1-\alpha)] = \delta \end{cases}$$

Therefore it is

$$\begin{aligned} &V^*(p_{s^*}; w'') + \delta V^* \left(p_{s^*}; \frac{w' - (1-\delta)u(a^*; p_{s^*})}{\delta} \right) \\ &= V^* \left(p_{s^*}; \beta w' + (1-\beta)\frac{w'' - (1-\delta)u(a^*; p_{s^*})}{\delta} \right) + \delta V^* \left(p_{s^*}; \alpha w' + (1-\alpha)\frac{w'' - (1-\delta)u(a^*; p_{s^*})}{\delta} \right) \\ &\geq (\beta + \delta\alpha)V^*(p_{s^*}; w') + [(1-\beta) + \delta(1-\alpha)]V^* \left(p_{s^*}; \frac{w'' - (1-\delta)u(a^*; p_{s^*})}{\delta} \right) \\ &= V^*(p_{s^*}; w') + \delta V^* \left(p_{s^*}; \frac{w'' - (1-\delta)u(a^*; p_{s^*})}{\delta} \right), \end{aligned} \quad \textbf{(7)}$$

where the inequality follows from the concavity of V^* with respect to w.

Moreover, by (6), for all $w' \in [m(p_{s^*}), \bar{w})$ there exists $w^o \in (w', \bar{w}]$ such that

$$\begin{aligned}(1-\delta)v(a^*;p_{s^*}) &= V^*(p_{s^*};w^o) - \delta V^*\left(p_{s^*};\frac{w^o-(1-\delta)u(a^*;p_{s^*})}{\delta}\right)\\ &\geq V^*(p_{s^*};w') - \delta V^*\left(p_{s^*};\frac{w'-(1-\delta)u(a^*;p_{s^*})}{\delta}\right),\end{aligned}$$

where the inequality follows from (7), and is equivalent to

$$V^*(p_{s^*},w') \leq (1-\delta)v(a^*,p_{s^*}) + \delta V^*\left(p_{s^*},\frac{w'-(1-\delta)u(a^*,p_{s^*})}{\delta}\right).$$

Since V^* is a value function, it is also

$$V^*(p_{s^*};w') \geq (1-\delta)v(a^*;p_{s^*}) + \delta V^*\left(p_{s^*};\frac{w'-(1-\delta)u(a^*;p_{s^*})}{\delta}\right).$$

Hence, (5) holds and for all $w' \in [m(p_{s^*}), \bar{w})$ it is

$$V^*(p_{s^*},w') = (1-\delta)v(a^*,p_{s^*}) + \delta V^*\left(p_{s^*},\frac{w'-(1-\delta)u(a^*,p_{s^*})}{\delta}\right),$$

It follows that, for all $w', w'' \in [m(p_{s^*}), \bar{w})$, we have

$$\begin{aligned}(1-\delta)v(a^*;p_{s^*}) &= V^*(p_{s^*};w') - \delta V^*\left(p_{s^*};\frac{w'-(1-\delta)u(a^*;p_{s^*})}{\delta}\right)\\ &= V^*(p_{s^*};w'') - \delta V^*\left(p_{s^*};\frac{w''-(1-\delta)u(a^*;p_{s^*})}{\delta}\right)\end{aligned}$$

or

$$V^*(p_{s^*},w')+\delta V^*\left(p_{s^*},\frac{w''-(1-\delta)u(a^*,p_{s^*})}{\delta}\right) = V^*(p_{s^*},w'')+\delta V^*\left(p_{s^*},\frac{w'-(1-\delta)u(a^*,p_{s^*})}{\delta}\right). \tag{8}$$

We now prove by contradiction that, for $w \in \left[m(p_{s^*}), \frac{\bar{w}-(1-\delta)u(a^*,p_{s^*})}{\delta}\right)$, $V^*(p_{s^*},\cdot)$ is linear in w. Suppose, to the contrary, that it is not; then, since $V^*(p_{s^*},\cdot)$ is concave in w, there exists $\bar{w}' \in \left(\frac{m(p_{s^*})-(1-\delta)u(a^*,p_{s^*})}{\delta}, \frac{\bar{w}-(1-\delta)u(a^*,p_{s^*})}{\delta}\right)$ such that, for all $\alpha \in (0,1)$

$$\alpha V^*(p_{s^*},m(p_{s^*})) + (1-\alpha)V^*(p_{s^*},\bar{w}') < V^*(p_{s^*},\alpha m(p_{s^*}) + (1-\alpha)\bar{w}').$$

Since $\bar{w}' > \frac{m(p_{s^*})-(1-\delta)u(a^*,p_{s^*})}{\delta}$, we then have that

$$\begin{cases} (1-\delta)u(a^*,p_{s^*})+\delta\bar{w}' \in (m(p_{s^*}),\bar{w}') \\ \frac{m(p_{s^*})-(1-\delta)u(a^*,p_{s^*})}{\delta} \in (m(p_{s^*}),\bar{w}') \end{cases}.$$

Therefore, there exists $\gamma,\eta \in (0,1)$ such that

$$\begin{cases} (1-\delta)u(a^*,p_{s^*})+\delta\bar{w}' = \eta m(p_{s^*})+(1-\eta)\bar{w}' \\ \frac{m(p_{s^*})-(1-\delta)u(a^*,p_{s^*})}{\delta} = \gamma m(p_{s^*})+(1-\gamma)\bar{w}' \end{cases}. \tag{9}$$

and

$$\begin{cases} V^*(p_{s^*},(1-\delta)u(a^*,p_{s^*})+\delta\bar{w}') > \eta V^*(p_{s^*},m(p_{s^*}))+(1-\eta)V^*(p_{s^*},\bar{w}') \\ \delta\left[V^*(p_{s^*},\frac{m(p_{s^*})-(1-\delta)u(a^*,p_{s^*})}{\delta})\right] > \delta\left[\gamma V^*(p_{s^*},m(p_{s^*}))+(1-\gamma)V^*(p_{s^*},\bar{w}')\right] \end{cases}.$$

Adding up lhs and rhs of the two inequalities, we obtain

$$\begin{aligned} & V^*(p_{s^*},(1-\delta)u(a^*,p_{s^*})+\delta\bar{w}')+\delta V^*\left(p_{s^*},\frac{m(p_{s^*})-(1-\delta)u(a^*,p_{s^*})}{\delta}\right) \\ & > (\eta+\delta\gamma)V^*(p_{s^*},m(p_{s^*}))+[(1-\eta)+\delta(1-\gamma)]V^*(p_{s^*},\bar{w}') \\ & = V^*(p_{s^*},m(p_{s^*}))+\delta V^*(p_{s^*},\bar{w}'), \end{aligned} \tag{10}$$

where the last equality follows from the fact that, by (9), $(\eta+\delta\gamma)=1$ and $[(1-\eta)+\delta(1-\gamma)]=\delta$, since

$$\begin{aligned} (1-\delta)u(a^*,p_{s^*})+\delta\bar{w}'+\delta\frac{m(p_{s^*})-(1-\delta)u(a^*,p_{s^*})}{\delta} &= m(p_{s^*})+\delta\bar{w}' \\ &= (\eta+\delta\gamma)m(p_{s^*})+[(1-\eta)+\delta(1-\gamma)]\bar{w}'. \end{aligned}$$

Letting $w''=(1-\delta)u(a^*,p_{s^*})+\delta\bar{w}'$ and $w'=m(p_{s^*})$, inequality (10) can be written as

$$V^*(p_{s^*},w'')+\delta V^*\left(p_{s^*},\frac{w'-(1-\delta)u(a^*,p_{s^*})}{\delta}\right) > V^*(p_{s^*},w')+\delta V^*\left(p_{s^*},\frac{w''-(1-\delta)u(a^*,p_{s^*})}{\delta}\right),$$

which contradicts (8) and hence proves that, for $w \in \left[m(p_{s^*}),\frac{\bar{w}-(1-\delta)u(a^*,p_{s^*})}{\delta}\right)$, $V^*(p_{s^*},\cdot)$ is linear in w.

Recall that, by Lemma 3, there exists a solution $\tau = (\lambda_s, p_s, w_s, a_s)_{s\in S}$ such that there is at most one $s^* \in S$ such that $a_{s^*} = a^*$. We now define an alternative policy τ' associating tuple $(\lambda'_s, p'_s, w'_s, a'_s)_{s\in S'}$ to each (p, w) and show that it is also optimal. First, pick any $\hat{w} \in \left(\bar{w}, \frac{\bar{w}-(1-\delta)u(a^*,p_{s^*})}{\delta}\right)$ and denote by $(\hat{\lambda}_s, \hat{p}_s, \hat{w}_s, \hat{a}_s)_{s\in\hat{S}}$ a solution to $T(V^*)(p_{s^*}, \hat{w})$. Note that, by the definition of $\bar{w}$, $\left(1, p_{s^*}, \frac{\hat{w}-(1-\delta)u(a^*,p_{s^*})}{\delta}, a^*\right)$ is not a solution to $T(V^*)(p_{s^*}, \hat{w})$. The alternative disclosure policy τ' is defined as follows, starting from any state (p, w): (i) For all signals $s \in S\backslash\{s^*\}$ the policy is the same as the solution $\tau = (\lambda_s, p_s, w_s, a_s)_{s\in S}$: signal s is send with probability λ_s and is associated with beliefs p_s, promised utility w_s and recommended action a_s; (ii) signal s^* is sent with probability $\lambda_{s^*}\left(\frac{\hat{w}-[(1-\delta)u(a^*,p_{s^*})+\delta w_{s^*}]}{\hat{w}-m(p_{s^*})}\right)$ and is associated with beliefs p_{s^*}, promised utility $m(p_{s^*})$, and recommended action a^*; (iii) with the remaining probability the solution $(\hat{\lambda}_s, \hat{p}_s, \hat{w}_s, \hat{a}_s)_{s\in\hat{S}}$ to $T(V^*)(p_{s^*}, \hat{w})$ is adopted. More precisely, signal $\hat{s} \in \hat{S}$ is send with probability $\lambda_{s^*}\hat{\lambda}_{\hat{s}}\left(\frac{(1-\delta)u(a^*,p_{s^*})+\delta w_{s^*}-m(p_{s^*})}{\hat{w}-m(p_{s^*})}\right)$ and is associated with beliefs $\hat{p}_{\hat{s}}$, promised utility $\hat{w}_{\hat{s}}$ and recommended action $\hat{a}_{\hat{s}}$.

We now argue that the policy τ' is a solution to $T(V^*)(p, w)$. First, for signals $s \in S\backslash\{s^*\}$, τ' coincides with the solution $\tau = (\lambda_s, p_s, w_s, a_s)_{s\in S}$. Second, since $V^*(p_{s^*}, w)$ is linear in $w \in \left[m(p_{s^*}), \frac{\bar{w}-(1-\delta)u(a^*,p_{s^*})}{\delta}\right)$ and τ' splits what the solution $\tau = (\lambda_s, p_s, w_s, a_s)_{s\in S}$ does at s^*, we only need to check that for the signals s^* and $\hat{s}$, τ' is a solution. This is the case because $(\hat{\lambda}_s, \hat{p}_s, \hat{w}_s, \hat{a}_s)_{s\in\hat{S}}$ is a solution to $T(V^*)(p_{s^*}, \hat{w})$ and, for signal s^* and state $(p_{s^*}, m(p_{s^*}))$,

$$V^*(p_{s^*}, m(p_{s^*})) = (1-\delta)v(a^*; p_{s^*}) + \delta V^*\left(p_{s^*}, \frac{m(p_{s^*}) - (1-\delta)u(a^*, p_{s^*})}{\delta}\right),$$

which implies that $\left(1, p_{s^*}, \frac{m(p_{s^*})-(1-\delta)u(a^*,p_{s^*})}{\delta}, a^*\right)$ is a solution to $T(V^*)(p_{s^*}, m(p_{s^*}))$. By Lemma 3, besides τ' there also exists a solution $\tau'' = (\lambda''_s, p''_s, w''_s, a''_s)_{s\in S''}$ such that there is at most one $s^* \in S''$ at which $a''_{s^*} = a^*$ and $\lambda''_{s^*} = \lambda'_{s^*} < \lambda_{s^*}$. We can then define an alternative optimal policy τ''' in the same way we have defined τ'. By Lemma 3, we can find a new solution such that there is at most one signal s^* at which a^* is recommended, and by Lemma 4, we can guarantee that the principal sends at most $|\Omega| + 1$ signals. Call such a solution $\tau^1 := ((\lambda^1_\omega, q_\omega, m(q_\omega), \hat{a}(q_\omega))_{\omega\in\Omega}, (\lambda^1_{s^*}, p^1_{s^*}, w^1_{s^*}, s^*))$. According to τ^1, after signal s^* the agent's belief is $p^1_{s^*}$, the principal recommends a^* and promises future payoff $w^1_{s^*}$. For any $\omega \in \Omega$, there exists a signal where the agent believes the state is ω for sure, the principal recommends the agent's optimal action $\hat{a}(q_\omega)$ and promises future payoff $m(q_\omega)$.

Iterating this process leads to defining a sequence of optimal policies $\{\tau^n\}_{n\geq 1}$, such that $\lambda^n_{s^*}$ is decreasing, while λ^n_ω is increasing, in n. Therefore, the sequences $\{\lambda^n_{s^*}\}$ and $\{\lambda^n_\omega\}$ must converge. Furthermore, the sequences are strictly decreasing and strictly increasing unless they converge in a finite number of steps. Denote by λ_{s^*} and λ_ω their limits and denote by p_{s^*} the solution of the Bayesian plausibility constraint

$$\lambda_{s^*} p_{s^*} + \sum_{\omega \in \Omega} \lambda_\omega = p.$$

We now argue that the policy sequence $\{\tau^n\}$ converges to the limit policy τ^∞,

$$\tau^\infty := \left((\lambda_\omega, q_\omega, m(q_\omega), \hat{a}(q_\omega))_{\omega\in\Omega}, \left(\lambda_{s^*}, p_{s^*}, \frac{m(p_{s^*}) - (1-\delta)u(a^*; p_{s^*})}{\delta}, a^* \right) \right)$$

Under this policy, when a^* is recommended at signal s^*, the principal promises future payoff $w^\infty_{s^*} = \frac{m(p_{s^*})-(1-\delta)u(a^*;p_{s^*})}{\delta}$, and it is thus the case that, as stated by the lemma,

$$(1-\delta)u(a^*; p_{s^*}) + \delta w^\infty_{s^*} = m(p_{s^*}).$$

As we have already argued that $\{\lambda^n_{s^*}\}$ and $\{\lambda^n_\omega\}$ converge to λ_{s^*} and λ_ω, to conclude the proof it only remains to show that $\{w^n_{s^*}\}$ converges to $w^\infty_{s^*}$. By (4) and the definition of τ^n, unless $\{\lambda^n_{s^*}\}$ and $\{\lambda^n_\omega\}$ have converged in a finite number of steps less than n, it is

$$w^n_{s^*} > \frac{m(p^n_{s^*}) - (1-\delta)u(a^*, p^n_{s^*})}{\delta},$$

and $\lambda^n_{s^*}$ is strictly decreasing. Hence in the limit, it must be the case that

$$w^\infty_{s^*} = \frac{m(p_{s^*}) - (1-\delta)u(a^*; p_{s^*})}{\delta}.$$

. □

A.4. **Value functions, $\Omega = \{\omega_0, \omega_1\}$.** This section characterizes the value function V_q induced by the policy τ_q in the binary case. As explained in the text, it is enough to characterize $V_{\overline{q}^1}$. We first start with the definition of important subsets of $[0,1]$.

A.4.1. *Construction of the sets* Q^k. Let $Q^0 := [0,1]$. We define inductively the set $Q^k \subseteq [0,1]$, $k \geq 0$. We write $\underline{q}^k$ (resp., $\overline{q}^k$) for $\inf Q^k$ (resp., $\sup Q^k$). For any $k \geq 0$, define the function $U^k : [\underline{q}^k, 1] \to \mathbb{R}$:

$$U^k(q) := \frac{1-q}{1-\underline{q}^k} m(\underline{q}^k) + \frac{q-\underline{q}^k}{1-\underline{q}^k} m(1),$$

with the convention that $U^k \equiv m(1)$ if $\underline{q}^k = 1$. Note that $U^0(q) = M(q)$ and $U^k(q) \geq m(q)$ for all k. We define Q^{k+1} as follows:

$$Q^{k+1} = \{q \in Q^k : (1-\delta)u(a^*, q) + \delta U^k(q) \geq m(q)\}.$$

For a graphical illustration, see Figure 10.

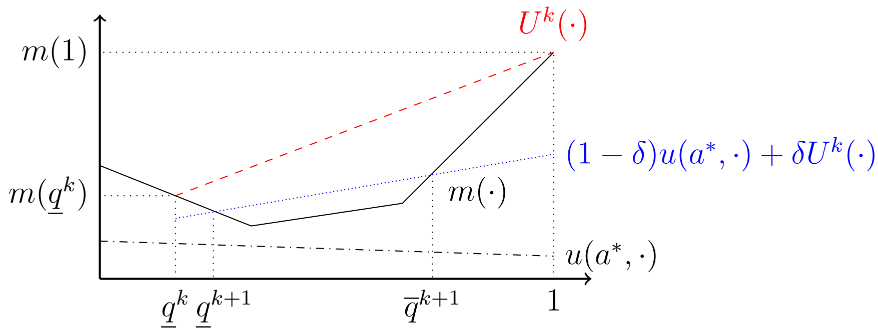

FIGURE 10. Construction of the thresholds

Few observations are worth making. First, we have that $P \subseteq Q^k$ for all k. Second, we have a decreasing sequence, i.e., $Q^{k+1} \subseteq Q^k$ for all k. Third, if Q^k and P are non-empty, then they are closed intervals. Fourth, the limit $Q^\infty = \lim_{k\to\infty} Q^k = \bigcap_k Q^k$ exists and includes P. Moreover, if $P \neq \emptyset$, then $\underline{q}^\infty = \underline{p}$, where $\underline{p} := \inf P$. If $P = \emptyset$, then $Q^\infty = \emptyset$. Consequently, there exists $k^* < \infty$ such that $\emptyset = Q^{k^*+1} \subset Q^{k^*} \neq \emptyset$.

The first to the third observations are readily proved, so we concentrate on the proof of the fourth observation. The limit exists as we have a decreasing sequence of sets.

We prove that if $P = \emptyset$, then $Q^\infty = \emptyset$. So, assume that $P = \emptyset$. We first argue that it cannot be that $Q^k = Q^{k-1} \neq \emptyset$ for some $k \geq 0$. To the contrary, assume that $Q^k = Q^{k-1} \neq \emptyset$ for some $k \geq 0$, hence $Q^{k'} = Q^{k-1}$ for all $k' \geq k$. From the convexity and continuity of m and the linearity of u, Q^{k-1} is the closed interval $[\underline{q}^{k-1}, \overline{q}^{k-1}]$, with the two boundary points solution to

$$(1-\delta)u(a^*, q) + \delta U^{k-2}(q) = m(q).$$

Therefore, if $(\underline{q}^k, \overline{q}^k) = (\underline{q}^{k-1}, \overline{q}^{k-1})$, we have that:

$$\begin{aligned}
m(\underline{q}^{k-1}) &= (1-\delta)u(a^*, \underline{q}^{k-1}) + \delta m(\underline{q}^{k-1}), \\
m(\overline{q}^{k-1}) &= (1-\delta)u(a^*, \overline{q}^{k-1}) + \delta \Big[\frac{1-\overline{q}^{k-1}}{1-\underline{q}^{k-1}} m(\underline{q}^{k-1}) + \frac{\overline{q}^{k-1} - \underline{q}^{k-1}}{1-\underline{q}^{k-1}} m(1)\Big], \\
&\leq (1-\delta)u(a^*, \overline{q}^{k-1}) + \delta m(\overline{q}^{k-1}).
\end{aligned}$$

This implies that $u(a^*, \underline{q}^{k-1}) = m(\underline{q}^{k-1})$ and $u(a^*, \overline{q}^{k-1}) = m(\overline{q}^{k-1})$ and, therefore, $\emptyset \neq Q^{k-1} \subseteq P$, a contradiction.

We thus have an infinite sequence of strictly decreasing non-empty closed intervals. Let $\varepsilon := \min_{p \in [0,1]} m(p) - u(a^*, p)$. Since $P = \emptyset$, we have that $\varepsilon > 0$. For all $p \in Q^\infty$, for all k,

$$\begin{aligned}
m(p) &\leq (1-\delta)u(a^*, p) + \delta U^k(p), \\
&\leq (1-\delta)(m(p) - \varepsilon) + \delta U^k(p).
\end{aligned}$$

Assume that Q^∞ is non-empty and let $\underline{q}^\infty$ its greatest lower bound. Since $\underline{q}^\infty \in Q^k$ for all k, we have that $U^k(\underline{q}^\infty) \geq m(\underline{q}^\infty) + \varepsilon(1-\delta)/\delta$ for all k. Since $\lim_{k\to\infty} U^k(\underline{q}^\infty) = m(\underline{q}^\infty)$, we have that $m(\underline{q}^\infty) \geq m(\underline{q}^\infty) + \varepsilon(1-\delta)/\delta$, a contradiction.

We now prove that if $P \neq \emptyset$, then $\underline{q}^\infty = \underline{p}$. From above, we have that if $Q^k = Q^{k-1} \neq \emptyset$ for some $k \geq 0$, hence $Q^{k'} = Q^{k-1}$ for all $k' \geq k$, then $P = Q^k$ since $P \subseteq Q^k$. If we have an infinite sequence of strictly decreasing sets, for all $q \in Q^\infty$,

$$(1-\delta)u(a^*, q) + \delta\Big[\frac{1-q}{1-\underline{q}^\infty} m(\underline{q}^\infty) + \frac{q-\underline{q}^\infty}{1-\underline{q}^\infty} m(1)\Big] \geq m(q).$$

Taking the limit $q \downarrow \underline{q}^\infty$, we obtain that $u(a^*, \underline{q}^\infty) = m(\underline{q}^\infty)$, i.e., $\underline{q}^\infty \in P$. Hence, $\underline{q}^\infty = \underline{p}$.

A.4.2. *Value functions.* We first derive $V_{\overline{q}^1}$ for all $(p, w) \in \mathcal{W} \setminus \mathcal{W}^2_{\overline{q}^1}$.

To start with, $V_{\overline{q}^1}(1, m(1)) = 0$ since a^* is not optimal at $p = 1$. Similarly, $V_{\overline{q}^1}(0, m(0)) = 0$ if a^* is not optimal at $p = 0$, while $V_{\overline{q}^1}(0, m(0) = v(a^*, 0)$ if a^* is optimal at $p = 0$. Also, $V_{\overline{q}^1}(\underline{q}^1, m(\underline{q}^1)) = (1-\delta)v(a^*, \underline{q}^1)$ if $\underline{q}^1 > 0$; while $V_{\overline{q}^1}(0, m(0)) = v(a^*, 0)$ if $\underline{q}^1 = 0$, since a^* is then optimal at $p = 0$.

With the function $V_{\overline{q}^1}$ defined at these three points, it is then defined at all points (p, w) in $\mathcal{W}^1_{\overline{q}^1} \cup \mathcal{W}^4_{\overline{q}^1}$. In particular, it is easy to show that

$$V_{\overline{q}^1}(\underline{q}^1, w) = \frac{M(\underline{q}^1) - w}{M(\underline{q}^1) - m(\underline{q}^1)}(1-\delta)v(a^*, \underline{q}^1) = \frac{M(\underline{q}^1) - w}{M(\underline{q}^1) - u(a^*, \underline{q}^1)} v(a^*, \underline{q}^1),$$

for all $w \in [m(\underline{q}^1), M(\underline{q}^1)]$.

At all points $(p, w) \in \mathcal{W}^3_{\overline{q}^1}$,

$$V_{\overline{q}^1}(p, w) = \frac{1-p}{1-\overline{q}^1} V_{\overline{q}^1}(\overline{q}^1, m(\overline{q}^1)).$$

Therefore, $V_{\overline{q}^1}$ is well-defined at all $(p, w) \in \mathcal{W} \setminus \mathcal{W}^2_{\overline{q}^1}$.

At all points $(p, w) \in \mathcal{W}^2_{\overline{q}^1}$, $V_{\overline{q}^1}(p, w)$ is defined via the recursive equation:

$$\begin{aligned} V_{\overline{q}^1}(p, w) &= \lambda(p, w)[(1-\delta)v(a^*, \varphi(p, w)) + \delta V_{\overline{q}^1}(\varphi(p, w), \mathbf{w}(\varphi(p, w))] \\ &= \lambda(p, w) V_{\overline{q}^1}(\varphi(p, w), m(\varphi(p, w))). \end{aligned}$$

Since $V_{\overline{q}^1}(p, w) = \lambda(p, w) V_{\overline{q}^1}(\varphi(p, w), m(\varphi(p, w))$, the value function is well-defined at all (p, w) if it is well-defined at all $(p, m(p))$, which we now prove.

By construction of the sets Q^k, observe that if $p \in Q^k \setminus Q^{k+1}$, then $\mathbf{w}(p) \in (U^k(p), U^{k+1}(p)]$ and, therefore, $\varphi(p, \mathbf{w}(p)) \in [\underline{q}^{k-1}, \underline{q}^k) \subset Q^{k-1} \setminus Q^k$. Moreover, $\varphi(\overline{q}^k, \mathbf{w}(\overline{q}^k)) = \underline{q}^k$. We now use these observations to complete the derivation of $V_{\overline{q}^1}$.

For all $p \in Q^1 \setminus Q^2$, we have that $\mathbf{w}(p) \in Q^0 \setminus Q^1$, so that $(p, \mathbf{w}(p)) \in \mathcal{W}^4_{\overline{q}^1}$. Since

$$V_{\overline{q}^1}(p, m(p)) = (1-\delta)v(a^*, p) + \delta V_{\overline{q}^1}(p, \mathbf{w}(p)),$$

$V_{\overline{q}^1}(p, m(p))$ is well-defined for all $p \in Q^1 \setminus Q^2$. By induction, assume that it is well-defined for all $p \in \bigcup_{\ell < k} Q^\ell \setminus Q^{\ell+1}$. We argue that it is well-defined for all $p \in Q^k \setminus Q^{k+1}$. Fix any $p \in Q^k \setminus Q^{k+1}$. From our initial observation, $\varphi(p, \mathbf{w}(p)) \in [\underline{q}^{k-1}, \underline{q}^k)$ and, therefore, $V_{\overline{q}^1}(p, m(p))$ is well-defined since

$$\begin{aligned} V_{\overline{q}^1}(p, m(p)) &= (1-\delta)v(a^*, p) + \delta V_{\overline{q}^1}(p, \mathbf{w}(p)) \\ &= (1-\delta)v(a^*, p) + \lambda(p, \mathbf{w}(p)) \underbrace{V_{\overline{q}^1}(\varphi(p, \mathbf{w}(p)), m(\varphi(p, \mathbf{w}(p))))}_{\text{defined by the induction step}}. \end{aligned}$$

Therefore, $V_{\overline{q}^1}(p, m(p))$ is well-defined for all $p \in \bigcup_\ell Q^\ell \setminus Q^{\ell+1} = Q^1 \setminus Q^\infty$. It remains to argue that it is well-defined for all $p \in Q^\infty$.

From the definition of Q^∞, we have that $\mathbf{w}(p) \leq \frac{1-p}{1-\underline{q}^\infty} m(\underline{q}^\infty) + \frac{p-\underline{q}^\infty}{1-\underline{q}^\infty} m(1)$ and, therefore, $\varphi(p, \mathbf{w}(p)) \in Q^\infty$. In other words, if $p \in Q^\infty$, then $\varphi(p, \mathbf{w}(p)) \in Q^\infty$, so that the restriction of $V_{\overline{q}^1}(\cdot, m(\cdot))$ to Q^∞ is entirely defined by its value on Q^∞ via the contraction:

$$V_{\overline{q}^1}(p, m(p)) = (1-\delta)v(a^*, p) + \delta\lambda(p, \mathbf{w}(p)) V_{\overline{q}^1}(\varphi(p, \mathbf{w}(p)), m(\varphi(p, \mathbf{w}(p)))).$$

The unique solution to this fixed point problem is given by:

$$V_{\overline{q}^1}(p, m(p)) = v(a^*, p) - \frac{m(p) - u(a^*, p)}{m(1) - u(a^*, 1)} v(a^*, 1),$$

for all $p \in Q^\infty$. To see this, with a slight abuse of notation, write (λ, φ) for $(\lambda(p, w), \varphi(p, \mathbf{w}(p)))$, and note that:

$$\begin{aligned} &(1-\delta)v(a^*, p) + \delta\lambda \left[v(a^*, \varphi) - \frac{m(\varphi) - u(a^*, \varphi)}{m(1) - u(a^*, 1)} v(a^*, 1) \right] \\ &= (1-\delta)v(a^*, p) + \delta\left[v(a^*, p) - (1-\lambda)v(a^*, 1)\right] \\ &\quad - \frac{m(p) - (1-\lambda)m(1) - u(a^*, p)(1-\delta)}{m(1) - u(a^*, 1)} v(a^*, 1) + \delta \frac{u(a^*, p) - (1-\lambda)u(a^*, 1)}{m(1) - u(a^*, 1)} v(a^*, 1) \\ &= v(a^*, p) - \frac{m(p) - u(a^*, p)}{m(1) - u(a^*, 1)} v(a^*, 1), \end{aligned}$$

where we use the identities $\lambda\varphi + (1-\lambda)1 = p$, $\lambda m(\varphi) + (1-\lambda)m(1) = \mathbf{w}(p)$, and $\delta\mathbf{w}(p) = m(p) - (1-\delta)u(a^*, p)$.

This completes the characterization of $V_{\overline{q}^1}$. Note that $V_{\overline{q}^1}$ and, therefore, all value functions V_q, are continuous functions.

A.4.3. *Value functions: another representation.* We now present another construction of V_q. For any $q \in [\underline{q}^1, \overline{q}^1]$, define the function $\overline{m}_q : [0,1] \to \mathbb{R}$ as

$$\begin{cases} \left(1 - \frac{p}{\underline{q}^1}\right) m(0) + \frac{p}{\underline{q}^1} m(\underline{q}^1) & \text{if } p \in [0, \underline{q}^1], \\ m(p) & \text{if } p \in (\underline{q}^1, q], \\ \frac{1-p}{1-q} m(q) + \frac{p-q}{1-q} m(1) & \text{if } p \in (q, 1]. \end{cases}$$

Note that $\overline{m}_q$ is convex, $\overline{m}_q(p) \geq m(p)$ for all $p \in [0,1]$, $\overline{m}_q(0) = m(0)$ and $\overline{m}_q(1) = m(1)$. For a graphical illustration, see Figure 11.

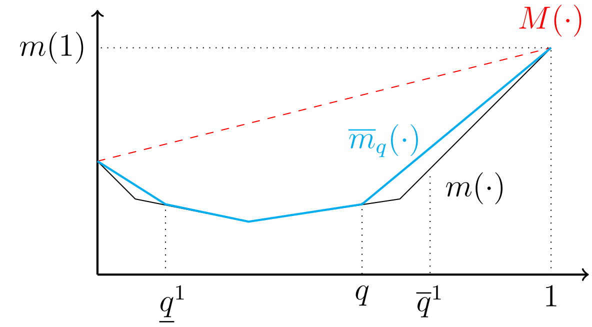

FIGURE 11. The function $\overline{m}_q$

It is straightforward to check that we have the following formula:

$$V_q(p, w) = \overline{\lambda}(p, w) V_q(\overline{\varphi}(p, w), \overline{m}_q(\overline{\varphi}(p, w)), \tag{11}$$

where the functions $\overline{\lambda}$ and $\overline{\varphi}$ are defined as in the main text, but with $\overline{m}_q$ instead of m. See Equation (3).

A.5. **Theorem 2.** To prove Theorem 2, we prove the following proposition and invoke Theorem 1.

Proposition 3. *Let V_{q^*} be the value function induced by the policy τ^*, with*

$$q^* = \sup\left\{p \in Q^1 : V_{\overline{q}^1}(p, m(p)) \geq V_{\overline{q}^1}(p, w) \text{ for all } w\right\}.$$

Then, V_{q^*} *is concave in* (p, w)*, decreasing in* w*, and satisfies:*

$$V_{q^*}(p, m(p)) \geq (1-\delta)v(a^*, p) + \delta V_{q^*}(p^*, \mathbf{w}(p)),$$

for all $p \in Q^1$.

We start with two preliminary observations.

OBSERVATION A. For all $q \in [\underline{q}^1, \overline{q}^1]$, we have the following identity:

$$V_q(p, w) = \frac{1-p}{1-p'} V_q\left(p', \frac{1-p'}{1-p} w + \frac{p'-p}{1-p} \overline{m}_q(1)\right).$$

The proof is as follows. Let $w' = \frac{1-p'}{1-p} w + \frac{p'-p}{1-p} \overline{m}_q(1)$.

Assume that $w' > \overline{m}_q(p')$. Since

$$\overline{\lambda}(p', w') \begin{pmatrix} \overline{\varphi}(p', w') \\ \overline{m}_q(\overline{\varphi}(p', w')) \end{pmatrix} + \left(1 - \overline{\lambda}(p', w')\right) \begin{pmatrix} 1 \\ \overline{m}_q(1) \end{pmatrix} = \begin{pmatrix} p' \\ w' \end{pmatrix},$$

we have

$$\frac{1-p}{1-p'} \overline{\lambda}(p', w') \begin{pmatrix} \overline{\varphi}(p', w') \\ \overline{m}_q(\overline{\varphi}(p', w')) \end{pmatrix} + \left(1 - \frac{1-p}{1-p'} \overline{\lambda}(p', w')\right) \begin{pmatrix} 1 \\ \overline{m}_q(1) \end{pmatrix} = \begin{pmatrix} p \\ w \end{pmatrix}.$$

Therefore, $\overline{\lambda}(p, w) = \frac{1-p}{1-p'} \overline{\lambda}(p', w')$ and $\overline{\varphi}(p', w') = \overline{\varphi}(p, w)$ since the solution $(\overline{\lambda}(p', w'), \overline{\varphi}(p', w'))$ is unique when $w' > m_q(p')$. The statement then follows from Equation (11).

Assume that $w' = \overline{m}_q(p')$. From the convexity of $\overline{m}_q$, this requires that $w = \overline{m}_q(p)$, so that $\overline{m}_q(p') = \frac{1-p'}{1-p} \overline{m}_q(p) + \frac{p'-p}{1-p} \overline{m}_q(1)$. The result follows from continuity as:

$$\begin{aligned}
V_q(p, \overline{m}_q(p)) &= \lim_{w \to \overline{m}_q(p)} V_q(p, w), \\
&= \lim_{w \to \overline{m}_q(p)} \frac{1-p}{1-p'} V_q\left(p', \frac{1-p'}{1-p} w + \frac{p'-p}{1-p} \overline{m}_q(1)\right), \\
&= \frac{1-p}{1-p'} V_q\left(p', \frac{1-p'}{1-p} \overline{m}_q(p) + \frac{p'-p}{1-p} \overline{m}_q(1)\right), \\
&= \frac{1-p}{1-p'} V_q\left(p', \overline{m}_q(p')\right).
\end{aligned}$$

Note that this implies that

$$V_q(p, \mathbf{w}(p) + c) = \overline{\lambda}(p, \mathbf{w}(p)) V_q \left(\overline{\varphi}(p, \mathbf{w}(p)), \overline{m}_q(\overline{\varphi}(p, \mathbf{w}(p))) + \frac{c}{\overline{\lambda}(p, \mathbf{w}(p))} \right),$$

where c is a positive constant.

OBSERVATION B. The value function $V_{\overline{q}^1}(p, \cdot) : [\overline{m}_{\overline{q}^1}(p), M(p)] \to \mathbb{R}$ is concave in w, for each p. See Lemma 6 in section A.6.

A.5.1. *Proposition 3(a).* We prove that V_{q^*} is decreasing in w. To start with, fix $p \in [0,1]$ and $(w, w') \in [\overline{m}_{q^*}(p), M(p)] \times [\overline{m}_{q^*}(p), M(p)]$, with $w' > w$.

First, assume that $p \leq q^*$. If $w = \overline{m}_{q^*}(p)$, then $V_{q^*}(p, w') \leq V_{q^*}(p, w)$ by construction of q^*. If $w > \overline{m}_{q^*}(p)$, we have that

$$\begin{aligned} \frac{V_{q^*}(p, w') - V_{q^*}(p, w)}{w' - w} &= \frac{V_{\overline{q}^1}(p, w') - V_{\overline{q}^1}(p, w)}{w' - w} \\ &\leq \frac{V_{\overline{q}^1}(p, w) - V_{\overline{q}^1}(p, \overline{m}_{q^*}(p))}{w - \overline{m}_{q^*}(p)} \\ &= \frac{V_{q^*}(p, w) - V_{q^*}(p, \overline{m}_{q^*}(p))}{w - \overline{m}_{q^*}(p)} \leq 0, \end{aligned}$$

where the inequality follows from the concavity of $V_{\overline{q}^1}$ with respect to w, for all $w \geq m_{\underline{q}^1}(p)$. (Recall that $m_{q^*}(p) = \overline{m}_{\overline{q}^1}(p)$ for all $p \leq q^*$.)

Second, assume that $p > q^*$. We show in detail how to make use of Observation A to deduce the result. We repeatedly use similar computations later on. We have

$$\begin{aligned} V_{q^*}(p, w') &= \overline{\lambda}(p, w') V_{q^*}(\overline{\varphi}(p, w'), \overline{m}_{q^*}(\overline{\varphi}(p, w'))) \\ &= \overline{\lambda}(p, w') \frac{1 - \overline{\varphi}(p, w')}{1 - \overline{\varphi}(p, w)} V_{q^*} \left(\overline{\varphi}(p, w), \frac{1 - \overline{\varphi}(p, w)}{1 - \overline{\varphi}(p, w')} \overline{m}_{q^*}(\overline{\varphi}(p, w')) + \left(1 - \frac{1 - \overline{\varphi}(p, w)}{1 - \overline{\varphi}(p, w')} \right) \overline{m}_{q^*}(1) \right) \\ &= \overline{\lambda}(p, w) V_{q^*} \left(\overline{\varphi}(p, w), \frac{\overline{\lambda}(p, w')}{\overline{\lambda}(p, w)} \overline{m}_{q^*}(\overline{\varphi}(p, w')) + \left(1 - \frac{\overline{\lambda}(p, w')}{\overline{\lambda}(p, w)} \right) \overline{m}_{q^*}(1) \right) \\ &= \overline{\lambda}(p, w) V_{q^*} \left(\overline{\varphi}(p, w), \overline{m}_{q^*}(\overline{\varphi}(p, w)) + \frac{w' - w}{\overline{\lambda}(p, w)} \right), \end{aligned}$$

where the first line follows from the construction of V_{q^*}, the second line from Observation A, the third line from the definition of the functions $\overline{\lambda}$ and $\overline{\varphi}$ and the last line from the following computations:

$$\begin{aligned}
\frac{\overline{\lambda}(p,w')}{\overline{\lambda}(p,w)}\overline{m}_{q^*}(\overline{\varphi}(p,w')) + \left(1-\frac{\overline{\lambda}(p,w')}{\overline{\lambda}(p,w)}\right)\overline{m}_{q^*}(1) &= \frac{1}{\overline{\lambda}(p,w)}w' + \left(1-\frac{1}{\overline{\lambda}(p,w)}\right)\overline{m}_{q^*}(1) \\
&= \frac{1}{\overline{\lambda}(p,w)}w' + \left(1-\frac{1}{\overline{\lambda}(p,w)}\right)\left[\frac{w-\overline{\lambda}(p,w)\overline{m}_{q^*}(\overline{\varphi}(p,w))}{1-\overline{\lambda}(p,w)}\right] \\
&= \overline{m}_{q^*}(\overline{\varphi}(p,w)) + \frac{w'-w}{\overline{\lambda}(p,w)}.
\end{aligned}$$

Thus, we are able to express $V_{q^*}(p,w')$ as $\overline{\lambda}(p,w)V_{q^*}(\overline{\varphi}(p,w),\tilde{w})$, with $\tilde{w}$ the above expression. Moreover, $\overline{\varphi}(p,w) \leq q^*$ as $w \geq \overline{m}_{q^*}(p)$. We can use the (already established) concavity of V_{q^*} in w for each $p \leq q^*$ to deduce the desired result. More precisely, we have that:

$$\begin{aligned}
\frac{V_{q^*}(p,w') - V_{q^*}(p,w)}{w'-w} &= \frac{\overline{\lambda}(p,w)\left(V_{q^*}\left(\overline{\varphi}(p,w),\overline{m}_{q^*}(\overline{\varphi}(p,w)) + \frac{w'-w}{\overline{\lambda}(p,w)}\right) - V_{q^*}\left(\overline{\varphi}(p,w),\overline{m}_{q^*}(\overline{\varphi}(p,w))\right)\right)}{w'-w} \\
&\leq 0,
\end{aligned}$$

where the inequality follows from the concavity of V_{q^*} in w at all $p \leq q^*$.

Lastly, since $V_{q^*}(p,w) = V_{q^*}(p,\overline{m}_{q^*}(p))$ for all $w \in [m(p),\overline{m}_{q^*}(p)]$, the result immediately follows for all (w,w'), with $w \in [m(p),\overline{m}_{q^*}(p)]$.

A.5.2. *Proposition 3(b).* We prove the concavity of V_{q^*} with respect to both arguments (p,w).

Let $\overline{\mathcal{W}} = \{(p,w) : w \geq \overline{m}_{q^*}(p)\}$. Let $(p,w) \in \overline{\mathcal{W}}$, $(p',w') \in \overline{\mathcal{W}}$ and $\alpha \in [0,1]$. Write (p_α, w_α) for

$$\alpha\begin{pmatrix} p \\ w \end{pmatrix} + (1-\alpha)\begin{pmatrix} p' \\ w' \end{pmatrix}.$$

Without loss of generality, assume that $p \leq p'$. We have that:

$$\begin{aligned}
&\alpha V_{q^*}(p,w)+(1-\alpha)V_{q^*}(p',w')\\
&=\alpha\frac{1-p}{1-p'}V_{q^*}\Big(p',\underbrace{\frac{1-p'}{1-p}w+\frac{p'-p}{1-p}\overline{m}_{q^*}(1)}_{\geq \overline{m}_{q^*}(p')}\Big)+(1-\alpha)V_{q^*}(p',w')\\
&\leq\left(\alpha\frac{1-p}{1-p'}+(1-\alpha)\right)V_{q^*}\left(p',\frac{\alpha\frac{1-p}{1-p'}\left(\frac{1-p'}{1-p}w+\frac{p'-p}{1-p}\overline{m}_{q^*}(1)\right)+(1-\alpha)w'}{\alpha\frac{1-p}{1-p'}+(1-\alpha)}\right)\\
&=\frac{1-p_\alpha}{1-p'}V_{q^*}\left(p',\frac{1-p'}{1-p_\alpha}w_\alpha+\frac{p'-p_\alpha}{1-p_\alpha}\overline{m}_{q^*}(1)\right)\\
&=V_{q^*}(p_\alpha,w_\alpha),
\end{aligned}$$

where the inequality follows from the concavity of $V_{\overline{q}^1}$ with respect to w for each p and the property that $V_{q^*}(p,w)=V_{\overline{q}^1}(p,w)$ for all (p,w) such that $w\geq \overline{m}_{q^*}(p)$. Notice that we use twice Observation A.

Finally, for all $(p,w)\in\mathcal{W}$, for all $(p',w')\in\mathcal{W}$ and for all α, we have that:

$$\begin{aligned}
\alpha V_{q^*}(p,w)+(1-\alpha)V_{q^*}(p',w') &= \alpha V_{q^*}(p,\max(w,\overline{m}_{q^*}(p)))+(1-\alpha)V_{q^*}(p',\max(w',\overline{m}_{q^*}(p')))\\
&\leq V_{q^*}(p_\alpha,\alpha\max(w,\overline{m}_{q^*}(p))+(1-\alpha)\max(w,\overline{m}_{q^*}(p')))\\
&\leq V_{q^*}(p_\alpha,w_\alpha),
\end{aligned}$$

since $\alpha\max(w,\overline{m}_{q^*}(p))+(1-\alpha)\max(w,\overline{m}_{q^*}(p'))\geq w_\alpha$ and the fact that V_{q^*} is decreasing in w for all p. This completes the proof of concavity.

A.5.3. *Proposition 3 (c).* We prove that $V_{q^*}(p,m(p))\geq(1-\delta)v(a^*,p)+\delta V_{q^*}(p,\mathbf{w}(p))$ for all $p\in Q^1$.

The statement is true for all $p\leq q^*$ by definition since $V_{q^*}(p,w)=V_{\overline{q}^1}(p,w)$ for all w.

Assume that $p>q^*$. From Lemma 7, there exists $\overline{q}$ such that $\varphi(p,\mathbf{w}(p))\geq\varphi(p',\mathbf{w}(p'))$ for all $p'\geq p\geq\overline{q}$. Moreover, it follows from A.6.3 and A.6.4 that $V(p,m(p))\geq V(p,w)$ for all w, for all $p\leq\overline{q}$. Therefore, we must have that $q^*\geq\overline{q}$. It follows that $\varphi(p,\mathbf{w}(p))<\varphi(q^*,\mathbf{w}(q^*))\leq q^*$, hence $\mathbf{w}(p)\geq\overline{m}_{q^*}(p)$. We therefore have that $V_{q^*}(p,\mathbf{w}(p))=V_{\overline{q}^1}(p,\mathbf{w}(p))$.

Since $V_{\overline{q}^1}(p, m(p)) = (1-\delta)v(a^*, p) + \delta V_{\overline{q}^1}(p, \mathbf{w}(p))$ for all $p \in Q^1$ and $V_{q^*}(p, m(p)) = V_{q^*}(p, \overline{m}_{q^*}(p)) = V_{\overline{q}^1}(p, \overline{m}_{q^*}(p))$, it is enough to prove that $V_{\overline{q}^1}(p, \overline{m}_{q^*}(p)) \geq V_{\overline{q}^1}(p, m(p))$. Clearly, there is nothing prove if $\overline{m}_{q^*}(p) = m(p)$ for all $p \in Q^1$, i.e., if $q^* = \overline{q}^1$ (remember that $\overline{m}_{\overline{q}^1}(p) = m(p)$ for all $p \in Q^1$).

So, assume that $\overline{m}_{q^*}(p) > m(p)$ for some $p \in (q^*, \overline{q}^1)$, hence $\overline{m}_{q^*}(p) > m(p)$ for all $p \in (q^*, \overline{q}^1)$. We now argue that if $V_{\overline{q}^1}(p, w) > V_{\overline{q}^1}(p, m(p))$ for some $w \geq \overline{m}_{q^*}(p)$, then

$$V_{\overline{q}^1}(p', m(p')) < \frac{1-p'}{1-p} V_{\overline{q}^1}(p, w),$$

for all $p' > p$. To see this, observe that $w > m(p)$ and, accordingly,

$$\frac{1-p'}{1-p} w + \frac{p'-p}{1-p} m(1) - m(p') > 0,$$

since m is convex. Hence,

$$\begin{aligned} 0 &< \frac{V_{\overline{q}^1}(p, w) - V_{\overline{q}^1}(p, m(p))}{w - m(p)} \\ &= \frac{\frac{1-p}{1-p'}\left[V_{\overline{q}^1}\left(p', \frac{1-p'}{1-p} w + \frac{p'-p}{1-p} m(1)\right) - V_{\overline{q}^1}\left(p', \frac{1-p'}{1-p} m(p) + \frac{p'-p}{1-p} m(1)\right)\right]}{w - m(p)} \\ &\leq \frac{V_{\overline{q}^1}\left(p', \frac{1-p'}{1-p} w + \frac{p'-p}{1-p} m(1)\right) - V_{\overline{q}^1}\left(p', m(p')\right)}{\frac{1-p'}{1-p} w + \frac{p'-p}{1-p} m(1) - m(p')}, \end{aligned}$$

where the equality follows Observation A and the inequality from the concavity of $V_{\overline{q}^1}$ in w for each p. Since

$$V_{\overline{q}^1}(p, w) = \frac{1-p}{1-p'} V_{\overline{q}^1}\left(p', \frac{1-p'}{1-p} w + \frac{p'-p}{1-p} m(1)\right),$$

we have the desired result.

Finally, from the definition of q^*, for all $n > 0$, there exist $p_n \in (q^*, \min(q^* + \frac{1}{n}, \overline{q}^1)]$ and $w_n \geq m(p_n)$ such that $V_{\overline{q}^1}(p_n, m(p_n)) < V_{\overline{q}^1}(p_n, w_n)$. From the concavity of $V_{\overline{q}^1}$ in w for all p, $V_{\overline{q}^1}(p_n, m(p_n)) < V_{\overline{q}^1}(p_n, \overline{m}_{q^*}(p_n))$ for all n.

From the above argument, for all p, for all n sufficiently large, i.e., such that $p_n < p$, we have that

$$V_{\overline{q}^1}(p, m(p)) < \frac{1-p}{1-p_n} V_{\overline{q}^1}(p_n, \overline{m}_{q^*}(p_n)).$$

Taking the limit as $n \to \infty$, we obtain that

$$V_{\overline{q}^1}(p, m(p)) < \frac{1-p}{1-q^*} V_{\overline{q}^1}(q^*, \overline{m}_{q^*}(q^*)) = V_{\overline{q}^1}(p, \overline{m}_{q^*}(p)),$$

which completes the proof.

A.6. Concavity of $V_{\overline{q}^1}$ with respect to w for each p.

Lemma 6. *The value function $V_{\overline{q}^1}(p, \cdot) : [\overline{m}_{\overline{q}^1}(p), M(p)] \to \mathbb{R}$ is concave in w, for each p.*

This section proves that $V_{\overline{q}^1}$ is concave in w for each p. To do so, we prove that

$$\frac{V_{\overline{q}^1}(p, \overline{m}_{\overline{q}^1}(p) + \eta(\overline{m}_{\overline{q}^1}(1) - u(a^*, 1))) - V_{\overline{q}^1}(p, \overline{m}_{\overline{q}^1}(p))}{\eta} \geq \frac{V_{\overline{q}^1}(p, \overline{m}_{\overline{q}^1}(p) + \eta'(\overline{m}_{\overline{q}^1}(1) - u(a^*, 1))) - V_{\overline{q}^1}(p, \overline{m}_{\overline{q}^1}(p))}{\eta'},$$

for all (η, η') such that $\eta' \geq \eta$. (See the observations on concave functions.) We start with some preliminary results.

A.6.1. *Preliminary Results.* We study how the function $\varphi(p, \mathbf{w}(p))$ varies with p.

Lemma 7. *There exists a non-empty interval $[\underline{q}, \overline{q}]$ such that:*

(1) For any $p' < p \leq \underline{q}$ or $p' > p \geq \bar{q}$, $\varphi(p, \mathbf{w}(p)) \geq \varphi(p', \mathbf{w}(p'))$,
(2) The ratio $\frac{m(1) - m(\varphi(p, \mathbf{w}(p)))}{1 - \varphi(p, \mathbf{w}(p))}$ is constant for all $p \in [\underline{q}, \overline{q}]$.

Proof of Lemma 7. Observe that

$$\frac{m(1) - \mathbf{w}(p)}{1-p} = \frac{m(1) - m(\varphi(p, \mathbf{w}(p)))}{1 - \varphi(p, \mathbf{w}(p))}.$$

Therefore, the convexity of m implies that if $\frac{m(1) - \mathbf{w}(p)}{1-p} < \frac{m(1) - \mathbf{w}(p')}{1-p'}$, then $\varphi(p, \mathbf{w}(p)) < \varphi(p', \mathbf{w}(p'))$.

Consider the function $h : [0,1] \to \mathbb{R}$, defined by $h(p) = \frac{m(1)-\mathbf{w}(p)}{1-p}$. We argue that h is quasi-concave. For all (p, p') and $\alpha \in [0,1]$, we have that

$$\begin{aligned}\frac{m(1)-\mathbf{w}(\alpha p+(1-\alpha)p')}{\alpha(1-p)+(1-\alpha)(1-p')} &\geq \frac{\alpha(m(1)-\mathbf{w}(p))+(1-\alpha)(m(1)-\mathbf{w}(p'))}{\alpha(1-p)+(1-\alpha)(1-p')} \\ &= \frac{\alpha(1-p)}{\alpha(1-p)+(1-\alpha)(1-p')}\frac{m(1)-\mathbf{w}(p)}{1-p}+ \\ &\quad \frac{(1-\alpha)(1-p')}{\alpha(1-p)+(1-\alpha)(1-p')}\frac{m(1)-\mathbf{w}(p')}{1-p'} \\ &\geq \min\left(\frac{m(1)-\mathbf{w}(p)}{1-p}, \frac{m(1)-\mathbf{w}(p')}{1-p'}\right),\end{aligned}$$

where the first inequality follows form the convexity of w. (Note that the inequality is strict if $\mathbf{w}(\alpha p+(1-\alpha)p') < \alpha\mathbf{w}(p)+(1-\alpha)\mathbf{w}(p')$.)

It follows that if $h(p') \geq h(p)$, then it is also true for all $p'' \in (p, p')$. Since h is quasi-concave and continuous, the set of maxima is a non-empty convex set $[\underline{q}, \overline{q}]$, and the function is increasing for all $p < \underline{q}$ and decreasing for all $p > \overline{q}$. (Note that $m(1) - \mathbf{w}(1) = \frac{(1-\delta)(u(a^*,1)-m(1))}{\delta} < 0$, hence the function is equal to $-\infty$ at $p = 1$.) □

We can make few additional observations about the interval $[\underline{q}, \overline{q}]$. Let $k^* := \sup\{k : Q^k \neq \emptyset\}$. Since $\varphi(\overline{q}^k, \mathbf{w}(\overline{q}^k)) = \underline{q}^k$, the function h is decreasing for all $p \geq \overline{q}^{k^*}$. Similarly, since $\varphi(\underline{q}^k, \mathbf{w}(\underline{q}^k)) = \underline{q}^{k-1}$, the function h is increasing for all $p \leq \underline{q}^{k^*}$. Therefore, $[\underline{q}, \overline{q}] \subset Q^{k^*}$.

If $P \neq \emptyset$, so that $k^* = \infty$, then for all $p \in P$, the function h is increasing by convexity of m since $\mathbf{w}(p) = m(p)$. (This is clearly true since $\varphi(p, m(p)) = p$ in that region.) Therefore, $\overline{p} \leq \underline{q}$ if $P \neq \emptyset$.

Finally, let $\tilde{p} := \inf\{p : m(p) = u(a^1, p)\}$. By construction, m is linear from $\tilde{p}$ to 1, i.e., $[\tilde{p}, 1]$ is the utmost right linear piece of m. We have that $\overline{q} < \tilde{p}$. To see this, observe that for all $p \geq \tilde{p}$,

$$\frac{m(1)-\mathbf{w}(p)}{1-p} = \frac{(1-\delta)\overbrace{(u(a^*,1)-u(a^1,1))}^{<0}}{1-p} + \frac{(u(a^1,0)-u(a^1,1))-(1-\delta)(u(a^*,0)-u(a^*,1))}{\delta},$$

hence it is decreasing in p. (If there are multiple optimal actions at $p = 1$, the argument applies to all of them and, therefore, to the one that induces the smallest $\tilde{p}$.)

The second preliminary result is technical. For any $p \in (0,1)$ and any $\eta \in \left[0, \frac{M(p)-\overline{m}_{\overline{q}^1}(p)}{\overline{m}_{\overline{q}^1}(1)-u(a^*,1)}\right]$, define $w(p;\eta)$ as

$$\overline{m}_{\overline{q}^1}(p) + \eta\left[\overline{m}_{\overline{q}^1}(1) - u(a^*,1)\right],$$

and write $(\lambda_\eta, \varphi_\eta)$ for $(\overline{\lambda}(p, w(p;\eta)), \overline{\varphi}(p, w(p;\eta)))$. To ease notation, we do not explicitly write the dependence of $(\lambda_\eta, \varphi_\eta)$ on p. We have the following:

Lemma 8. *φ_η, λ_η, and $\frac{1-\lambda_\eta}{\eta}$ are all decreasing in η.*

The proof follows directly from the definition of $(\lambda_\eta, \varphi_\eta)$ and is omitted.

Finally, we conclude with the following implication of Observation A, which wel use throughout. For all (p, w, w') with $w \leq w'$, we have that:

$$V_{\overline{q}^1}(p,w) - V_{\overline{q}^1}(p,w') = \overline{\lambda}(p,w)\left[V_{\overline{q}^1}(\overline{\varphi}(p,w), \overline{m}_{\overline{q}^1}(p,w)) - V_{\overline{q}^1}\left(\overline{\varphi}(p,w), \overline{m}_{\overline{q}^1}(p,w) + \frac{w'-w}{\overline{\lambda}(p,w)}\right)\right].$$

A.6.2. *Proof of Lemma 6.* We now prove that the gradient $\mathcal{G}(p;\eta) := \frac{V_{\overline{q}^1}(p,\overline{m}_{\overline{q}^1}(p)) - V_{\overline{q}^1}(p,w(p;\eta))}{\eta}$ is increasing in $\eta \in \left[0, \frac{M(p)-\overline{m}_{\overline{q}^1}(p)}{\overline{m}_{\overline{q}^1}(1)-u(a^*,1)}\right]$, for all p. We prove it on three separate intervals $\mathcal{I}_1$, $\mathcal{I}_2$ and $\mathcal{I}_3$. If $P = \emptyset$, the three intervals are $[0, \underline{q}]$, $(\underline{q}, \overline{q}]$ and $(\overline{q}, 1]$, respectively. If $P \neq \emptyset$, the three intervals are $[0, \underline{p}]$, $(\underline{p}, \overline{q}^\infty]$ and $(\overline{q}^\infty, 1]$, respectively.

A.6.3. *For all $p \in \mathcal{I}_1$, $\mathcal{G}(p;\eta)$ is increasing in η.* We limit attention to the case $P \neq \emptyset$. (The case $P = \emptyset$ is identical.) The proof is by induction. First, consider the interval $[0, \underline{q}^1]$. Remember that at $\underline{q}^1$, we have a closed-form solution for $V_{\overline{q}^1}(\underline{q}^1, w)$ for all w given by

$$V_{\overline{q}^1}(\underline{q}^1, w) = \frac{M(\underline{q}^1) - w}{M(\underline{q}^1) - u(a^*, \underline{q}^1)} v(a^*, \underline{q}^1).$$

Therefore,

$$\frac{V_{\overline{q}^1}(\underline{q}^1, \overline{m}_{\overline{q}^1}(\underline{q}^1)) - V_{\overline{q}^1}(\underline{q}^1, w(\underline{q}^1;\eta))}{\eta} = \frac{1}{\eta}\left[\frac{M(\underline{q}^1) - \overline{m}_{\overline{q}^1}(\underline{q}^1)}{M(\underline{q}^1) - u(a^*, \underline{q}^1)} v(a^*, \underline{q}^1) - \frac{M(\underline{q}^1) - w(\underline{q}^1;\eta)}{M(\underline{q}^1) - u(a^*, \underline{q}^1)} v(a^*, \underline{q}^1)\right]$$

$$= \frac{v(a^*, \underline{q}^1)}{M(\underline{q}^1) - u(a^*, \underline{q}^1)} \frac{[\overline{m}_{\overline{q}^1}(\underline{q}^1) + \eta(\overline{m}_{\overline{q}^1}(1) - u(a^*, 1))] - \overline{m}_{\overline{q}^1}(\underline{q}^1)}{\eta}$$

$$= \frac{\underline{q}^1 v(a^*, 1) + (1 - \underline{q}^1) v(a^*, 0)}{\underline{q}^1[\overline{m}_{\overline{q}^1}(1) - u(a^*, 1)] + (1 - \underline{q}^1)[\overline{m}_{\overline{q}^1}(0) - u(a^*, 0)]} \frac{w(\underline{q}^1;\eta) - \overline{m}_{\overline{q}^1}(\underline{q}^1)}{\eta}$$

$$= v(a^*, 1) \underbrace{\frac{\underline{q}^1 + (1 - \underline{q}^1)\frac{v(a^*,0)}{v(a^*,1)}}{\underline{q}^1 + (1 - \underline{q}^1)\frac{\overline{m}_{\overline{q}^1}(0) - u(a^*,0)}{\overline{m}_{\overline{q}^1}(1) - u(a^*,1)}}}_{\geq 1 \text{ since } \frac{v(a^*,0)}{v(a^*,1)} \geq \frac{\overline{m}_{\overline{q}^1}(0) - u(a^*,0)}{\overline{m}_{\overline{q}^1}(1) - u(a^*,1)}} \geq v(a^*, 1).$$

We now consider any $p \in [0, \underline{q}^1)$. From Observation A, we have that:

$$\begin{cases} V_{\overline{q}^1}(p, \overline{m}_{\overline{q}^1}(p)) = \frac{1-p}{1-\underline{q}^1} V_{\overline{q}^1}\left(\underline{q}^1, \frac{1-\underline{q}^1}{1-p}\overline{m}_{\overline{q}^1}(p) + \left(1 - \frac{1-\underline{q}^1}{1-p}\right)\overline{m}_{\overline{q}^1}(1)\right) \\ V_{\overline{q}^1}(p, w(p;\eta)) = \frac{1-p}{1-\underline{q}^1} V_{\overline{q}^1}\left(\underline{q}^1, \frac{1-\underline{q}^1}{1-p}\overline{m}_{\overline{q}^1}(p) + \left(1 - \frac{1-\underline{q}^1}{1-p}\right)\overline{m}_{\overline{q}^1}(1) + \frac{1-\underline{q}^1}{1-p}\eta\left[\overline{m}_{\overline{q}^1}(1) - u(a^*, 1)\right]\right) \end{cases}$$

It follows that

$$\frac{V_{\overline{q}^1}(p, \overline{m}_{\overline{q}^1}(p)) - V_{\overline{q}^1}(p, w(p;\eta))}{\eta}$$

$$= \frac{1-p}{1-\underline{q}^1} \frac{V_{\overline{q}^1}\left(\underline{q}^1, \frac{1-\underline{q}^1}{1-p}\overline{m}_{\overline{q}^1}(p) + \left(1 - \frac{1-\underline{q}^1}{1-p}\right)\overline{m}_{\overline{q}^1}(1)\right) - V_{\overline{q}^1}\left(\underline{q}^1, \frac{1-\underline{q}^1}{1-p}\overline{m}_{\overline{q}^1}(p) + \left(1 - \frac{1-\underline{q}^1}{1-p}\right)\overline{m}_{\overline{q}^1}(1) + \frac{1-\underline{q}^1}{1-p}\eta\left[\overline{m}_{\overline{q}^1}(1) - u(a^*, 1)\right]\right)}{\eta}$$

$$= \frac{1-p}{1-\underline{q}^1}\frac{1-\underline{q}^1}{1-p}\frac{\overline{m}_{\overline{q}^1}(1) - u(a^*, 1)}{M(\underline{q}^1) - u(a^*, \underline{q}^1)} v(a^*, \underline{q}^1) = \frac{1-p}{1-\underline{q}^1}\frac{1-\underline{q}^1}{1-p} v(a^*, 1) \frac{\underline{q}^1 + (1 - \underline{q}^1)\frac{v(a^*,0)}{v(a^*,1)}}{\underline{q}^1 + (1 - \underline{q}^1)\frac{\overline{m}_{\overline{q}^1}(0) - u(a^*,0)}{\overline{m}_{\overline{q}^1}(1) - u(a^*,1)}}$$

$$\geqslant \frac{1-p}{1-\underline{q}^1}\frac{1-\underline{q}^1}{1-p} v(a^*, 1) = v(a^*, 1).$$

Therefore, $\mathcal{G}(p;\eta) \geq v(a^*, 1)$ for all η, for all $p \in [0, \underline{q}^1]$. Moreover, the gradient $\mathcal{G}(p;\eta)$ is independent of η for all $p \in [0, \underline{q}^1]$, hence is (weakly) increasing.

By induction, assume that $\mathcal{G}(p;\eta) \geq v(a^*, 1)$ for all $p \in [0, \underline{q}^k]$ and is increasing in η, we want to prove that both properties also hold for all $p \in (\underline{q}^k, \underline{q}^{k+1}]$.

We rewrite $V_{\overline{q}^1}(p, w(p;\eta))$ as follows:

$$V_{\overline{q}^1}(p, w(p;\eta)) = \lambda_\eta V_{\overline{q}^1}(\varphi_\eta, \overline{m}_{\overline{q}^1}(\varphi_\eta)) = \lambda_\eta \left[(1-\delta)v(a^*, \varphi_\eta) + \delta V_{\overline{q}^1}(\varphi_\eta, \mathbf{w}(\varphi_\eta))\right]$$
$$= (1-\delta)\lambda_\eta v(a^*, \varphi_\eta) + \delta \lambda_\eta V_{\overline{q}^1}(\varphi_\eta, \mathbf{w}(\varphi_\eta))$$
$$= (1-\delta)\lambda_\eta v(a^*, \varphi_\eta) + \delta V_{\overline{q}^1}\left(p, \lambda_\eta \mathbf{w}(\varphi_\eta) + [1-\lambda_\eta]\overline{m}_{\overline{q}^1}(1)\right)$$
$$= (1-\delta)\lambda_\eta v(a^*, \varphi_\eta) + \delta V_{\overline{q}^1}\left(p, \mathbf{w}(p) + \frac{\eta - (1-\delta)(1-\lambda_\eta)}{\delta}[\overline{m}_{\overline{q}^1}(1) - u(a^*, 1)]\right).$$

The second to last equality follows from Observation A, while the last equality follows from:

$$\lambda_\eta \mathbf{w}(\varphi_\eta) + [1-\lambda_\eta]\overline{m}_{\overline{q}^1}(1) = \lambda_\eta \frac{-(1-\delta)u(a^*, \varphi_\eta) + \overline{m}_{\overline{q}^1}(\varphi_\eta)}{\delta} + [1-\lambda_\eta]\overline{m}_{\overline{q}^1}(1)$$
$$= \frac{-(1-\delta)}{\delta}\lambda_\eta u(a^*, \varphi_\eta) + \frac{1}{\delta}\lambda_\eta \overline{m}_{\overline{q}^1}(\varphi_\eta) + [1-\lambda_\eta]\overline{m}_{\overline{q}^1}(1)$$
$$= \frac{-(1-\delta)}{\delta}\left[u(a^*, p) - (1-\lambda_\eta)u(a^*, 1)\right] + \frac{1}{\delta}\left[w(p;\eta) - (1-\lambda_\eta)\overline{m}_{\overline{q}^1}(1)\right] + [1-\lambda_\eta]\overline{m}_{\overline{q}^1}(1)$$
$$= \frac{-(1-\delta)}{\delta}\left[u(a^*, p) - (1-\lambda_\eta)u(a^*, 1)\right] + \frac{1}{\delta}\left[\overline{m}_{\overline{q}^1}(p) + \eta(\overline{m}_{\overline{q}^1}(1) - u(a^*, 1)) - (1-\lambda_\eta)\overline{m}_{\overline{q}^1}(1)\right] + [1-\lambda_\eta]\overline{m}_{\overline{q}^1}$$
$$= \left[\frac{-(1-\delta)}{\delta}u(a^*, p) + \frac{1}{\delta}\overline{m}_{\overline{q}^1}(p)\right] + \frac{\eta - (1-\delta)(1-\lambda_\eta)}{\delta}[\overline{m}_{\overline{q}^1}(1) - u(a^*, 1)].$$

For future reference, recall that

$$\lambda_\eta \mathbf{w}(\varphi_\eta) + (1-\lambda_\eta)\overline{m}_{\overline{q}^1}(1) = \lambda_\eta \left[\overline{\lambda}(\varphi_\eta, \mathbf{w}(\varphi_\eta))\overline{m}_{\overline{q}^1}(\overline{\varphi}(\varphi_\eta, \mathbf{w}(\varphi_\eta))) + (1 - \overline{\lambda}(\varphi_\eta, \mathbf{w}(\varphi_\eta))\overline{m}_{\overline{q}^1}(1)\right]$$
$$+ (1-\lambda_\eta)\overline{m}_{\overline{q}^1}(1),$$

so that

$$\overline{\varphi}\left(p, \mathbf{w}(p) + \frac{\eta - (1-\delta)(1-\lambda_\eta)}{\delta}[\overline{m}_{\overline{q}^1}(1) - u(a^*, 1)]\right) = \overline{\varphi}(\varphi_\eta, \mathbf{w}(\varphi_\eta)),$$

and

$$\overline{\lambda}\left(p, \mathbf{w}(p) + \frac{\eta - (1-\delta)(1-\lambda_\eta)}{\delta}[\overline{m}_{\overline{q}^1}(1) - u(a^*, 1)]\right) = \lambda_\eta \overline{\lambda}(\varphi_\eta, \mathbf{w}(\varphi_\eta)).$$

Since φ_η is decreasing in η, we have that $\varphi_{\eta'} \leq \varphi_\eta$ when $\eta' > \eta$ and, therefore, we have that $\overline{\varphi}(\varphi_\eta, \mathbf{w}(\varphi_\eta)) \leq \overline{\varphi}(\varphi_{\eta'}, \mathbf{w}(\varphi_{\eta'}))$ since $\varphi_{\eta'} \leq \varphi_\eta \leq p \leq \underline{q}$. Similarly, since $\varphi_\eta < p \leq \underline{q}$, we have that $\overline{\varphi}(\varphi_\eta, \mathbf{w}(\varphi_\eta)) \leq \overline{\varphi}(p, \mathbf{w}(p))$ and, therefore, $\frac{\eta-(1-\delta)(1-\lambda_\eta)}{\delta} > 0$.

We now return to the computation of the gradient. We have:

$$= \frac{\Big[(1-\delta)v(a^*,p) + \delta V_{\overline{q}^1}(p,\mathbf{w}(p))\Big] - \Big[(1-\delta)\lambda_\eta v(a^*,\varphi_\eta) + \delta V_{\overline{q}^1}\Big(p,\mathbf{w}(p) + \frac{\eta-(1-\delta)(1-\lambda_\eta)}{\delta}[m(1)-u(a^*,1)]\Big)\Big]}{\eta}$$

$$= \frac{(1-\delta)}{\eta}\left[v(a^*,p) - \lambda_\eta v(a^*,\varphi_\eta)\right] + \frac{\delta}{\eta}\left[V_{\overline{q}^1}(p,\mathbf{w}(p)) - V_{\overline{q}^1}\left(p,\mathbf{w}(p) + \frac{\eta-(1-\delta)(1-\lambda_\eta)}{\delta}[m(1)-u(a^*,1)]\right)\right]$$

$$= \frac{(1-\delta)}{\eta}(1-\lambda_\eta)v(a^*,1) + \frac{\delta}{\eta}\left[V_{\overline{q}^1}(p,\mathbf{w}(p)) - V_{\overline{q}^1}\left(p,\mathbf{w}(p) + \frac{\eta-(1-\delta)(1-\lambda_\eta)}{\delta}[m(1)-u(a^*,1)]\right)\right]. \tag{12}$$

We further develop the above expression. To ease notation, we write $(\varphi(p),\lambda(p))$ for $(\overline{\varphi}(p,\mathbf{w}(p)),\overline{\lambda}(p,\mathbf{w}(p)))$. Note that $\varphi(p)\in(\underline{q}^{k-1},\underline{q}^{k}]$, since $p\in(\underline{q}^{k},\underline{q}^{k+1}]$. As $\frac{\eta-(1-\delta)(1-\lambda_\eta)}{\delta}>0$, we have that:

$$= \frac{(1-\delta)}{\eta}(1-\lambda_\eta)v(a^*,1) + \frac{\delta}{\eta}\left[V_{\overline{q}^1}(p,\mathbf{w}(p)) - V_{\overline{q}^1}\left(p,\mathbf{w}(p) + \frac{\eta-(1-\delta)(1-\lambda_\eta)}{\delta}[m(1)-u(a^*,1)]\right)\right]$$

$$= \frac{(1-\delta)}{\eta}(1-\lambda_\eta)v(a^*,1) + \frac{\delta}{\eta}\frac{\eta-(1-\delta)(1-\lambda_\eta)}{\delta}\frac{V_{\overline{q}^1}(p,\mathbf{w}(p)) - V_{\overline{q}^1}\left(p,\mathbf{w}(p)+\frac{\eta-(1-\delta)(1-\lambda_\eta)}{\delta}[m(1)-u(a^*,1)]\right)}{\frac{\eta-(1-\delta)(1-\lambda_\eta)}{\delta}}$$

$$= \frac{(1-\delta)}{\eta}(1-\lambda_\eta)v(a^*,1) + \left[1-\frac{(1-\delta)\,(1-\lambda_\eta)}{\eta}\right]\frac{\lambda(p)\left[V_{\overline{q}^1}(\varphi(p),\overline{m}_{\overline{q}^1}(\varphi(p))) - V_{\overline{q}^1}\left(\varphi(p),\overline{m}_{\overline{q}^1}(\varphi(p)) + \frac{\eta-(1-\delta)(1-\lambda_\eta)}{\delta\lambda(p)}[m(1)-u(a^*,1)]\right.\right.}{\frac{\eta-(1-\delta)(1-\lambda_\eta)}{\delta}}$$

$$= \frac{(1-\delta)}{\eta}(1-\lambda_\eta)v(a^*,1) + \left[1-\frac{(1-\delta)\,(1-\lambda_\eta)}{\eta}\right]\frac{V_{\overline{q}^1}(\varphi(p),\overline{m}_{\overline{q}^1}(\varphi(p))) - V_{\overline{q}^1}\left(\varphi(p),\overline{m}_{\overline{q}^1}(\varphi(p)) + \frac{\eta-(1-\delta)(1-\lambda_\eta)}{\delta\lambda(p)}[m(1)-u(a^*,1)]\right)}{\frac{\eta-(1-\delta)(1-\lambda_\eta)}{\delta\lambda(p)}}$$

$$\geqslant \frac{(1-\delta)}{\eta}(1-\lambda_\eta)v(a^*,1) + \left[1-\frac{(1-\delta)\,(1-\lambda_\eta)}{\eta}\right]v(a^*,1) = v(a^*,1),$$

where we use Observation A and the induction step.

We now show that the gradient is increasing in η. To start with, note that $\frac{\eta-(1-\delta)(1-\lambda_\eta)}{\delta}$ is increasing in η since $\frac{1-\lambda_\eta}{\eta}$ is decreasing in η (see Lemma 8). For any $\eta > \eta'$, we have the following

$$\frac{V_{\overline{q}^1}(p, \mathbf{w}(p)) - V_{\overline{q}^1}\left(p, \mathbf{w}(p) + \frac{\eta-(1-\delta)(1-\lambda_\eta)}{\delta}\left[\overline{m}_{\overline{q}^1}(1) - u(a^*,1)\right]\right)}{\frac{\eta-(1-\delta)(1-\lambda_\eta)}{\delta}}$$

$$= \frac{\lambda(p)V_{\overline{q}^1}(\varphi(p), \overline{m}_{\overline{q}^1}(\varphi(p))) - \lambda(p)V_{\overline{q}^1}\left(\varphi(p), \overline{m}_{\overline{q}^1}(\varphi(p)) + \frac{\eta-(1-\delta)(1-\lambda_\eta)}{\delta\lambda(p)}\left[\overline{m}_{\overline{q}^1}(1) - u(a^*,1)\right]\right)}{\frac{\eta-(1-\delta)(1-\lambda)}{\delta}}$$

$$= \frac{V_{\overline{q}^1}(\varphi(p), \overline{m}_{\overline{q}^1}(\varphi(p))) - V_{\overline{q}^1}\left(\varphi(p), \overline{m}_{\overline{q}^1}(\varphi(p)) + \frac{\eta-(1-\delta)(1-\lambda_\eta)}{\delta\lambda(p)}\left[\overline{m}_{\overline{q}^1}(1) - u(a^*,1)\right]\right)}{\frac{\eta-(1-\delta)(1-\lambda)}{\delta\lambda(p)}}$$

$$\geqslant \frac{V_{\overline{q}^1}(\varphi(p), \overline{m}_{\overline{q}^1}(\varphi(p))) - V_{\overline{q}^1}\left(\varphi(p), \overline{m}_{\overline{q}^1}(\varphi(p)) + \frac{\eta'-(1-\delta)(1-\lambda_{\eta'})}{\delta\lambda(p)}\left[\overline{m}_{\overline{q}^1}(1) - u(a^*,1)\right]\right)}{\frac{\eta'-(1-\delta)(1-\lambda_{\eta'})}{\delta\lambda(p)}}$$

$$= \frac{V_{\overline{q}^1}(p, \mathbf{w}(p)) - V_{\overline{q}^1}\left(p, \mathbf{w}(p) + \frac{\eta'-(1-\delta)(1-\lambda_{\eta'})}{\delta}\left[\overline{m}_{\overline{q}^1}(1) - u(a^*,1)\right]\right)}{\frac{\eta'-(1-\delta)(1-\lambda_{\eta'})}{\delta}},$$

where the inequality follows from the fact that $\varphi(p) \in (\underline{q}^{k-1}, \underline{q}^k]$ and, therefore, the gradient $\mathcal{G}(\varphi(p); \eta)$ being increasing in η by the induction hypothesis.

Finally, we have that

$$\frac{1}{\eta}\left[V_{\overline{q}^1}(p, \overline{m}_{\overline{q}^1}(p)) - V_{\overline{q}^1}(p, w(p;\eta))\right] =$$

$$\frac{(1-\delta)(1-\lambda_\eta)}{\eta}v(a^*,1) + \left[1 - \frac{(1-\delta)(1-\lambda_\eta)}{\eta}\right]\frac{V_{\overline{q}^1}(p, \mathbf{w}(p)) - V_{\overline{q}^1}\left(p, \mathbf{w}(p) + \frac{\eta-(1-\delta)(1-\lambda_\eta)}{\delta}\left[m(1) - u(a^*,1)\right]\right)}{\frac{\eta-(1-\delta)(1-\lambda_\eta)}{\delta}}$$

$$\geqslant \frac{(1-\delta)(1-\lambda_\eta)}{\eta}v(a^*,1) + \left[1 - \frac{(1-\delta)(1-\lambda_\eta)}{\eta}\right]\frac{V_{\overline{q}^1}(p, \mathbf{w}(p)) - V_{\overline{q}^1}\left(p, \mathbf{w}(p) + \frac{\eta'-(1-\delta)(1-\lambda_{\eta'})}{\delta}\left[\overline{m}_{\overline{q}^1}(1) - u(a^*,1)\right]\right)}{\frac{\eta'-(1-\delta)(1-\lambda_{\eta'})}{\delta}}$$

$$= \frac{(1-\delta)(1-\lambda_{\eta'})}{\eta'}v(a^*,1) + \left[1 - \frac{(1-\delta)(1-\lambda_{\eta'})}{\eta'}\right]\frac{V_{\overline{q}^1}(p, \mathbf{w}(p)) - V_{\overline{q}^1}\left(p, \mathbf{w}(p) + \frac{\eta'-(1-\delta)(1-\lambda_{\eta'})}{\delta}\left[\overline{m}_{\overline{q}^1}(1) - u(a^*,1)\right]\right)}{\frac{\eta'-(1-\delta)(1-\lambda_{\eta'})}{\delta}}$$

$$+ \left[\frac{(1-\delta)(1-\lambda_{\eta'})}{\eta'} - \frac{(1-\delta)(1-\lambda_\eta)}{\eta}\right]\left[\frac{V_{\overline{q}^1}(p, \mathbf{w}(p)) - V_{\overline{q}^1}\left(p, \mathbf{w}(p) + \frac{\eta'-(1-\delta)(1-\lambda_{\eta'})}{\delta}\left[\overline{m}_{\overline{q}^1}(1) - u(a^*,1)\right]\right)}{\frac{\eta'-(1-\delta)(1-\lambda_{\eta'})}{\delta}} - v(a^*,1)\right]$$

$$\geqslant \frac{1}{\eta'}\left[V_{\overline{q}^1}(p, \overline{m}_{\overline{q}^1}(p)) - V_{\overline{q}^1}(p, w(p;\eta'))\right]$$

$$+ \left[\frac{(1-\delta)(1-\lambda_{\eta'})}{\eta'} - \frac{(1-\delta)(1-\lambda_\eta)}{\eta}\right]$$

$$\times \left[\frac{V_{\overline{q}^1}(\varphi(p), \overline{m}_{\overline{q}^1}(\varphi(p))) - V_{\overline{q}^1}\left(\varphi(p), \overline{m}_{\overline{q}^1}(\varphi(p)) + \frac{\eta'-(1-\delta)(1-\lambda_{\eta'})}{\delta\lambda(p)}\left[\overline{m}_{\overline{q}^1}(1) - u(a^*,1)\right]\right)}{\frac{\eta'-(1-\delta)(1-\lambda_{\eta'})}{\delta\lambda(p)}} - v(a^*,1)\right]$$

$$\geqslant \frac{1}{\eta'}\left[V_{\overline{q}^1}(p, \overline{m}_{\overline{q}^1}(p)) - V_{\overline{q}^1}(p, w(p;\eta'))\right].$$

The last inequality follows from the fact that the gradient in the second bracket is weakly larger than $v(a^*,1)$ by the induction hypothesis and the fact that $\frac{1-\lambda_\eta}{\eta} < \frac{1-\lambda_{\eta'}}{\eta'}$ (Lemma 8).

Since $\lim_{k\to\infty} \underline{q}^k = \underline{p}$ when $P \neq \emptyset$, this completes the proof that the gradient is greater than $v(a^*,1)$ for all $p \in [0,\underline{p}]$.

A.6.4. *For all $p \in \mathcal{I}_2$, $\mathcal{G}(p;\eta)$ is increasing in η.* We first treat the case $P \neq \emptyset$. Recall that for all $p \in (\underline{p}, \overline{q}^\infty]$, we have an explicit definition of the value function $V_{\overline{q}^1}(p, \overline{m}_{\overline{q}^1}(p))$ as:

$$v(a^*,p) - \frac{\overline{m}_{\overline{q}^1}(p) - u(a^*,p)}{\overline{m}_{\overline{q}^1}(1) - u(a^*,1)} v(a^*,1).$$

Define $\bar{\eta}(p)$ as the solution to $\varphi_{\bar{\eta}(p)} = \overline{\varphi}(p, w(p;\bar{\eta}(p))) = \underline{p}$. Note that for any $p \in (\underline{p}, \overline{q}^\infty]$, for any $\eta \leq \bar{\eta}$, $\varphi_\eta \in [\underline{p}, \overline{q}^\infty]$. Therefore,

$$\begin{aligned} V_{\overline{q}^1}(p, w(p;\eta)) = \lambda_\eta V_{\overline{q}^1}(\varphi_\eta, \overline{m}_{\overline{q}^1}(\varphi_\eta)) &= \lambda_\eta \left[v(a^*,\varphi_\eta) - \frac{\overline{m}_{\overline{q}^1}(\varphi_\eta) - u(a^*,\varphi_\eta)}{\overline{m}_{\overline{q}^1}(1) - u(a^*,1)} v(a^*,1) \right] \\ &= v(a^*,p) - \frac{w(p;\eta) - u(a^*,p)}{\overline{m}_{\overline{q}^1}(1) - u(a^*,1)} v(a^*,1). \end{aligned}$$

It follows that the gradient is equal to $v(a^*,1)$ for all $p \in (\underline{p}, p^*]$, for all $\eta \leq \bar{\eta}$.

Consider now $\eta > \bar{\eta}$. We rewrite the gradient $\mathcal{G}(p;\eta)$ as follows:

$$\begin{aligned} & \frac{V_{\overline{q}^1}(p, \overline{m}_{\overline{q}^1}(p)) - V_{\overline{q}^1}(p, w(p;\eta))}{\eta} \\ &= \frac{V_{\overline{q}^1}(p, \overline{m}_{\overline{q}^1}(p)) - V_{\overline{q}^1}(p, w(p;\eta_1(p)))}{\eta} + \frac{V_{\overline{q}^1}(p, w(p;\eta_1(p))) - V_{\overline{q}^1}(p, w(p;\eta))}{\eta} \\ &= \frac{\eta_1(p)}{\eta} \frac{V_{\overline{q}^1}(p, \overline{m}_{\overline{q}^1}(p)) - V_{\overline{q}^1}(p, w(p,\eta_1(p)))}{\eta_1(p)} + \frac{\eta - \eta_1(p)}{\eta} \frac{V_{\overline{q}^1}(p, w(p;\eta_1(p))) - V_{\overline{q}^1}(p, w(p;\eta))}{\eta - \eta_1(p)} \\ &= \frac{\eta_1(p)}{\eta} v(a^*,1) + \frac{\eta - \eta_1(p)}{\eta} \frac{\frac{1-p}{1-\underline{p}} \left[V_{\overline{q}^1}(\underline{p}, \overline{m}_{\overline{q}^1}(\underline{p})) - V_{\overline{q}^1}\left(\underline{p}, w\left(\underline{p}; \frac{\eta-\eta_1(p)}{\frac{1-p}{1-\underline{p}}}\right)\right)\right]}{\eta - \eta_1(p)} \\ &= \frac{\eta_1(p)}{\eta} v(a^*,1) + \frac{\eta - \eta_1(p)}{\eta} \mathcal{G}\left(\underline{p}; \frac{\eta - \eta_1(p)}{\frac{1-p}{1-\underline{p}}}\right). \end{aligned}$$

Since we have already shown that $\mathcal{G}(\underline{p};\eta)$ is increasing in η and weakly larger than $v(a^*,1)$, we have that the gradient $\mathcal{G}(p;\eta)$ is also weakly increasing in η (and greater than $v(a^*,1)$).

We now treat the case $P = \emptyset$. Define $\bar{\eta}(p)$ as the solution to $\varphi_{\bar{\eta}(p)} = \overline{\varphi}(p, w(p; \bar{\eta}(p))) = \underline{q}$. Note that for any $p \in [\underline{q}, \overline{q}]$, for any $\eta \leq \bar{\eta}$, $\varphi_\eta \in [\underline{q}, \overline{q}]$. Therefore, for all $\eta \leq \bar{\eta}$, $\eta = (1-\delta)(1-\lambda_\eta)$ since the ratio $\frac{\overline{m}_{\overline{q}^1}(1) - \mathbf{w}(\varphi_\eta)}{1-\varphi_\eta}$ is constant in η and so is $\overline{\varphi}(\varphi_\eta, \mathbf{w}(\varphi_\eta))$. (Recall that we vary η at a fixed p.) It follows then from Equation (12) that

$$\mathcal{G}(p;\eta) = \frac{(1-\delta)}{\eta}(1-\lambda_\eta)v(a^*,1) + \frac{\delta}{\eta}\left[V_{\overline{q}^1}(p, \mathbf{w}(p)) - V_{\overline{q}^1}\left(p, \mathbf{w}(p) + \frac{\eta - (1-\delta)(1-\lambda_\eta)}{\delta}[m(1) - u(a^*,1)]\right)\right]$$

$$= \frac{(1-\delta)}{\eta}(1-\lambda_\eta)v(a^*,1) = v(a^*,1).$$

We have that the gradient $\mathcal{G}(p;\eta)$ is equal to $v(a^*,1)$ for all $p \in (\underline{q}, \overline{q}]$, for all $\eta \leq \bar{\eta}$. Finally, when $\eta > \bar{\eta}$, the same decomposition as in the case $P \neq \emptyset$ completes the proof.

A.6.5. *For all $p \in \mathcal{I}_3$, the gradient $\mathcal{G}(p;\eta)$ is increasing in η.*

We only treat the case $P \neq \emptyset$. (The case $P = \emptyset$ is treated analogously.) Define $\bar{\eta}(p)$ as the solution to $\varphi_{\bar{\eta}(p)} = \overline{\varphi}(p, w(p; \bar{\eta}(p))) = \overline{q}^\infty$. By construction, for all $p \in (\overline{q}^\infty, 1]$, for all $\eta \leq \bar{\eta}(p)$, we have that $\varphi_\eta \in (\overline{q}^\infty, 1]$. Therefore, $\varphi_\eta > \overline{q}$.

Choose $\bar{\eta}(p) \leq \eta' \leq \eta$. We have that $\varphi_{\eta'} \geq \varphi_\eta \geq \overline{q}$ since $\overline{q}^\infty \geq \overline{q}$ and, therefore,

$$\overline{\varphi}\left(p, \mathbf{w}(p) + \frac{\eta - (1-\delta)(1-\lambda_\eta)}{\delta}[\overline{m}_{\overline{q}^1}(1) - u(a^*,1)]\right) = \overline{\varphi}(\varphi_\eta, \mathbf{w}(\varphi_\eta)) \geq$$

$$\overline{\varphi}(\varphi_{\eta'}, \mathbf{w}(\varphi_{\eta'}) = \overline{\varphi}\left(p, \mathbf{w}(p) + \frac{\eta' - (1-\delta)(1-\lambda_{\eta'})}{\delta}[\overline{m}_{\overline{q}^1}(1) - u(a^*,1)]\right).$$

Also, since $\overline{q} \leq \varphi_\eta \leq p$, we have that $\overline{\varphi}(\varphi_\eta, \mathbf{w}(\varphi_\eta)) \geq \overline{\varphi}(p, \mathbf{w}(p))$ and, therefore, $\frac{\eta - (1-\delta)(1-\lambda_\eta)}{\delta} \leq 0$. The same applies to η'. Finally, as already shown,

$$\frac{\eta - (1-\delta)(1-\lambda_\eta)}{\delta} < \frac{\eta' - (1-\delta)(1-\lambda_{\eta'})}{\delta}.$$

To ease notation, define $(\tilde{\lambda}_\eta, \tilde{\varphi}_\eta)$ as follows:

$$\begin{cases} \tilde{\lambda}_\eta = \lambda\left(p, \mathbf{w}(p) - \dfrac{(1-\delta)(1-\lambda_\eta) - \eta}{\delta}[m(1) - u(a^*,1)]\right) \\ \tilde{\varphi}_\eta = \varphi\left(p, \mathbf{w}(p) - \dfrac{(1-\delta)(1-\lambda_\eta) - \eta}{\delta}[m(1) - u(a^*,1)]\right) \end{cases} \tag{13}$$

Notice that $\tilde{\varphi}_\eta = \overline{\varphi}(\varphi_\eta, \mathbf{w}(\varphi_\eta)) \in \mathcal{I}_1$ since $\varphi_\eta > \overline{q}^\infty$.

The rest of the proof is purely algebraic and mirrors the case $p \in \mathcal{I}_1$. First, we have the following:

$$\frac{V_{\overline{q}^1}(p, \mathbf{w}(p)) - V_{\overline{q}^1}\left(p, \mathbf{w}(p) - \frac{(1-\delta)(1-\lambda_\eta)-\eta}{\delta}\left[\overline{m}_{\overline{q}^1}(1) - u(a^*,1)\right]\right)}{\frac{(1-\delta)(1-\lambda_\eta)-\eta}{\delta}}$$
$$= \frac{\tilde{\lambda}_\eta V_{\overline{q}^1}\left(\tilde{\varphi}_\eta, \overline{m}_{\overline{q}^1}(\tilde{\varphi}_\eta) + \frac{(1-\delta)(1-\lambda_\eta)-\eta}{\delta\tilde{\lambda}_\eta}\left[\overline{m}_{\overline{q}^1}(1) - u(a^*,1)\right]\right) - \tilde{\lambda}_\eta V_{\overline{q}^1}\left(\tilde{\varphi}_\eta, \overline{m}_{\overline{q}^1}(\tilde{\varphi}_\eta)\right)}{\frac{(1-\delta)(1-\lambda_\eta)-\eta}{\delta}}$$
$$= \frac{V_{\overline{q}^1}\left(\tilde{\varphi}_\eta, w\left(\tilde{\varphi}_\eta; \frac{(1-\delta)(1-\lambda_\eta)-\eta}{\delta\tilde{\lambda}_\eta}\right)\right) - V_{\overline{q}^1}\left(\tilde{\varphi}_\eta, \overline{m}_{\overline{q}^1}(\tilde{\varphi}_\eta)\right)}{\frac{(1-\delta)(1-\lambda_\eta)-\eta}{\delta\tilde{\lambda}_\eta}},$$

where we again use Observation A. Similarly, we have:

$$\frac{V_{\overline{q}^1}(p, w(p)) - V_{\overline{q}^1}\left(p, w(p) - \frac{(1-\delta)(1-\lambda_{\eta'})-\eta'}{\delta}\left[\overline{m}_{\overline{q}^1}(1) - u(a^*,1)\right]\right)}{\frac{(1-\delta)(1-\lambda_{\eta'})-\eta'}{\delta}}$$
$$= \frac{\tilde{\lambda}_\eta V_{\overline{q}^1}\left(\tilde{\varphi}_\eta, w\left(\tilde{\varphi}_\eta; \frac{(1-\delta)(1-\lambda_\eta)-\eta}{\delta\tilde{\lambda}_\eta}\right)\right) - \tilde{\lambda}_\eta V_{\overline{q}^1}\left(\tilde{\varphi}_\eta, w\left(\tilde{\varphi}_\eta; \frac{(1-\delta)(1-\lambda_\eta)-\eta}{\delta\tilde{\lambda}_\eta} - \frac{(1-\delta)(1-\lambda_{\eta'})-\eta'}{\delta\tilde{\lambda}_\eta}\right)\right)}{\frac{(1-\delta)(1-\lambda_{\eta'})-\eta'}{\delta}}$$
$$= \frac{V_{\overline{q}^1}\left(\tilde{\varphi}_\eta, w\left(\tilde{\varphi}_\eta; \frac{(1-\delta)(1-\lambda_\eta)-\eta}{\delta\tilde{\lambda}_\eta}\right)\right) - V_{\overline{q}^1}\left(\tilde{\varphi}_\eta, w\left(\tilde{\varphi}_\eta; \frac{(1-\delta)(1-\lambda_\eta)-\eta}{\delta\tilde{\lambda}_\eta} - \frac{(1-\delta)(1-\lambda_{\eta'})-\eta'}{\delta\tilde{\lambda}_\eta}\right)\right)}{\frac{(1-\delta)(1-\lambda_{\eta'})-\eta'}{\delta\tilde{\lambda}_\eta}},$$

where again we use Observation A and the fact

$$\frac{(1-\delta)(1-\lambda_\eta)-\eta}{\delta\tilde{\lambda}_\eta} > \frac{(1-\delta)(1-\lambda_{\eta'})-\eta'}{\delta\tilde{\lambda}_\eta}.$$

Since $\tilde{\varphi}_\eta \in \mathcal{I}_1$, we have that:

$$\frac{V_{\overline{q}^1}\left(\tilde{\varphi}_\eta, w\left(\tilde{\varphi}_\eta; \frac{(1-\delta)(1-\lambda_\eta)-\eta}{\delta\tilde{\lambda}_\eta}\right)\right) - V_{\overline{q}^1}\left(\tilde{\varphi}_\eta, w\left(\tilde{\varphi}_\eta; \frac{(1-\delta)(1-\lambda_\eta)-\eta}{\delta\tilde{\lambda}_\eta} - \frac{(1-\delta)(1-\lambda_{\eta'})-\eta'}{\delta\tilde{\lambda}_\eta}\right)\right)}{\frac{(1-\delta)(1-\lambda_{\eta'})-\eta'}{\delta\tilde{\lambda}_\eta}}$$
$$\leqslant \frac{V_{\overline{q}^1}\left(\tilde{\varphi}_\eta, w\left(\tilde{\varphi}_\eta; \frac{(1-\delta)(1-\lambda_\eta)-\eta}{\delta\tilde{\lambda}_\eta}\right)\right) - V_{\overline{q}^1}\left(\tilde{\varphi}_\eta, \overline{m}_{\overline{q}^1}(\tilde{\varphi}_\eta)\right)}{\frac{(1-\delta)(1-\lambda_\eta)-\eta}{\delta\tilde{\lambda}_\eta}},$$

where the inequality follows from our previous argument on the interval $\mathcal{I}_1$.

It follows that:

$$\frac{V_{\overline{q}^1}(p,w(p)) - V_{\overline{q}^1}\left(p, w(p) - \frac{(1-\delta)(1-\lambda_{\eta'})-\eta'}{\delta}\left[\overline{m}_{\overline{q}^1}(1) - u(a^*,1)\right]\right)}{\frac{(1-\delta)(1-\lambda_{\eta'})-\eta'}{\delta}}$$

$$\leqslant \frac{V_{\overline{q}^1}(p,w(p)) - V_{\overline{q}^1}\left(p, w(p) - \frac{(1-\delta)(1-\lambda_{\eta})-\eta}{\delta}\left[\overline{m}_{\overline{q}^1}(1) - u(a^*,1)\right]\right)}{\frac{(1-\delta)(1-\lambda_{\eta})-\eta}{\delta}}.$$

From Equation (12), we then have that

$$\frac{1}{\eta}\left[V_{\overline{q}^1}(p,\overline{m}_{\overline{q}^1}(p)) - V_{\overline{q}^1}(p,\mathbf{w}(p;\eta))\right] =$$

$$\frac{(1-\delta)(1-\lambda_\eta)}{\eta}v(a^*,1) + \left[\frac{(1-\delta)(1-\lambda_\eta)}{\eta} - 1\right]\frac{V_{\overline{q}^1}(p,w(p)) - V_{\overline{q}^1}\left(p, w(p) - \frac{(1-\delta)(1-\lambda_{\eta})-\eta}{\delta}\left[m(1) - u(a^*,1)\right]\right)}{\frac{(1-\delta)(1-\lambda_{\eta})-\eta}{\delta}}$$

$$\geqslant \frac{(1-\delta)(1-\lambda_\eta)}{\eta}v(a^*,1) + \left[\frac{(1-\delta)(1-\lambda_\eta)}{\eta} - 1\right]\frac{V_{\overline{q}^1}(p,w(p)) - V_{\overline{q}^1}\left(p, w(p) - \frac{(1-\delta)(1-\lambda_{\eta'})-\eta'}{\delta}\left[\overline{m}_{\overline{q}^1}(1) - u(a^*,1)\right]\right)}{\frac{(1-\delta)(1-\lambda_{\eta'})-\eta'}{\delta}}$$

$$= \frac{(1-\delta)(1-\lambda_\eta)}{\eta}v(a^*,1) + \left[1 - \frac{(1-\delta)(1-\lambda_\eta)}{\eta}\right]\frac{V_{\overline{q}^1}\left(p, w(p) - \frac{(1-\delta)(1-\lambda_{\eta'})-\eta'}{\delta}\left[\overline{m}_{\overline{q}^1}(1) - u(a^*,1)\right]\right) - V_{\overline{q}^1}(p,w(p))}{\frac{(1-\delta)(1-\lambda_{\eta'})-\eta'}{\delta}}$$

$$= \frac{(1-\delta)(1-\lambda_{\eta'})}{\eta'}v(a^*,1) + \left[1 - \frac{(1-\delta)(1-\lambda_{\eta'})}{\eta'}\right]\frac{V_{\overline{q}^1}\left(p, w(p) - \frac{(1-\delta)(1-\lambda_{\eta'})-\eta'}{\delta}\left[\overline{m}_{\overline{q}^1}(1) - u(a^*,1)\right]\right) - V_{\overline{q}^1}(p,w(p))}{\frac{(1-\delta)(1-\lambda_{\eta'})-\eta'}{\delta}}$$

$$+ \left[\frac{(1-\delta)(1-\lambda_{\eta'})}{\eta'} - \frac{(1-\delta)(1-\lambda_\eta)}{\eta}\right]\left[\frac{V_{\overline{q}^1}\left(p, w(p) - \frac{(1-\delta)(1-\lambda_{\eta'})-\eta'}{\delta}\left[\overline{m}_{\overline{q}^1}(1) - u(a^*,1)\right]\right) - V_{\overline{q}^1}(p,w(p))}{\frac{(1-\delta)(1-\lambda_{\eta'})-\eta'}{\delta}} - v(a^*,1)\right]$$

$$\geqslant \frac{1}{\eta'}\left[V_{\overline{q}^1}(p,\overline{m}_{\overline{q}^1}(p)) - V_{\overline{q}^1}(p,w(p;\eta'))\right],$$

where the last inequality follows from:

$$\frac{V_{\overline{q}^1}\left(p, w(p) - \frac{(1-\delta)(1-\lambda_{\eta'})-\eta'}{\delta}\left[\overline{m}_{\overline{q}^1}(1) - u(a^*,1)\right]\right) - V_{\overline{q}^1}(p,w(p))}{\frac{(1-\delta)(1-\lambda_{\eta'})-\eta'}{\delta}}$$

$$= \frac{\tilde{\lambda}_{\eta'}V_{\overline{q}^1}(\tilde{\varphi}_{\eta'},\overline{m}_{\overline{q}^1}(\tilde{\varphi}_{\eta'})) - \tilde{\lambda}_{\eta'}V_{\overline{q}^1}\left(\tilde{\varphi}_{\eta'}, w\left(\tilde{\varphi}_{\eta'}; \frac{(1-\delta)(1-\lambda_{\eta'})-\eta'}{\delta\tilde{\lambda}_{\eta'}}\right)\right)}{\frac{(1-\delta)(1-\lambda_{\eta'})-\eta'}{\delta}} \geqslant v(a^*,1).$$

We now show that the the gradient $\mathcal{G}(p;\eta)$ is smaller than $v(a^*,1)$ for any $\eta \leq \bar{\eta}(p)$. From Equation (12), we have that:

$$\frac{1}{\eta}\left[V_{\overline{q}^1}(p,\overline{m}_{\overline{q}^1}(p)) - V_{\overline{q}^1}(p,\mathbf{w}(p;\eta))\right]$$
$$= \frac{(1-\delta)(1-\lambda_\eta)}{\eta}v(a^*,1) - \left[\frac{(1-\delta)\,(1-\lambda_\eta)}{\eta} - 1\right]\frac{V_{\overline{q}^1}\left(p, w(p) - \frac{(1-\delta)(1-\lambda_\eta)-\eta}{\delta}\left[m(1)-u(a^*,1)\right]\right) - V_{\overline{q}^1}(p,w(p))}{\frac{(1-\delta)(1-\lambda_\eta)-\eta}{\delta}}$$
$$= v(a^*,1) - \left[\frac{(1-\delta)\,(1-\lambda_\eta)}{\eta} - 1\right]\left[\frac{V_{\overline{q}^1}\left(p, w(p) - \frac{(1-\delta)(1-\lambda_\eta)-\eta}{\delta}\left[m(1)-u(a^*,1)\right]\right) - V_{\overline{q}^1}(p,w(p))}{\frac{(1-\delta)(1-\lambda_\eta)-\eta}{\delta}} - v(a^*,1)\right]$$
$$= v(a^*,1) - \left[\frac{(1-\delta)\,(1-\lambda_\eta)}{\eta} - 1\right]\left[\frac{\tilde{\lambda}_\eta V_{\overline{q}^1}(\tilde{\varphi}_\eta,\overline{m}_{\overline{q}^1}(\tilde{\varphi}_\eta)) - \tilde{\lambda}_\eta V_{\overline{q}^1}\left(\tilde{\varphi}_\eta, w\left(\tilde{\varphi}_\eta;\frac{(1-\delta)(1-\lambda_\eta)-\eta}{\delta\tilde{\lambda}_\eta}\right)\right)}{\frac{(1-\delta)(1-\lambda_\eta)-\eta}{\delta}} - v(a^*,1)\right]$$
$$= v(a^*,1) - \underbrace{\left[\frac{(1-\delta)\,(1-\lambda_\eta)}{\eta} - 1\right]}_{\geq 0}\underbrace{\left[\frac{V_{\overline{q}^1}(\tilde{\varphi}_\eta,\overline{m}_{\overline{q}^1}(\tilde{\varphi}_\eta)) - V_{\overline{q}^1}\left(\tilde{\varphi}_\eta, w\left(\tilde{\varphi}_\eta;\frac{(1-\delta)(1-\lambda_\eta)-\eta}{\delta\tilde{\lambda}_\eta}\right)\right)}{\frac{(1-\delta)(1-\lambda_\eta)-\eta}{\delta\tilde{\lambda}_\eta}} - v(a^*,1)\right]}_{\geq 0}$$
$$\leqslant v(a^*,1),$$

where the inequality follows from the fact that $\tilde{\varphi}_\eta \leq \underline{p}$ (therefore, from our arguments on the interval $\mathcal{I}_1$, where we show that the gradient is larger than $v(a^*,1)$).

Finally, we can use a similar decomposition as in the case $p \in \mathcal{I}_2$ to prove that the gradient is increasing for all η.

Appendix B. A Formal Discussion of Other Policies

B.1. **Non-uniqueness and comparison with the KG policy.** Our policy is not always uniquely optimal. We demonstrate the non-uniqueness with the help of a simple example and then discuss how our policy compares with the KG policy (for Kamenica-Gentzkow's policy).

Example 2.

The agent has two possible actions a_0 and a_1, with a_0 (resp., a_1) the agent's optimal action when the state is ω_0 (resp., ω_1). The principal wants to induce a_0 as often as possible, i.e., $a^* = a_0$. The discount factor is $1/2$. The payoffs are in Table 3, with the first coordinate corresponding to the principal's payoff.

TABLE 3. Payoff table of Example 2

	a_0	a_1
ω_0	$1,1$	$0,0$
ω_1	$1,0$	$0,1$

In Example 2, we have that: $m(p) = \max(1-p,p)$, $M(p) = 1$ and $u(a^*,p) = 1-p$. Thus, a^* is optimal for all $p \in P = [0,1/2]$. Moreover, $Q^1 = [0,2/3]$ and $\mathrm{w}(p) = 3p-1$ for $p \in (1/2,2/3]$.

We now provide an explicit characterization of the value function. We first compute the value function $V_{\overline{q}^1}(\cdot,m(\cdot))$ and check whether it is concave. For $p \in [0,1/2]$, the policy recommends a^* and promises a continuation payoff of $m(p)$. That is, since a^* is optimal, the principal does not need to incentivize the agent. For $p \in (1/2,2/3]$, the policy recommends a^* and promises a continuation payoff of $\mathrm{w}(p)$. At $(p,\mathrm{w}(p))$ with $p \in (1/2,2/3]$, the policy splits p into $\varphi(p,\mathrm{w}(p))$ and 1, with probability $\lambda(p,\mathrm{w}(p))$ and $1-\lambda(p,\mathrm{w}(p))$ respectively. (See Equation (3).)

We obtain that $\lambda(p,\mathrm{w}(p)) = (3-4p)$ and $\varphi(p,\mathrm{w}(p)) = \frac{2-3p}{3-4p}$. Note that $\varphi(p,\mathrm{w}(p)) = \frac{2-3p}{3-4p} < \frac{1}{2}$ since $p \in (1/2,2/3]$. After splitting p into $\varphi(p,\mathrm{w}(p))$, the principal therefore obtains a payoff of 1 in all subsequent periods. It follows that the principal's expected payoff is

$$\frac{1}{2} + \frac{1}{2}\lambda(p,\mathrm{w}(p)) = 2(1-p).$$

Finally, if $p \in (2/3,1]$, the policy splits p into $2/3$ and 1 with probability $3(1-p)$ and $(1-3(1-p))$, respectively. The principal's expected payoff is then

$$3(1-p) \times \left[\frac{1}{2} + \frac{1}{2}\lambda\left(\frac{2}{3},\mathrm{w}\left(\frac{2}{3}\right)\right)\right] = 3(1-p) \times 2\left(1-\frac{2}{3}\right) = 2(1-p).$$

So, the value function $V_{\overline{q}^1}$ induced by the policy $\tau_{\underline{q}^1}$ is such $V_{\overline{q}^1}(p,m(p)) = 1$ for all $p \in [0,1/2]$ and $V_{\overline{q}^1}(p,m(p)) = 2(1-p)$ for all $p \in (1/2,1]$. Since it is concave in p, this guarantees that $q^* = \overline{q}^1$ and, thus, the policy is indeed optimal.

We now consider another policy, which we call the KG policy. The aim of the KG policy is to persuade the agent to choose a^* as often as possible by disclosing information at the initial stage only. The best payoff the principal can obtain with a KG policy is:

$$\max_{(\lambda_s, p_s, a_s)} \sum_s \lambda_s v(a_s, p_s),$$

subject to

$$\forall s, u(a_s, p_s) \geq m(p_s), \text{and} \sum_s \lambda_s p_s = p.$$

In Example 2, the KG policy differs from our policy only when $p \geq 1/2$, and consists in splitting p into $1/2$ and 1, with probability $2(1-p)$ and $1-2(1-p)$ respectively. The KG policy induces the same value function as our policy, hence is also optimal. We now prove that this is not accidental.

Suppose that there are only two actions, a_0 and a_1, such that a_0 (resp., a_1) is optimal at state ω_0 (resp., ω_1). The principal aims at implementing a_0 as often as possible, i.e., $a^* = a_0$.[16] Remember that a_0 is optimal at all beliefs in $[\underline{p}, \overline{p}]$. Since a_0 is optimal at 0, $\underline{p} = 0$. To streamline the exposition, assume that the prior $p_0 > \overline{p}$. (If $p_0 \leq \overline{p}$, an optimal policy is to never reveal any information.) It is then immediate to see that the KG policy consists in splitting the prior p_0 into $\overline{p}$ and 1, with probability $\frac{1-p_0}{1-\overline{p}}$ and $1 - \frac{1-p_0}{1-\overline{p}}$, respectively. Intuitively, the principal designs a binary experiment, with one signal perfectly informing the agent that the state is ω_1 and the other partially informing the agent so that his posterior beliefs is $\overline{p}$.

We can contrast the KG policy with our policy. Unlike the KG policy, our policy does not reveal information to the agent at the first period, and only reveals information to the agent if he plays a_0. If the agent plays a_0 at the first period, the policy splits p_0 into $\varphi(p_0, \mathrm{w}(p_0))$ and 1 with probability $\lambda(p_0, \mathrm{w}(p_0))$ and $1 - \lambda(p_0, \mathrm{w}(p_0))$, respectively. Note that $\varphi(p_0, \mathrm{w}(p_0)) \leq \overline{p}$ since $\mathrm{w}(p_0) \geq m(p_0)$. Thus, our policy guarantees that the agent plays a^* for sure at the first period. However, this comes at a cost: the principal needs to reveal more information to the agent at the next period and, consequently, inducing the agent to play a_0 with a lower probability. Somewhat surprisingly, both policies are optimal, regardless of the discount factor.

[16]If $a^* = a_1$, then $0 = m(1) - u(a_1, 1) \geq (m(0) - u(a_1, 0))\frac{v(a^*,1)}{v(a^*,0)} = (u(a_0, 0) - u(a_1, 0))\frac{v(a^*,1)}{v(a^*,0)} \geq 0$, i.e., a_1 is also optimal when the agent believes that the state is ω_0 with probability 1.

Corollary 2. *If there are only two actions, then the KG policy is also optimal.*

As Example 1 shows, the KG policy is not always optimal. Yet, if a^* is not strictly dominated and the function m is linear from $\overline{p}$ to 1, then the KG policy is also optimal at all priors above $\overline{p}$. (A proof is available upon request.) More generally, whenever the value function V^* is linear in (p, w), the KG policy is also optimal. We conjecture, however, that the value function V^* is generically non-linear.

B.2. **Comparison with the "random disclosure" policy.** Remember that the policy of fully disclosing the state with delay plays a prominent role in the work of Ball (2019) and Orlov et al. (2019). Since we study a discrete time model, we do not directly compare our policy with the policy of fully disclosing the state with delay, but with the "random disclosure" policy. The "random disclosure" policy consists in fully disclosing the state with probability α at period $t+1$ (and to withhold all information with the complementary probability) if the agent plays a^* at period t.[17]

We first compute the principal's payoff if he commits to the best "random disclosure" policy. To ease the exposition, we assume that a^* is not optimal at the belief $p = 0$.[18] Assume that $p \in Q^1$. The best "random disclosure" policy is solution to the maximization problem:

$$V = \max_{\alpha \in [0,1]} (1-\delta) v(a^*, p) + \delta(1-\alpha)V,$$

subject to

$$U = (1-\delta)u(a^*, p) + \delta\left[\alpha M(p) + (1-\alpha)U\right] \geq m(p).$$

The optimal solution is

$$\alpha^* = \frac{\mathbf{w}(p) - m(p)}{M(p) - m(p)} = \frac{1-\delta}{\delta} \frac{m(p) - u(a^*, p)}{M(p) - m(p)},$$

[17]In continuous time, the policy of fully disclosing the state with delay yields the same payoff as the "random disclosure" policy.

[18]When a^* is optimal at $p = 0$, we need to add the term $\delta\alpha(1-p)v(a^*, p)$ to the objective, which corresponds to the payoff the principal obtains when the disclosed state is ω_0.

inducing the value

$$(1-\delta)\sum_t \delta^t \left(\frac{M(p)-\mathbf{w}(p)}{M(p)-m(p)}\right)^t v(a^*,p) = \frac{M(p)-m(p)}{M(p)-u(a^*,p)} v(a^*,p).$$

The formula has a natural interpretation. Whenever the agent is recommended to play a^*, no information has been revealed yet, so that the maximal value of information the principal can create is $M(p)-m(p)$. To incentivize the agent, the principal needs to promise a continuation payoff of $\mathbf{w}(p)$ in the future and thus needs to create an information value of $\mathbf{w}(p)-m(p)$. To create an information value of $\mathbf{w}(p)-m(p)$, the principal commits to fully disclose the state with some probability, hence foregoing the opportunity to incentivize the agent to play a^* in the future. Therefore, the highest probability with which the principal can incentivize the agent to play a^* is $(M(p)-\mathbf{w}(p))/(M(p)-m(p))$.

To understand why and when the principal can do better than following the "random recursive policy," we study the *relaxed* version of our problem, where only the (ex-ante) participation constraint needs to be satisfied. Consider the following policy. The principal discloses information at the ex-ante stage, i.e., chooses a splitting $(\lambda_s, p_s)_s$ of p, and recommends the agent to play a^* at all periods with probability β_s when the realized signal is s. We continue to assume that $p \in Q^1$. The policy satisfies the participation constraint if

$$\sum_s \lambda_s \left[\beta_s u(a^*,p_s) + (1-\beta_s)m(p_s)\right] \geq m(p).$$

We can rewrite the participation constraint as:

$$\sum_s \lambda_s (1-\beta_s)(m(p_s)-u(a^*,p_s)) \geq m(p)-u(a^*,p), \tag{14}$$

where $m(p_s)-u(a^*,p_s)$ is the opportunity cost of following the recommendation at belief p_s. The principal maximizes $\sum_s \lambda_s \beta_s v(a^*,p_s)$ subject to the participation constraint. Clearly, the participation constraint binds at a maximum. Moreover, since m is convex, the best for the principal is to fully disclose all information, i.e., to split p into 0 and 1.

Note that if the principal recommends a^* with the same probability at all s, his payoff is

$$\frac{M(p) - m(p)}{M(p) - u(a^*, p)} v(a^*, p),$$

which is precisely the payoff of the "random recursive" policy.[19]

The principal can do better by exploiting the difference in opportunity costs at the two extreme beliefs 0 and 1. Writing β_1 (resp., β_0) for the probability of recommending a^* conditional on the posterior being 1 (resp., 0), the principal maximizes $p\beta_1 v(a^*, 1) + (1-p)\beta_0 v(a^*, 0)$ subject to:

$$p\beta_1(m(1) - u(a^*, 1)) + (1-p)\beta_0(m(0) - u(a^*, 0)) \leq M(p) - m(p).$$

The right-hand side is the maximal value of information the principal can create, while the left-hand side is the expected opportunity cost of following the recommendation. As with the "random disclosure" policy, the principal needs to generate the maximal value of information; this is the maximal value the principal can use to incentivize the agent. However, unlike the "random disclosure" policy, the principal needs to use the surplus created asymmetrically, as it is easier to incentivize the agent in state ω_0 than ω_1.

More precisely, the problem is linear in (β_0, β_1). Therefore, since the slope $\frac{v(a^*,0)}{v(a^*,1)}$ is larger than the slope $\frac{m(0)-u(a^*,0)}{m(1)-u(a^*,1)}$, the optimal solution is to set β_0 as high as possible. For instance, if $M(p) - m(p) \leq (1-p)(m(0) - u(a^*, 0))$, the best is to set $(\beta_0, \beta_1) = (\frac{M(p)-m(p)}{(1-p)(m(0)-u(a^*,0))}, 0)$, resulting in a payoff of

$$\frac{M(p) - m(p)}{m(0) - u(a^*, 0)} v(a^*, 0) \geq \frac{M(p) - m(p)}{M(p) - u(a^*, p)} v(a^*, p),$$

with a strict inequality if the opportunity cost is strictly higher in state ω_1.[20] This is the solution to the relaxed constraint.

While our policy also needs to incentivize the agent to follow the recommendation, it exploits the same asymmetries in opportunity costs as the above policy, which explains why it outperforms the "random disclosure" policy.

[19] When a^* is optimal at $p = 0$, we need to add the term $(1-p)\left(1 - \frac{M(p)-m(p)}{M(p)-u(a^*,p)}\right) v(a^*, p)$.

[20] See Appendix B.4 for the full characterization.

To conclude, note that if $\frac{v(a^*,0)}{v(a^*,1)} = \frac{m(0)-u(a^*,0)}{m(1)-u(a^*,1)}$, then the random disclosure policy solves the relaxed problem and, therefore, is also optimal.

B.3. **Proof of Corollary 2.** We first compute the principal's payoff induced by our policy. To ease notation, we write φ for $\varphi(p, \mathbf{w}(p))$. We first assume that $q^* = \underline{q}^1$, compute the value function $V_{\overline{q}^1}(p, m(p))$ for all p and check that it is concave. By construction, the principal's payoff satisfies:

$$V_{\overline{q}^1}(p, m(p)) = (1-\delta)v(a^*, p) + \delta V_{\overline{q}^1}(p, \mathbf{w}(p)) = (1-\delta)v(a^*, p) + \delta \frac{1-p}{1-\varphi} v(a^*, \varphi).$$

Remember that

$$\mathbf{w}(p) = \frac{m(p) - (1-\delta)u(a_0, p)}{\delta} = \frac{1-p}{1-\varphi} m(\varphi) + \frac{p-\varphi}{1-\varphi} m(1).$$

Since $\mathbf{w}(p) = m(p) = u(a_0, p)$ when $p \leq \overline{p}$, we have that $\varphi = p$ and, therefore, the principal payoff is 1 when $p \leq \overline{p}$. Assume that $p > \overline{p}$. We have that:

$$\mathbf{w}(p) = \frac{u(a_1, p) - (1-\delta)u(a_0, p)}{\delta} = \frac{1-p}{1-\varphi} u(a_0, \varphi) + \frac{p-\varphi}{1-\varphi} u(a_1, 1),$$

since $m(\varphi) = u(a_0, \varphi)$ and $\varphi \leq \overline{p}$. (To see this, if $\varphi > \overline{p}$, then $m(\varphi) = u(a_1, \varphi)$, hence $\mathbf{w}(p) = m(p)$, a contradiction with $\mathbf{w}(p) > m(p)$ when $p > \overline{p}$.) The above equation is equivalent to:

$$(1-\varphi)[u(a_1, p) - (1-\delta)u(a_0, p)] = \delta[(1-p)u(a_0, \varphi) + (p-\varphi)u(a_1, 1)].$$

Observing that $u(a, p) = (1-p)(u(a, 0) - u(a, 1)) + u(a, 1)$ for all a and, similarly, for φ, we can simplify the above expression to

$$\delta \frac{1-p}{1-\varphi} = \delta - p + (1-p) \frac{u(a_0, 0) - u(a_1, 0)}{u(a_1, 1) - u(a_0, 1)}.$$

Lastly, remember that the threshold $\overline{p}$ is given by:

$$1 - \overline{p} = \frac{u(a_1, 1) - u(a_0, 1)}{u(a_0, 0) - u(a_0, 1) + u(a_1, 1) - u(a_1, 0)},$$

and, therefore,

$$\begin{aligned} V_{\overline{q}^1}(p, m(p)) &= v(a^*, p) + \delta\left(1 - \frac{1-p}{1-\varphi}\right) v(a^*, 1) \\ &= \frac{1-p}{1-\overline{p}} v(a^*, \overline{p}) + \left[1 - \frac{1-p}{1-\overline{p}} + \delta\left(1 - \frac{1-p}{1-\varphi}\right)\right] v(a^*, 1) \\ &= \frac{1-p}{1-\overline{p}} v(a^*, \overline{p}). \end{aligned}$$

Since the KG policy induces the same payoff, it is also optimal.

B.4. **First best.** This section provides detail about the solution to the first-best problem, which we study when comparing our policy with the random disclosure policy. Let

$$\alpha_1^* = 1 - \frac{m(p) - u(a^*, p)}{p(m(1) - u(a^*, 1))} = \frac{M(p) - m(p) - (1-p)(m(0) - u(a^*, 0))}{p(m(1) - u(a^*, 1))}.$$

Note that $\alpha_1^* \leq 1$, with equality if $m(p) = u(a^*, p)$), and $\alpha_1^* < 0$ if $M(p) - m(p) - (1-p)(m(0) - u(a^*, 0)) < 0$.

At an optimum, the participation constraint clearly binds. If $m(0) - u(a^*, 0) = 0$, the solution is clearly $(1, \frac{M(p)-m(p)}{p(m(1)-u(a^*,1))})$. Assume that $m(0) - u(a^*, 0) > 0$. We can rewrite the principal's objective as a function of α_1:

$$\begin{cases} p\alpha_1 v(a^*, 1) + (1-p) v(a^*, 0) & \text{if } \alpha_1 \leq \max(0, \alpha_1^*), \\ p\alpha_1 \left(v(a^*, 1) - v(a^*, 0) \frac{m(1)-u(a^*,1)}{m(0)-u(a^*,0)}\right) + \frac{M(p)-m(p)}{m(0)-u(a^*,0)} v(a^*, 0) & \text{if } \max(0, \alpha_1^*) \leq \alpha_1 \leq \frac{M(p)-m(p)}{p(m(1)-u(a^*,1))}, \\ -\infty & \text{otherwise}. \end{cases}$$

Note that the objective is continuous in α_1. The optimal payoff is therefore:

$$p \max(0, \alpha_1^*) v(a^*, 1) + (1-p) \max\left(\frac{M(p) - m(p)}{(1-p)(m(0) - u(a^*, 0))}, 1\right) v(a^*, 0),$$

obtained with $(\alpha_0, \alpha_1) = \left(\frac{M(p)-m(p)}{(1-p)(m(0)-u(a^*,0))}, 0\right)$ if $\frac{M(p)-m(p)}{(1-p)(m(0)-u(a^*,0))} \leq 1$ and $(\alpha_0, \alpha_1) = (1, \alpha_1^*)$, otherwise.

References

AU, P. H. (2015): "Dynamic Information Disclosure," *The RAND Journal of Economics*, 46, 791–823.

AUMANN, R. J., M. B. MASCHLER, AND R. E. STEARNS (1995): *Repeated Games with Incomplete Information*, Cambridge (Mass.) London: MIT press.

BALL, I. (2019): "Dynamic Information Provision: Rewarding the Past and Guiding the Future," *Available at SSRN 3103127*.

BIZZOTTO, J., J. RÜDIGER, AND A. VIGIER (Forthcoming): "Dynamic Persuasion with Outside Information," *American Economic Journal: Microeconomics*.

CHE, Y.-K., K. KIM, AND K. MIERENDORFF (2020): "Keeping the Listener Engaged: a Dynamic Model of Bayesian Persuasion," *arXiv preprint arXiv:2003.07338*.

CRAWFORD, V. P. AND J. SOBEL (1982): "Strategic Information Transmission," *Econometrica*, 50, 1431–1451.

ELY, J. AND M. SZYDLOWSKI (2020): "Moving the Goalposts," *Journal of Political Economy*, 128.

ELY, J. C. (2015): "Beeps," Tech. rep., Northwestern University.

——— (2017): "Beeps," *American Economic Review*, 107, 31–53.

ESCUDE, M. AND L. SINANDER (2020): "Strategic Research Funding," *Working Paper*.

FUDENBERG, D. AND L. RAYO (2019): "Training and Effort Dynamics in Apprenticeship," *American Economic Review*, 109, 3780–3812.

GARICANO, L. AND L. RAYO (2017): "Relational Knowledge Transfers," *American Economic Review*, 107, 2695–2730.

HENRY, E. AND M. OTTAVIANI (2019): "Research and the Approval Process: The Organization of Persuasion," *American Economic Review*, 109, 911–55.

KAMENICA, E. (2019): "Bayesian persuasion and information design," *Annual Review of Economics*, 11, 249–72.

KAMENICA, E. AND M. GENTZKOW (2011): "Bayesian Persuasion," *American Economic Review*, 101, 2590–2615.

MAKRIS, M. AND L. RENOU (2021): "Information Design in Multi-Stage Games," ArXiv:2102.13482.

ORLOV, D., A. SKRZYPACZ, AND P. ZRYUMOV (2019): "Persuading the principal to wait," *forthcoming Journal of Political Economics*.

RENAULT, J., E. SOLAN, AND N. VIEILLE (2017): "Optimal dynamic information provision," *Games and Economic Behavior*, 104, 329 – 349.

SMOLIN, A. (2018): "Dynamic Evaluation Design," *R&R at American Economic Journal: Microeconomics*.

SPEAR, S. E. AND S. SRIVASTAVA (1987): "On Repeated Moral Hazard with Discounting," *The Review of Economic Studies*, 54, 599–617.

WEI ZHAO, HEC PARIS AND GREGHEC-CNRS, 1 RUE DE LA LIBÉRATION, 78351 JOUY-EN-JOSAS, FRANCE

Email address: wei.zhao1(at)hec.edu

CLAUDIO MEZZETTI, SCHOOL OF ECONOMICS, THE UNIVERSITY OF QUEENSLAND, LEVEL 6, COLIN CLARK BUILDING, ST LUCIA, BRISBANE QLD 4072, AUSTRALIA

Email address: c.mezzetti(at)uq.edu.au

LUDOVIC RENOU, QUEEN MARY UNIVERSITY OF LONDON, CEPR AND UNIVERSITY OF ADELAIDE, MILES END, E1 4NS, LONDON, UK

Email address: lrenou.econ(at)gmail.com

TRISTAN TOMALA, HEC PARIS AND GREGHEC-CNRS, 1 RUE DE LA LIBÉRATION, 78351 JOUY-EN-JOSAS, FRANCE

Email address: tomala(at)hec.fr