Information Design in Repeated Interaction*

João Correia-da-Silva† Takuro Yamashita‡

December 5, 2024

Abstract

How does a long-term relationship affect communication among economic entities? In this paper, we study dynamic information design in repeated interaction, where the state is imperfectly persistent. We observe that dynamics introduce a novel distortion relative to the static counterpart. Even if both parties' preferences are aligned in some state, revealing that state (which is myopically optimal for the informed) could make it difficult to persuade the uninformed in future periods (when the state switched to those with preference misalignment). This motif can distort the information in two different ways. If the state is moderately persistent, then the optimal mechanism may exhibit inefficient bad news; while if the state is very persistent, then it may exhibit inefficient good news.

*We thank the audiences at Minho, Tokyo, Doshisha, Osaka (OEIO), UECE 2023, EWET 2023, PEJ 2023, and SAET 2023. João Correia-da-Silva acknowledges financing by Portuguese public funds through Fundação para a Ciência e a Tecnologia (2022.09548.PTDC and UIDB/04105/2020). Takuro Yamashita acknowledges the funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 714693), ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program), and JSPS KAKENHI (JP23K01311, JP24K04841).

†University of Porto, School of Economics and Management, CEF.UP. joao@fep.up.pt

‡Osaka School of International Public Policy, Osaka University. yamashita.takuro.osipp@osaka-u.ac.jp

1 Introduction

How does a long-term relationship affect communication among relevant economic entities? For example, does a retailer better informed about its product quality disclose this information to a loyal customer? Does a public authority disclose information about a vaccine's side effect if periodic doses could be important for avoiding a future pandemic? To better understand this fundamental question, in this paper, we consider a repeated interaction of two parties, one having access to information about each period's payoff-relevant state, while the other not. In each period, the informed party can provide some information about the state (including its past realizations). Then, they play a simple "agreement game", where they earn a state-dependent payoff in each period if and only if both play an "agree" action, while they earn zeros otherwise. The set of states where the informed player earns a positive agreement payoff is a strict superset of that where the uninformed player earns a positive agreement payoff, representing their imperfect preference alignment. We work with this payoff environment because it is simple, yet includes various interesting economic applications, such as (stylized versions of) buyer-seller trading or public-good contribution (see Section 1.1).

We assume that the informed party has commitment power in designing a rule about information disclosure to the uninformed party, making the problem a *dynamic information design* problem. As a benchmark scenario, if it is a one-shot environment (or if both parties are myopic), then the optimal rule is obtained as a solution of a simple Bayesian persuasion exercise: They play the "agree" actions whenever their agreement payoffs are both positive; they do not whenever their agreement payoffs are both negative; and they agree with some probability when the informed earns a positive agreement payoff while the uninformed earns a negative agreement payoff, where this probability is chosen so that the uninformed player's obedience constraint (in case he is recommended to agree) is binding. Note that this outcome is *Pareto efficient* in the sense that, whenever they have the same preferred action profile, that is played for sure.

How would this change in case their relationship can continue? We find that *less* efficient communication may happen, in the sense explained below. That

is, a long-term relationship can be *detrimental* to communication. In fact, the basic intuition is quite simple. Compared to the static environment, the informed party's problem is more complicated, because today's information disclosure can affect the uninformed party's incentive in the future, in case of (imperfect) state persistence. More specifically, imagine that today's state is such that both players' agreement payoffs are negative. In the static counterpart, there is no reason to hide that information, because revealing it would lead to a non-agreement outcome, preferable to both.[1] However, in the dynamic setting where the state is *imperfectly* persistent, revealing that information today would make the uninformed party reluctant to agree *tomorrow*. This would not be ideal for the informed party, if the state tomorrow is likely to be one where his agreement payoff is positive (recall that there are states where the informed prefers agreement while the uninformed does not). If this concern is sufficiently severe, then the informed player may prefer recommending the agree action even if both players' agreement payoffs are negative ("*inefficient agreement*"). In fact, this concern could be so strong that, under certain conditions, we show that it is optimal for the informed player to not disclose any information at all (the *no-disclosure* mechanism).

Under some other conditions, the optimal mechanism may exhibit a different form of inefficiency, though based on the same concern. Instead of recommending the agree action in a state where both players' agreement payoffs are negative, the informed player can recommend the non-agree action in some states where both players' agreement payoffs are *positive*. That is, the optimal mechanism may exhibit "*inefficient disagreement*". The point of this inefficiency is to make the uninformed player *less pessimistic* in case non-agreement is recommended: Conditional on recommendation of the non-agree action, there is some probability that the state is one where the uninformed player's agreement payoff is positive today, implying (due to persistence) a higher probability that tomorrow's agreement payoff is positive. As a consequence, it becomes easier to persuade the uninformed player to agree tomorrow. As far as we know, this paper is the first to show that those two "opposite" forms of inefficiency (inefficient agreement / disagreement)

[1] The same is true in dynamic environments if the state is iid or perfectly persistent.

can be effective in a single environment, and under what conditions they are so.[2]

Although the intuition is as simple as explained above, the main theoretical challenge is to prove that a certain mechanism such as the no-disclosure mechanism is optimal. Even if the candidate mechanism is simple, in order to prove its optimality, in principle one must examine a large class of alternative mechanisms. Indeed, the "standard" approach in the existing studies would set up a value function as a function of multi-dimensional arguments including the uninformed player's belief about the current state and his promised continuation payoff (Abreu et al., 1990; Carrasco et al., 2019).[3] This multi-dimensionality makes it difficult to solve the value function, especially with a larger space of the payoff-state. We take an alternative approach based on Lagrangian (weak) duality, following, e.g., Ball (2023) and Lyu (2023). The basic idea is to formulate a Lagrangian adding the agents' (infinitely many) obedience constraints to the objective with some multipliers. An advantage of the Lagrangian problem over the original (constrained) problem is that there no longer exist *forward-looking constraints*, and hence the solution is time-consistent, admitting a Bellman-type recursive formulation.[4] Moreover, if the candidate optimal policy is simple, then the corresponding Bellman equation could also be simple.

1.1 Applications

In order to fix the idea, among possible applications, it would be useful to focus on one application, where a hospital recommends some form of preventive or palliative health care to a patient. This can take the form of regular medical examinations (e.g., blood tests, cancer screening, or psychotherapy) or medication (e.g., for asthma, hypertension, cancer, or mental depression). The hospital may be more informed about the effectiveness and possible side effects of health care

[2]Another way to think of the two forms of inefficiency is that inefficient agreement is to avoid sending the bad signal, while inefficient disagreement is to reduce the informational content of the bad signal.

[3]See also Spear and Srivastava (1987).

[4]Even if long-term commitment is necessary to implement the optimal policy, the fictitious Lagrangian-maximizer only needs single-period commitment (as in static information design) to solve the auxiliary problem (Lagrangian-maximization) whose solution is the optimal policy.

than the patient. Also, their preferences may not be fully aligned. For example, the hospital may care about the revenue and cost aspect differently from the patient; or there may exist some externality that only one party cares about, such as the patient's carers' absence from work and cost of admitting hospitalization. Given that periodic examinations and/or medicine intakes are necessary to reduce health risks, the hospital faces a dynamic information design problem about the effectiveness or side effects to the patient, as his current and future incentives to adhere to treatments depends on the way the hospital communicates with him.[5]

We explain the main insights of the paper mainly within this application. However, an agreement game has been considered as a stylized description of various economic environments, such as trading (between a seller and a buyer), matching (between an employer and an employee), public good contribution (among a public-good provider and residents), investment (between an investor and an entrepreneur), and so on. We believe that the main insights given in this paper would be useful in those other contexts too.

1.2 Related literature

Our paper is closely related to the literature of information design, in particular, its dynamic version. In *static* information design, an informed party commits to a mechanism which maps a payoff-relevant state to a signal. The uninformed party observes this signal, updates his belief about the state, and then takes an action.[6]

Its dynamic counterpart shows a large variety of models. Ball (2023) is the closest to ours, as he considers a designer with full commitment who observes an evolving state and seeks to persuade a forward-looking agent. His model is a continuous-time extension of the model by Crawford and Sobel (1982) with quadratic preferences, where the state follows a (possibly explosive) Ornstein-

[5]The way the hospital's physicians should communicate with a patient is often guided by a third party, such as a medical association the physicians belong to. Our information design problem may also be interpreted as the association's optimal guideline design problem.

[6]See Bergemann and Morris (2019) and Kamenica (2019) for the surveys. It is related to a *cheap-talk* problem, which concerns the same setup but without the informed party's commitment power. It also differs from *(costly) signaling* where the informed party's signal takes the form of a payoff-relevant action.

Uhlenbeck process,[7] and the agent continually adjusts the action. Their optimal policy turns out to be a *delayed-revelation* policy: the principal discloses the true realization of date t without noise, but only at a later date $t + \tau$. Moreover, this delay τ vanishes to 0 in finite time.

We consider a different model of repeated interaction: a discrete-time, binary-action environment with flexible agreement payoffs.[8] Though the binary-action assumption is a restriction relative to Ball (2023), the flexibility in the payoff structure enriches the possible properties of the optimal mechanism: We show that the optimal mechanism may exhibit *inefficient agreement* or *inefficient disagreement* depending on the parameters; while in Ball (2023), the optimal mechanism always exhibits the delayed disclosure, based on the crucial observation that the players only care about the future paths of the bias and variance under the quadratic payoff structure. In this sense, our paper contributes to the study of information design in repeated environments, complementing Ball (2023).

Zhao et al. (2024) investigated a related environment, but with a fully persistent binary state. Assuming that the principal has a (state-independent) preferred action and is indifferent across all other actions, they were able to solve this problem using the standard approach (with the agent's belief and continuation payoff as arguments of the principal's value function), and concluded that the principal's preferred action is either played in all periods or until the agent learns the state.

The model of Renault et al. (2017) is closer to ours in that the state follows a discrete-time Markov process. However, they assume that the agent is myopic (equivalently, that the designer faces a sequence of short-lived agents) and that the principal's payoff only depends on the action (does not depend on the state).[9] They show that the policy that maximizes current-period payoff ("greedy policy") is optimal, as long as the prior is not too distant from the invariant distribution

[7] An Ornstein-Uhlenbeck process is a continuous-time analogue of an AR(1) process.

[8] Our model is also different from his in that the informed player does not only send a costless message to the uninformed player, but also plays a payoff-relevant action.

[9] A myopic agent does not care about future payoff, and thus the designer's ability to credibly promise some future payoff (which is made possible by long-term commitment) becomes unimportant. Technically, in this case the designer's problem admits a recursive formulation without the promised future payoff of the agent as a state variable (and a Lagrangian formulation without lagged multipliers).

of the Markov process.

Other existing contributions mostly address a stopping problem, a different form of dynamic games: at each point in time, the agent decides whether to continue or to choose an action that ends the game (Au, 2015; Ely, 2017; Nikandrova and Pancs, 2018; Che and Mierendorff, 2019; Ely and Szydlowski, 2020; Orlov et al., 2020; Zhao, 2022). A common idea is that the optimal dynamic information disclosure generates a distribution over posteriors such that the agent is made either exactly indifferent to continue the interaction, or made sufficiently pessimistic so that he quits immediately (in case the continuation is the designer's desired action).

Following some papers in the literature, we identify the optimal mechanisms based on Lagrangian duality. This can be advantageous in dynamic information design: as briefly explained above, the alternative value-function approach based on Abreu et al. (1990) must solve a multi-dimensional functional equation, which can be more complicated and intractable. The above mentioned paper by Ball (2023) applies a version of the Lagrangian approach in his continuous-time environment. Lyu (2023) emphasizes the usefulness of Lagrangian duality in the study of a dynamic persuasion problem with a sequence of myopic agents, whose actions influence information generation.[10] The Lagrangian-based approach is also used in the dynamic macro / public-finance literature. For example, Marcet and Marimon (2019) establish a strong-duality result, and a recursive formulation for a saddle-point value function.

2 Model

2.1 Basic ingredients

There exist two players: Principal and Agent. Time is discrete, denoted by $t = 0, 1, 2, \ldots$. The players have a common discount factor $\delta \in (0, 1)$.

[10] The model of Lyu (2023) also differs from ours in that the state is perfectly persistent and gradually learned by the principal (instead of imperfectly persistent and instantaneously observed by the principal).

At each t, the payoff-state $\theta_t \in \Theta$ (finite) is realized, according to a first-order Markov distribution $P = (p_{i\to j})_{i,j}$, where $p_{i\to j} = \mathbb{P}(\theta_t = j|\theta_{t-1} = i)$ for each $i, j \in \Theta$ and each $t \geq 1$. Let μ denote the steady-state distribution (i.e., $\mu = P\mu$),[11] and assume that the prior for θ_0 is the same as μ.

Only Principal observes θ_t, which happens at the beginning of time t. Principal then sends a message to Agent. Based on the revelation principle,[12] we assume (without loss) that the message takes the form of a *recommendation* of Agent's action $a_t \in \{0, 1\}$. The essential role of this communication is to (possibly imperfectly) disclose the information about the (current and/or past) state to Agent. We explain later more in detail how the disclosure formally works.

After the disclosure, they play an *agreement game* as a stage game: Principal plays $q_t \in \{0, 1\}$, and then (after observing q_t) Agent plays $a_t \in \{0, 1\}$. Their action 1 is interpreted as an "agree" action, and action 0 is a "disagree" action.[13] For example, a physician (Principal) offers a certain type of medication ($q_t = 1$) or not ($q_t = 0$); and a patient (Agent) follows it ($a_t = 1$) or not ($a_t = 0$). Principal's stage-game payoff is $q_t a_t v(\theta_t)$, while Agent's is $q_t a_t u(\theta_t)$. That is, they earn state-dependent payoffs if both agree, and 0 otherwise.

Let $(\emptyset \neq)\Theta_A \subsetneq \Theta_P \subsetneq \Theta$ be such that

$$\begin{aligned} v(\theta) \geq 0 \quad &\Leftrightarrow \quad \theta \in \Theta_P \\ u(\theta) \geq 0 \quad &\Leftrightarrow \quad \theta \in \Theta_A. \end{aligned}$$

The assumption that $\Theta_A \neq \Theta_P$ reflects partial (mis)alignment of their preferences. In particular, Principal is more biased toward the agreement as $\Theta_A \subsetneq \Theta_P$.[14] Note that we do not assume that $v(\theta) \geq u(\theta)$ for all $\theta \in \Theta$. To simplify the analysis, let us assume that $v(\theta), u(\theta) \neq 0$ for all $\theta \in \Theta$. The cases where some $v(\theta)$

[11] Pick one arbitrarily as μ if there are multiple steady-state distributions.

[12] In our environment, there is no chance that Agent observes a probability-zero event. Thus, we adopt Nash / Bayesian equilibrium as a solution concept. See Forges (1986).

[13] In the application we will consider, obedience constraints for action 0 are not binding. In such cases, it is equivalent to allow Principal to terminate the relationship ($q_t = 0$ terminates the game, while $q_t = 1$ continues the game) instead of choosing an agree/disagree action.

[14] The assumption that $\Theta_P \subsetneq \Theta$ is only necessary for the possibility of *inefficient agreement* (action 1 being chosen in a state in which both players prefer action 0).

or $u(\theta)$ can be 0 are similar, but more complicated.

Thus, the timing of the stage game is summarized as follows: At each t:

(t.0) Principal observes θ_t.

(t.1) Principal takes action q_t (publicly observed), and recommends action $\tilde{a}_t$.

(t.2) Agent takes action a_t (publicly observed).

The (full) history of the game up until (t.0) is denoted by:

$$h_t = ((\theta_\tau)_{\tau \leq t}, (\tilde{a}_\tau)_{\tau \leq t-1}, (q_\tau)_{\tau \leq t-1}, (a_\tau)_{\tau \leq t-1}) = (\theta_{\leq t}, \tilde{a}_{\leq t-1}, q_{\leq t-1}, a_{\leq t-1}).$$

This is also Principal's private history. Agent's private history (which is also the public history) is denoted by:

$$h_t^A = (\tilde{a}_{\leq t-1}, q_{\leq t-1}, a_{\leq t-1}).$$

For each $\tau \leq t$, we denote $h_\tau \preceq h_t$ ($h_\tau^A \preceq h_t^A$) if h_τ (h_τ^A) corresponds to the first τ-period sub-history of h_t (h_t^A).

2.2 Mechanism

We assume that Principal ex ante commits to any mechanism, denoted by $\alpha = (\alpha_t(h_t))_{t,h_t}$, where $\alpha_t(h_t) \in [0,1]$ denotes the probability that Principal recommends action 1 to Agent (i.e., $\tilde{a}_t = 1$) at time t given history h_t up until then. Regarding Principal's own action, it is without loss to assume that Principal always takes the same action as what she recommends to Agent (i.e., $q_t = \tilde{a}_t$), and hence we omit it for brevity.

By the revelation principle, the mechanism must satisfy the following *obedience* constraint: for each t and the history of recommended actions which has a positive

probability given α, Agent finds it optimal to obey the recommended action $\tilde{a}_t$:[15]

$$U_t(h_t^A, \tilde{a}_t; \alpha) \equiv \mathbb{P}(h_t^A, \tilde{a}_t|\alpha) \cdot \mathbb{E}\left[\sum_{s=t}^{\infty} \delta^t \alpha_s(h_s) u(\theta_s) \Big| h_t^A, \tilde{a}_t; \alpha\right] \geq 0,$$

where 0 on the right-hand side is because Principal can always revert to "perpetual disagreement": $q_\tau \equiv 0$ for $\tau \geq t$.[16]

Because Agent's disobedience implies this trivial outcome of perpetual disagreement in the continuation game, from here on, we focus only on the full and public histories without past disobedience. By abuse of notation, we denote them by $h_t = (\theta_{\leq t}, a_{\leq t-1})$ and $h_t^A = a_{\leq t-1}$ (with the interpretation that they correspond to the full and public histories such that $\tilde{a}_{(.)} = a_{(.)}$).

3 (In)efficiency of optimal dynamic mechanism

In order to highlight the role of dynamics and partial persistence, it is useful to introduce the notion of *efficient mechanisms*, which is based on the idea of Pareto-efficient action choices in a static *ex post* sense.

Definition 1. Mechanism α is an **efficient mechanism** if $\alpha_t(h_t)$ is Pareto efficient for each t, h_t, in the sense that: $\alpha_t = 1$ if $\theta_t \in \Theta_A$, and $\alpha_t = 0$ if $\theta_t \notin \Theta_P$. It is **ex ante efficient** if for some $\lambda \geq 0$, α maximizes $\mathbb{E}[\sum_{t=0}^{\infty} \delta^t (v(\theta_t) + \lambda u(\theta_t)) \alpha_t(h_t)]$.

An ex ante efficient mechanism is a special case of an efficient mechanism, while the two notions coincide if there is a single state (in $\Theta_P \setminus \Theta_A$) where interests are not aligned. This is the case in the examples and applications along the paper. Notice also that any efficient mechanism is at least partially informative.

The two extreme instances of an efficient mechanism are (i) the P-optimal mechanism where $\alpha_t(h_t) = 1_{\{\theta_t \in \Theta_P\}}$; and (ii) the A-optimal mechanism where

[15] Agent's action given any off-path event does not matter for Principal's ex ante payoff given α. By convention, we let $U_t(h_t^A, \tilde{a}_t; \alpha) \equiv 0$ if $\mathbb{P}(h_t^A, \tilde{a}_t|\alpha) = 0$.

[16] Conversely, Agent's continuation payoff can never be negative, because he can always guarantee 0 by playing $a_\tau \equiv 0$ for $\tau \geq t$.

$\alpha_t(h_t) = 1_{\{\theta_t \in \Theta_A\}}$. The A-optimal mechanism is always obedient, while the P-optimal mechanism may not be obedient. In fact, if the P-optimal mechanism is obedient, it is clearly optimal.

3.1 Benchmark cases

An ex ante efficient mechanism is optimal either in the static environment or with extreme (i.e., no or full) persistence. In this sense, ex ante efficient mechanisms can serve as a benchmark in studying dynamics with partial persistence.

Proposition 1. The optimal mechanism is ex ante efficient in the following cases:

(a) $\delta = 0$

(b) state is iid ($p_{i \to j} = p_{i' \to j}$ for all i, i' and each j)

(c) state is fully persistent ($p_{i \to j} > 0$ if and only if $i = j$).

The logic of Proposition 1 is the following. Ex ante efficiency of optimal mechanisms would be straightforward in a relaxed problem where Agent's obedience constraints are replaced with a single ex ante participation constraint, because to maximize $\mathbb{E}[\sum_{t=0}^{\infty} \delta^t v(\theta_t)\alpha_t(h_t)]$ subject to $\mathbb{E}[\sum_{t=0}^{\infty} \delta^t u(\theta_t)\alpha_t(h_t)] \geq 0$, we can set up a Lagrangian $\mathbb{E}[\sum_{t=0}^{\infty} \delta^t (v(\theta_t) + \lambda u(\theta_t))\alpha_t(h_t)]$. If a solution to this relaxed problem satisfies all the obedience constraints of the original problem, it is an optimal mechanism (i.e., the solution of the original problem) and thus optimal mechanisms are ex ante efficient.[17]

3.2 Suboptimality of efficient mechanisms

In our problem with dynamics and *partially* persistent state, efficient mechanisms could perform poorly.

Example 1. Let $\Theta = \{0, 1, 2\}$, and $P = (p_{i \to j})_{i,j}$ be such that:

[17] Ex ante efficiency of optimal mechanisms in a static environment with binary action is also a corollary of the main result of Ichihashi (2019).

	$\theta_{t+1}=2$	$\theta_{t+1}=1$	$\theta_{t+1}=0$
$\theta_t=2$	1	0	0
$\theta_t=1$	0	$\frac{1}{2}$	$\frac{1}{2}$
$\theta_t=0$	0	$\frac{1}{2}$	$\frac{1}{2}$

Let the prior μ satisfy $\mu(2)=\frac{1}{2}$ and $\mu(1)=\mu(0)=\frac{1}{4}$, which is one of the steady-state distributions. It is as if nature flips a fair coin at the beginning of the game, determining whether the state is 2 forever, or it is iid between 0 and 1.

Principal's stage-game payoff is $v(0)<0<v(1)=v(2)$ (i.e., $\Theta_P=\{1,2\}$); and Agent's is $u(0)=u(1)<0<u(2)$ (i.e., $\Theta_A=\{2\}$). So that Agent prefers agreement to disagreement if no information is disclosed, assume $u(2)+u(1)\geq 0$.

Observe that any efficient mechanism recommends $a_t=1$ if $\theta_t=2$, and $a_t=0$ if $\theta_t=0$. Because state 2 and states 0/1 are separated from each other, once action 0 is recommended, Agent realizes that the state is never 2, and therefore, plays action 0 forever after. This means that, as time goes by, any obedient efficient mechanism in this example converges (in probability) to the A-optimal policy. □

3.3 Inefficient agreement

We propose two ways to potentially improve the mechanism, by introducing some *inefficiency*. The first candidate is the *no-disclosure* mechanism:[18]

Definition 2. α is the **no-disclosure** mechanism if $\alpha_t(h_t)=1$ for all h_t such that $h_t^A=(1,\ldots,1)$.

This mechanism always recommends action 1 as long as Agent has been obedient in the past. It is obedient because we are assuming $u(2)+u(1)\geq 0$. It is not an efficient mechanism, because Principal recommends action 1 even if $\theta\notin\Theta_P$. More precisely, it induces *inefficient agreement* in some states.

In the example above, the no-disclosure mechanism yields the expected payoff $\frac{\mathbb{E}[v(\theta)]}{1-\delta}$ to Principal. On the other hand, in any efficient mechanism, Principal's

[18] It is clear that there are many other mechanisms that do not disclose any information. We refer to the particular mechanism that always recommends agreement as the *no-disclosure* mechanism to stress that its goal is to avoid disclosing information.

expected continuation payoff converges to what the A-optimal mechanism induces: $\frac{\mathbb{E}[v(\theta)1_{\{\theta=2\}}]}{1-\delta}$. Therefore, if $\mathbb{E}[v(\theta)] > \mathbb{E}[v(\theta)1_{\{\theta=2\}}]$; equivalently, if $v(1)+v(0)>0$, then there exists $\delta^* \in (0,1)$ such that, for $\delta > \delta^*$, Principal strictly prefers the no-disclosure mechanism to any efficient mechanism.

The intuition for the improvement is simple. In any efficient mechanism, Agent realizes that the state is 1 or 0 forever as soon as action 0 is recommended, and the mechanism reduces to the A-optimal one since then. In the no-disclosure mechanism, such a moment never comes, as the recommendation is state-independent. Of course, its cost is inefficient agreement in state 0, but the benefit may outweigh the cost.

Obviously, this exercise does not itself mean that the no-disclosure mechanism is optimal. Nevertheless, we show later that it is optimal in some cases.

3.4 Inefficient disagreement

The second candidate for improvement is to introduce *inefficient disagreement*, that is, to recommend action 0 in some state where both Principal's and Agent's payoffs are positive (i.e., state 2 in the example).

In the same example, consider the following "full obfuscation" mechanism: At any history where Agent has been obedient, Principal recommends action 1 (i) with probability 1 if $\theta_1 = 1$, and (ii) with probability $\frac{1}{2}$ if $\theta_1 = 2$, while (iii) with probability 0 if $\theta_1 = 0$.[19] Agent's obedience condition is the same: $u(2)+u(1) \geq 0$.

The motivation for this inefficiency is again to prevent revealing that the state is 0 when action 0 is recommended. Indeed, Agent's belief for state 2 is kept constant regardless of the recommendation. However, as opposed to the no-disclosure mechanism which generates inefficient *agreement*, this full obfuscation mechanism generates inefficient *disagreement*, by recommending action 0 in state 2.

This mechanism yields a periodic payoff of $\frac{v(2)}{4}+\frac{v(1)}{4}$ in expectation. Therefore, if $\frac{v(2)}{4}+\frac{v(1)}{4} > \mathbb{E}[v(\theta)1_{\theta=2}] = \frac{v(2)}{2}$; equivalently, $v(2) < v(1)$, then there exists $\delta^* \in$

[19] We use the term "obfuscation" because the point of recommending action 0 if $\theta_1 = 2$ is to reduce the information content of recommending action 0 if $\theta_1 = 0$. With this mechanism, obfuscation is "full" because information (about future states) is fully eliminated.

$(0, 1)$ such that, for $\delta > \delta^*$, Principal strictly prefers the obfuscation mechanism to any efficient mechanism.

Again, the exercise does not show that this obfuscation mechanism is optimal. In fact, in the environment considered later, it is never optimal. However, we find that the optimal mechanism requires partial and temporary obfuscation (with inefficient disagreement) for some parameter regions, in particular, when the states are moderately persistent.

4 Lagrangian approach

The previous example shows that the optimal mechanism in our dynamic, partially-persistent environment can be inefficient in a variety of ways. The no-disclosure mechanism is inefficient as it yields agreement even if *both* players' agreement payoffs are negative ("*inefficient agreement*"). The obfuscation mechanism is inefficient as it implies disagreement even if *both* players' agreement payoffs are positive ("*inefficient disagreement*"). In order to better understand whether / when those different types of distortions are useful for Principal, it is important to characterize the optimal mechanism among all the obedient mechanisms.

Characterizing the optimal mechanism with partial persistence is a complex problem, simply because there are many candidates. For example, as typical in the static environment, the optimal mechanism may necessarily be a stochastic mechanism, in order to satisfy some obedience constraints with equality. Also, history dependence may be crucial because of partial persistence (recall Ball (2023)).

A well-known approach in the related literature is to be based on the value function recursion proposed by Abreu et al. (1990) (see Carrasco et al., 2019).[20] In the current context, it would concern a value function that has three kinds of arguments: Today's state realization, Agent's belief about it, and Agent's promised continuation payoff. Thus, the space of value functions is a complex infinite-dimensional space, and solving for it is challenging.

[20]Formally, this method is for the characterization of public-perfect equilibria. In the current context, Agent's private history coincides with the public history, which implies that the set of public-perfect-equilibrium payoffs coincides with the entire equilibrium payoffs set.

Following some papers in the literature (such as Ball (2023)), our approach is based on Lagrangian (weak) duality. As we see in the course, an advantage of this approach is that the problem of verifying the optimality of a candidate mechanism could be much simpler than the above standard approach, especially if the candidate mechanism is simple. Moreover, if one mechanism is proven optimal in some parameter region but not in a neighboring region, the violated optimality condition often suggests *how the mechanism should be modified* to obtain the correct optimal mechanism in that neighboring region. This allows us to virtually fully characterize how the optimal mechanism depends on parameter values in the application studied in Section 5.

In the rest of this section, we outline how we set up the Lagrangian, and show the weak duality lemma as the key to verify the optimality of candidate mechanisms. The Lagrangian method seems standard in economics, so we only provide a minimal explanation, and those who are familiar with it can skip the rest of this section.

Principal's (primal) problem is as follows:

$$\begin{aligned} \max_{\alpha} \quad & V(\alpha) \equiv \mathbb{E}\left[\sum_{t=0}^{\infty} \delta^t v(\theta_t)\alpha_t(h_t)\right] \\ \text{sub. to} \quad & U_\tau(h_\tau^A, a_\tau; \alpha) \geq 0, \ \forall \tau, h_\tau^A, a_\tau. \end{aligned}$$

Recall our convention that $U_\tau(h_\tau^A, a_\tau; \alpha) \equiv 0$ if the probability of having (h_τ^A, a_τ) is 0 under α.

Consider its Lagrangian:

$$\begin{aligned} L\Big(\alpha, \lambda\Big) &= V(\alpha) + \sum_{\tau, h_\tau^A, a_\tau} \lambda_\tau(h_\tau^A, a_\tau) U_\tau(h_\tau^A, a_\tau; \alpha) \\ &= \sum_{t\geq 1} \sum_{h_t, a_t} \left[v_t(\theta_t) + u_t(\theta_t) \sum_{\tau \leq t; (h_\tau^A, a_\tau) \preceq (h_t^A, a_t)} \lambda_\tau(h_\tau^A, a_\tau) \right] a_t \mathbb{P}(h_t, a_t; \alpha), \end{aligned}$$

for some multiplier function $\lambda = (\lambda_\tau(h_\tau^A, a_\tau))_{\tau, h_\tau^A, a_\tau} \geq 0$, where $(h_\tau^A, a_\tau) \preceq (h_t^A, a_t)$ means that (h_τ^A, a_τ) is the first τ-subsequence of (h_t^A, a_t). It is useful to introduce

the *cumulative* multiplier:

$$\Lambda_t(h_t^A, a_t) = \sum_{\tau \leq t; (h_\tau^A, a_\tau) \preceq (h_t^A, a_t)} \lambda_\tau(h_\tau^A, a_\tau).$$

That λ is non-negative implies that Λ is non-decreasing: for each $t \leq t'$ and $(h_t^A, a_t) \preceq (h_{t'}^A, a_{t'})$, we have $\Lambda_t(h_t^A, a_t) \leq \Lambda_t(h_{t'}^A, a_{t'})$. They are related one-to-one (analogous to the one-to-one relationship between a pmf and a cdf).

With the cumulative multiplier function Λ, we have:

$$L\Big(\alpha, \lambda\Big) = \sum_{t \geq 1} \sum_{h_t, a_t} \left[v_t(\theta_t) + u_t(\theta_t)\Lambda_t(h_t^A, a_t)\right] a_t \mathbb{P}(h_t, a_t; \alpha).$$

The following lemma is the basis for our "guess-and-verify" procedure.

Lemma 1. (Weak duality) If $\exists(\alpha^*, \lambda^*)$ s.t. (i) α^* is obedient and $\lambda^* \geq 0$, (ii) α^* is optimal for $L(\cdot, \lambda^*)$, and (iii) $L(\alpha^*, \lambda^*) = V(\alpha^*)$, then α^* is optimal.

Proof. Given any obedient mechanism α, we have:

$$V(\alpha) \leq L(\alpha, \lambda^*) \leq L(\alpha^*, \lambda^*) = V(\alpha^*),$$

where the first inequality is by (i), the second by (ii), and the equality by (iii). □

The next key observation is that, once the above Lagrangian is set up, its maximization is a single-person problem (by a "fictitious Lagrangian-maximizer"), and therefore, the solution is time-consistent (*Bellman's principle*). Also, it admits a recursive formulation and the one-shot deviation principle in the current context.[21]

5 Application: Hospital-Patient Relationship

To illustrate the above Lagrangian-based approach, we consider a simple information design environment motivated by the following hospital-patient relationship.

[21] See Marcet and Marimon (2019) for establishing strong duality in a relevant context and the recursive formula for a saddle-point value function.

A hospital (Principal) wants to persuade a patient (Agent) to adhere to a health care plan, which comprises periodic medication (e.g., palliative medication for cancer, or preventive medication for mental depression). Let $\theta_t = (\theta_t^c, \theta_t^p) \in \Theta = \{0,1\}^2$, where $\theta_t^c \in \{0,1\}$ denotes the state of the efficacy of the medication, and $\theta_t^p \in \{0,1\}$ denotes the state of its cost (which Agent does not care about but Principal does). For example: $\theta_t^c = 1$ may mean that the disease is so strong that the medication's benefit is high relative to its side effects, while $\theta_t^c = 0$ means the opposite; and $\theta_t^p = 1$ may mean that medication (free for Agent) is cheaper than with $\theta_t^p = 0$.[22]

Only Principal can observe efficacy and cost of the medication, which can change periodically (e.g., scientific evidences, drug markets,...). Depending on Principal's recommendation in each period, Agent might update his belief about the future payoff of adhering to the medication. Therefore, it is important for Principal to design how to communicate the information about the state to Agent.

Principal's agreement payoff at t (i.e., Principal offers certain medication and Agent adheres to it) is $v(\theta_t^c, \theta_t^p)$, and Agent's is $u(\theta_t^c)$; while their non-agreement payoffs (i.e., either Principal does not offer any medication or Agent does not adhere) are 0. Assume that (i) v, u are increasing in their arguments; (ii) $\Theta_P = \{01, 10, 11\}$ and $\Theta_A = \{10, 11\}$. Thus, if $\theta^c = 1$, then both Principal and Agent prefer action 1. If $\theta^c = 0$, then Agent prefers action 0, while Principal's preference depends on his private state: Principal prefers action 0 if and only if $\theta^c = \theta^p = 0$.

In order to highlight the effect of the state persistence that is payoff-relevant to Agent, we assume that θ_t^p is iid over time with $\mathbb{P}(\theta^p = 0) = \mathbb{P}(\theta^p = 1) = \frac{1}{2}$, independently from θ_t^c; while $\mathbb{P}(\theta_{t+1}^c \neq \theta_t^c | \theta_t^c) = \beta \leq \frac{1}{2}$ for each θ_t^c. Hence, $1 - \beta$ represents the persistence of θ_t^c. In the medication context, the assumption may be interpreted as the extreme case where the efficacy changes much less frequently than the cost. The prior for θ_0 is the steady-state distribution $\mu(00) = \mu(01) = \mu(10) = \mu(11) = \frac{1}{4}$. Henceforth, we also denote $v(\theta)$ and $u(\theta^c)$ by v_θ and u_{θ^c}.

We normalize $v_{00} = -1$ and $u_0 = -1$,[23] and we assume:

[22] Alternatively, $\theta_t^p = 0$ may mean that hospitalization (again free for Agent) is costlier (either in the monetary or non-monetary sense) than with $\theta_t^p = 1$.

[23] To revert this normalization, replace v_θ and u_1 with $\frac{v_\theta}{|v_{00}|}$ and $\frac{u_1}{|u_0|}$ everywhere below.

Assumption 1. $u_1 > 1$ and $v_{01} > 1$.

That $u_1 > 1$ corresponds to $v(1) + v(2) \geq 0$ in the Example in Section 3.2: In the no-disclosure mechanism, Agent plays action 1. That $v_{01} > 1$ is to avoid the less interesting case where the no-disclosure mechanism is worse for Principal than the A-optimal mechanism (and hence is never optimal).[24]

5.1 First best

The following condition fully characterizes the possibility of the first-best (i.e., P-optimal) mechanism. Agent is most reluctant to play action 1 if the mechanism recommended action 0 (thereby revealing state 00) in the previous period, obeying if and only if his continuation payoff is nonnegative.

Proposition 2. If $\mathbb{E}[\sum_{t=1}^{\infty} \delta^{t-1} u(\theta_t) 1_{\{\theta_t \neq 00\}} | \theta_0 = 00] \geq 0$, then the P-optimal mechanism is obedient, and thus optimal.

From now on, we assume the converse: Agent would be too pessimistic to play action 1 at t if he knew that the common state was 0 at $t-1$. Thus, the P-optimal mechanism is not obedient.

Assumption 2. $\mathbb{E}[\sum_{t=1}^{\infty} \delta^{t-1} u(\theta_t) 1_{\{\theta_t \neq 00\}} | \theta_0 = 00] < 0$; equivalently, $\delta < \frac{1-2u_1\beta-\beta}{1-2\beta}$.

5.2 Optimal Mechanism and Comparative Statics with respect to patience and persistence

In what follows, we characterize the optimal mechanism and discuss its properties as the parameters of the patience and persistence change. We also illustrate how our Lagrangian approach could be useful in identifying the optimal mechanisms "piece by piece" for a variety of parameter regions.

[24] Both the no-disclosure mechanism and the A-optimal mechanism are obedient, and hence, they can be simply compared based on Principal's payoff. In the no-disclosure mechanism, Principal's average expected payoff is $\mathbb{E}[v(\theta)]$; while in the A-optimal mechanism, it is $\mathbb{E}[v(\theta) 1_{\{\theta_c = 1\}}]$. Therefore, the former is larger if and only if $\mathbb{E}[v(\theta) 1_{\{\theta_c = 0\}}] = \frac{v_{01}-1}{2} \geq 0$.

5.2.1 Low δ

First, consider very impatient players (i.e., very low δ). In this case, future gains of distorting today's information would be too small to be profitable. As a consequence, an efficient mechanism is to be optimal. The next theorem provides a sufficient condition under which efficient mechanisms are optimal, clearly satisfied when δ is sufficiently low.

Theorem 1. Under Assumptions 1-2, an efficient mechanism is optimal if:

$$\begin{cases} \frac{1}{v_{01}} \geq \frac{1-\beta-\delta+2\beta\delta-2\beta u_1}{2-3\delta+\delta^2+\beta\delta(3-2\delta)}\delta \\ \frac{v_{10}}{v_{01}} \geq \frac{u_1(2-\delta-2\beta+2\beta\delta)+\beta}{2-3\delta+\delta^2+\beta\delta(3-2\delta)}\delta. \end{cases}$$

To explain this result, recall first that any efficient mechanism recommends action 1 if $\theta_t = 11$ or 10, and action 0 if $\theta_t = 00$. Thus, the only degree of freedom is in state $\theta_t = 01$. It is reasonable to guess that Agent's obedience constraint is binding at each $t+1$ conditional on being recommended action 0 at t, which pins down the probability of recommending action 1 at each t when $\theta_t = 01$.

Let us first observe that such a mechanism indeed exists. Fix $q \in (0,1)$, and consider the following mechanism. Conditional on $\theta_t = 01$:

- Phase P: If $t = 1$ or $h_t^A = (1, \ldots, 1)$, recommend action 1;
- Phase E: Otherwise, recommend action 1 (0) with probability q $(1-q)$.

The mechanism begins with Phase P, where the outcome is the same as in the P-optimal mechanism. This Phase P continues until action 0 is first recommended (because $\theta_t = 00$), at which point the mechanism moves to Phase E.

Lemma 2. Under Assumptions 1-2, there exists $q \in (0,1)$ such that Agent's obedience constraint for action 1 is binding at each t conditional on being recommended action 0 at $t-1$ (and all the other obedience constraints are satisfied).

This result exploits the fact that Agent's continuation payoff in Phase E only depends on the (unobserved) realization of the common state, θ_t^c. Agent is most pessimistic (and thus reluctant to obey $a_t = 1$) if $a_{t-1} = 0$, which reveals that

$\theta^c_{t-1} = 0$ and thus $\mathbb{P}(\theta^c_t = 1 | a_{t-1} = 0, a_t = 1) = \frac{\beta}{(1-\beta)\frac{q}{2}+\beta}$, independently of the previous history. Since Agent's continuation payoff from t onwards, conditionally upon $(a_{t-1}, a_t) = (0, 1)$, is strictly positive if $q = 0$ (A-optimal policy) and strictly negative if $q = 1$ (P-optimal policy), there exists $q \in (0, 1)$ for which it is zero and thus the corresponding obedience constraints hold with equality.

Now we provide the idea of the proof of Theorem 1 using the Lagrangian approach. Notice that Agent's obedience condition for action 1 is slack in Phase P, which suggests $\Lambda^P \equiv \Lambda_t(1, \ldots, 1) = 0$. In Phase E, the obedience condition is binding, and moreover, the probability of recommending action 1 is interior. This suggests that the (fictitious) Lagrangian maximizer must be indifferent between action 1 and action 0 at any t with $\theta_t = 01$, implying $\Lambda^E \equiv v_{01}$.[25]

Once the (cumulative) multipliers are so constructed, the optimality conditions of the mechanism imply the corresponding parameter region, as follows. Because the problem of maximizing the Lagrangian is a single-person problem, it can be solved by applying a Bellman principle. More specifically, consider first any period in Phase E (by backward induction). The one-shot deviation principle implies that, in this phase: action 1 must be taken if $\theta = 10, 11$ because $v_\theta + v_{01} u_{\theta^c} > 0$; action 0 must be taken if $\theta = 00$ because $v_\theta + v_{01} u_{\theta^c} < 0$; and any action is optimal if $\theta = 01$. Thus, the above efficient mechanism satisfies all the requirements.

In Phase P, the one-shot deviation principle requires: $v_\theta + \delta L^P(\theta^c) \geq \delta L^E(\theta^c),\ \forall \theta$, where $L^j(\theta^c)$ is the continuation value of the Lagrangian as a function of next period's phase $j \in \{P, E\}$ and current period's common state θ^c. At $\theta = 01, 11$, they are always satisfied, while at $\theta = 00, 10$:

$$\begin{aligned} v_{00} + \delta L^P(0) \leq \delta L^E(0) \quad &\Leftrightarrow \quad \tfrac{1}{v_{01}} \geq \tfrac{1-\beta-\delta+2\beta\delta-2\beta u_1}{2-3\delta+\delta^2+\beta\delta(3-2\delta)}\delta, \\ v_{10} + \delta L^P(1) \geq \delta L^E(1) \quad &\Leftrightarrow \quad \tfrac{v_{10}}{v_{01}} \geq \tfrac{u_1(2-\delta-2\beta+2\beta\delta)+\beta}{2-3\delta+\delta^2+\beta\delta(3-2\delta)}\delta; \end{aligned}$$

hence, Theorem 1 follows.

As its corollaries, we present comparative statics results with respect to pa-

[25] Thus, in terms of the (de-cumulative version of) Lagrange multiplier λ, it begins with zero until action 0 is ever recommended, and (only) right after the period where action 0 was first recommended, it takes the value of v_{01} (and 0 forever after).

tience and persistence parameters, when δ is sufficiently low:

Corollary 1. Under Assumptions 1-2:

1. If $\delta \leq \min\left\{\frac{2}{v_{01}+1}, \frac{v_{10}}{u_1 v_{01}+v_{10}}\right\}$, then the above efficient mechanism is optimal for any β;

2. If $\frac{2}{v_{01}+1} < \delta \leq \frac{v_{10}}{u_1 v_{01}+v_{10}}$, then for each δ in this region, there exists $\beta(\delta) > 0$ such that the above efficient mechanism is optimal if $\beta \geq \beta(\delta)$.

Therefore, the efficient mechanism is optimal either when δ is very low (in which case β does not matter), or when δ is moderately low but β is higher than the threshold $\beta(\delta)$. The intuition for the former case is straightforward: With a very low δ, the environment becomes close to the static case, where the potential future gain of the present inefficient action is small. The intuition for the latter case is as explained in Section 3: In a more persistent environment, Principal is more reluctant to share the pessimistic information, because revealing the information today would have a larger impact on Agent's belief about the future states. Of course, hiding $\theta_t = 00$ reduces today's payoff ("*inefficient agreement*"), but the informational benefit of the inefficient agreement outweighs its instantaneous cost.

Notice that $\beta(\delta)$ is increasing in δ. Intuitively, this is because the motif of hiding today's information is more relevant when the players are more patient.

In case $\delta \in \left(\frac{2}{v_{01}+1}, \frac{v_{10}}{u_1 v_{01}+v_{10}}\right)$ and $\beta < \beta(\delta)$, the above efficient mechanism is no longer optimal. A natural next question is then what the optimal mechanism is in that case. Recall that, in that case, one of the optimality conditions for the efficient mechanism, that $v_{00} + \delta L^P(0) \leq \delta L^E(0)$, is violated. That is, given the above structure of the Lagrange multipliers, the (fictitious) Lagrangian maximizer prefers action 1 to action 0 at state $\theta = 00$ in Phase P. Therefore, he prefers action 1 *in all the states* in Phase P.

This argument suggests that the no-disclosure mechanism seems optimal in this region. Indeed, consider the following construction:

- Phase I ($\Lambda = 0$): Take $a = 1$ for all θ; once $a = 0$ is taken, move to Phase E.
- Phase E ($\Lambda = \Lambda^E$): Same as in Phase E of the above efficient mechanism.

Note that, in this construction, Phase E is "off-path", but this phase is important in order to prevent the Lagrangian maximizer from deviating to action 0 in Phase I. As for the efficient mechanism, the critical optimality conditions for the no-disclosure mechanism are at states $\theta = 10, 00$ in Phase I:

$$\begin{aligned} v_{00} + \delta L^I(0) \leq \delta L^E(0) \quad &\Leftrightarrow \quad \tfrac{1}{v_{01}} \leq \tfrac{1-\beta-\delta+2\beta\delta-2\beta u_1}{2-3\delta+\delta^2+\beta\delta(3-2\delta)}\delta, \\ v_{10} + \delta L^I(1) \geq \delta L^E(1) \quad &\Leftrightarrow \quad v_{10} \geq \tfrac{\beta-\beta v_{01}+2u_1 v_{01}(1-\beta-\delta+2\beta\delta)}{2(1-\delta)[1-\delta(1-2\beta)]}\delta; \end{aligned}$$

where $L^j(\theta^c)$ is the continuation value of the Lagrangian as a function of next period's phase $j \in \{P, E\}$ and current period's common state θ^c.

Theorem 2. The no-disclosure mechanism is optimal if

$$\begin{cases} \frac{1}{v_{01}} \leq \frac{1-\beta-\delta+2\beta\delta-2\beta u_1}{2-3\delta+\delta^2+\beta\delta(3-2\delta)}\delta \\ v_{10} \geq \frac{\beta-\beta v_{01}+2u_1 v_{01}(1-\beta-\delta+2\beta\delta)}{2(1-\delta)[1-\delta(1-2\beta)]}\delta. \end{cases}$$

These conditions are satisfied if $\frac{2}{v_{01}+1} < \delta \leq \frac{v_{10}}{u_1 v_{01}+v_{10}}$ and $\beta \leq \beta(\delta)$.

Figure 1 illustrates the parameter regions defined in Theorems 1 and 2.

5.2.2 Intermediate δ

So far, the optimal mechanisms are identified for low δ. Let us now explain what happens for δ (slightly) above the threshold imposed in Theorems 1 and 2. Both in the no-disclosure mechanism and the efficient mechanism, setting such δ corresponds to the violation of the Lagrangian maximizer's optimality condition in state $\theta = 10$ in the initial phase (i.e., Phase I for the no-disclosure mechanism and Phase P for the efficient mechanism). More specifically, if we employ the above structure of the Lagrange multipliers, then the Lagrangian maximizer would prefer action 0 to action 1 in state $\theta = 10$. Despite the loss of the present payoff v_{10}, such a deviation would imply the future gain from *the slackness of the obedience constraint* (recall that $\Lambda = \Lambda^E > 0$ in Phase E, while $\Lambda = 0$ in Phase I / Phase P). Higher δ implies that this future gain outweighs the present loss.

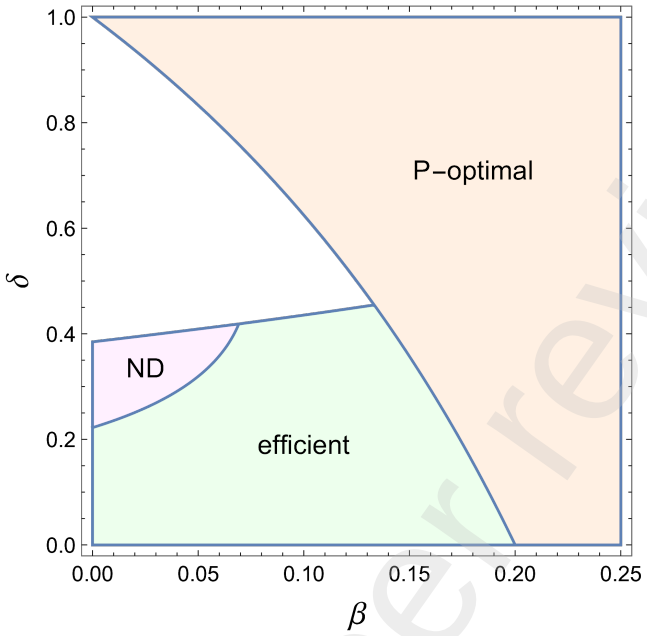

Figure 1: Optimality of *best efficient* policy (green), *no disclosure* policy (pink), and P-optimal policy (orange), with $v_{01} = 8$, $v_{10} = 10$, and $u_1 = 2$.

A natural solution to this motif of "enjoying the slackness" is to *slow down* the evolution of the cumulative Lagrange multiplier. More specifically, consider a three-phase structure of the Lagrange multipliers: $\Lambda = 0$ in the initial phase; $\Lambda \in (0, \Lambda^E)$ in the next phase; $\Lambda = \Lambda^E$ in the last phase (where the transition from one phase to the subsequent one occurs as soon as action 0 is taken).[26]

Let us examine two mechanisms with this three-phase structure of the Lagrange multipliers that build, respectively, on the no-disclosure mechanism and on the best efficient mechanism defined before. The following (three-phase) no-disclosure mechanism will be shown to be sometimes optimal for small β:

- Phase I ($\Lambda = 0$): Take $a = 1$ for all θ; once $a = 0$, then move to Phase P.

[26] Another possibility is to maintain the two-phase structure, but with a stochastic transition: If action 0 is recommended, transition to Phase E occurs only with some probability. Under some conditions, this stochastic structure allows us to establish optimality of no-disclosure if $\beta = 0$ (for any δ above the threshold required in Theorems 1 and 2). See Appendix C.1 and the gray region in Figure 3.

- Phase P ($\Lambda = \Lambda^P \in (0, v_{01})$):[27] Take $a = 1$ for all $\theta \neq 00$ (and $a = 0$ if $\theta = 00$); once $a = 0$, then move to Phase E.

- Phase E ($\Lambda = v_{01}$): Same as Phase E of the above mechanisms.

This is the no-disclosure mechanism in that, on-path, the action recommendation is always 1. Phase P and E are used only off-path, in order to avoid the (fictitious) Lagrangian maximizer's deviation from Phase I. As discussed above, the key to this construction is the delayed evolution of the cumulative multiplier.

For intermediate β, the following *(one-)obfuscation mechanism* will be shown to be sometimes optimal.[28]

- Phase O ($\Lambda^O = 0$): $a = 0$ if $\theta = 00$; $a = 1$ if $\theta \in \{01, 11\}$; mix between $a = 0$ and $a = 1$ if $\theta = 10$; if $a = 0$, move to next phase;

- Phase P ($\Lambda^P \in (0, v_{01}]$): P-optimal policy; if $a = 0$, move to next phase;

- Phase E ($\Lambda^E = v_{01}$): Same as Phase E of the above mechanisms.

The logic that underlies the obfuscation policy is that recommending action 0 in state 10 (in addition to state 00) reduces the informational content of $a_t = 0$ (Agent no longer learns that $\theta_t^c = 0$), and thereby it becomes easier to convince Agent to obey $a = 1$ in the future. Of course, this policy is inefficient because both Principal and Agent prefer $a = 1$ in state 10, which is the type of inefficiency ("*inefficient disagreement*") opposite from that in the no-disclosure mechanism.

To avoid lengthy analytical expressions in establishing optimality of the above three-phase mechanisms, let us make the following assumption.

Assumption 3. $\delta = \frac{1}{2}$ and $u_1 = 2$.

Let us also assume that v_{01}, v_{10} are sufficiently large and not too different.[29]

[27] See the appendix for the expression

[28] The threshold β^* that separates the optimality regions of obfuscation and no-disclosure (in Theorem 3) originates in the no-deviation constraint in the initial phase when state 00 occurs.

[29] The upper bound $\frac{v_{10}}{v_{01}} \leq \frac{3}{2}$ guarantees that the value of Λ^P that is consistent with randomization in Phase O is not greater than v_{10}, and that the no-deviation constraint if $\theta_t = 01$ in Phase P holds. The lower bound $\frac{v_{10}}{v_{01}} \geq \frac{2}{3}$ guarantees that the no-deviation constraint if $\theta_t = 10$ in Phase P holds. The condition $v_{01} + v_{10} > 9$ is not necessary for optimality of the obfuscation mechanism. On the contrary, it makes the obfuscation mechanism suboptimal for small β.

Assumption 4. $v_{01} + v_{10} > 9$ and $\frac{v_{10}}{v_{01}} \in \left[\frac{2}{3}, \frac{3}{2}\right]$.

Under Assumption 3, the P-optimal mechanism is obedient if and only if $\beta \geq \frac{1}{8}$. The following result characterizes the optimal mechanism for $\beta < \frac{1}{8}$.

Theorem 3. Under Assumptions 3-4, there exists $\beta^* \in (0, \frac{1}{8})$ such that the no-disclosure mechanism is optimal if $\beta \in [0, \beta^*]$, and the obfuscation mechanism is optimal if $\beta \in \left[\beta^*, \frac{1}{8}\right)$.

Figure 2 represents the parameter regions for which optimality of no-disclosure and obfuscation mechanisms is established using three-phase multipliers.

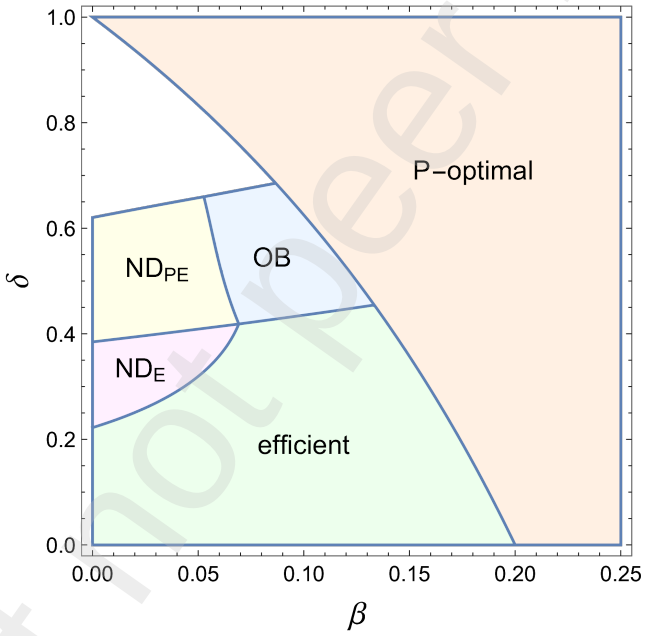

Figure 2: Optimality of *no disclosure* policy (yellow) and *obfuscation* policy (blue) established using "three-phase" multipliers, with $v_{01} = 8$, $v_{10} = 10$, and $u_1 = 2$.

5.2.3 Higher δ

As explained above, as δ becomes higher, we need even slower evolution of Lagrange multipliers, and this in turns suggests the shape of the corresponding optimal mechanism. We do not explain the detail (see Appendix C), as the intuition

is basically the same, but this procedure allows us to systematically identify the optimal mechanism for a large space of parameters, as illustrated in Figure 3.

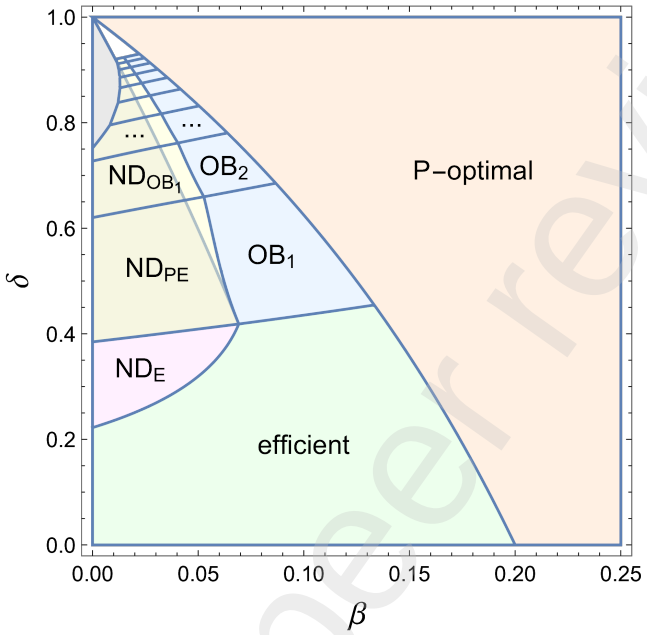

Figure 3: Optimality of *no disclosure* (pink, yellow, gray), and *obfuscation* policies with $n \in \{1, ..., 9\}$ obfuscation phases (blue), with $v_{01} = 8$, $v_{10} = 10$, and $u_1 = 2$.

Although the structure of the optimal mechanisms becomes more complex[30], the qualitative feature remains robust: With sufficiently high persistence (low β), the no-disclosure mechanism is optimal ("inefficient agreement"); with moderate persistence and patience, the obfuscation mechanisms are optimal ("inefficient disagreement"); in other regions, efficient mechanisms are optimal.

6 Conclusion

This paper studies dynamic information design in a repeated environment with (imperfectly) persistent state. The fully committed Principal may find optimal

[30] As δ increases, successive thresholds for δ arise, above which an additional obfuscation phase is inserted right before Phase P.

to deviate from an efficient mechanism, especially with a moderate degree of patience. The optimal mechanism may either exhibit *inefficient agreement* as in the no-disclosure mechanism, which tends to be optimal with high persistence; or *inefficient disagreement* as in obfuscation mechanisms, which tends to be optimal with moderate persistence. The methodology based on Lagrangian (weak) duality is illustrated as useful in not only verifying the optimality of any given guess, but also in educating how a candidate mechanism should be modified in the neighboring region once it became suboptimal.

We believe that the methodology here would be amenable to a variety of alternative modelling of dynamic design, such as to the case with monetary transfer, multiple agents, their private information, and so on. Obviously, the problem would be more complex in those alternative models. Nevertheless, the basic idea of constructing a Lagrangian stacking all the relevant constraints and systematically conducting a "piece-by-piece" guess-and-verify approach could be useful in gaining some key insights.

A Proofs for Section 3

A.1 Proof of Proposition 1 (a)

Proof. Because both Principal and Agent care only about the outcomes at $t = 0$, the problem reduces to the one maximizing $\mathbb{E}[v(\theta_0)]$ subject to $\mathbb{E}[u(\theta_0)] \geq 0$. In case the P-optimal mechanism is obedient (and hence optimal), it is ex ante efficient (with $\lambda = 0$). In case the P-optimal mechanism violates the obedience constraint, we can set up a Lagrangian:

$$L(\alpha, \lambda) = \mathbb{E}[(v(\theta_0) + \lambda u(\theta_0))\alpha_0(\theta_0)]$$

for $\lambda > 0$. Maximizing the Lagrangian with given λ, we obtain $\alpha_0(\theta_0) = 1$ if $v(\theta_0) + \lambda u(\theta_0) > 0$, $\alpha_0(\theta_0) = 0$ if $v(\theta_0) + \lambda u(\theta_0) < 0$, and any $\alpha_0(\theta_0) \in [0, 1]$ if $v(\theta_0) + \lambda u(\theta_0) = 0$. Clearly, it is ex ante efficient.

It is easy to see that, for some λ^* and some choice of $\alpha_0(\theta_0)$ for θ_0 with $v(\theta_0) + \lambda^* u(\theta_0) = 0$, Agent's expected payoff in this mechanism (call it α^*) can be made 0. To show that α^* is optimal, let $V(\alpha)$ denote Principal's ex ante expected payoff given any obedient mechanism α. We have:

$$V(\alpha) \leq L(\alpha, \lambda^*) \leq L(\alpha^*, \lambda^*) = V(\alpha^*),$$

implying $V^* \geq V(\alpha)$, that is, α^* is optimal. □

A.2 Proof of Proposition 1 (b)

Proof. As in the myopic case, consider the relaxed problem which only concerns the obedience constraint at $t = 0$. In case the P-optimal mechanism is obedient (and hence optimal), it is ex ante efficient (with $\lambda = 0$). In case the P-optimal mechanism violates the obedience constraint, we can set up a Lagrangian:

$$L(\alpha, \lambda) = \mathbb{E}[\sum_{t=0}^{\infty} \delta^t (v(\theta_t) + \lambda u(\theta_t))\alpha_t(h_t)]$$

for $\lambda > 0$. Maximizing the Lagrangian with given λ, for any t, we obtain $\alpha_t(\theta_t) = 1$ if $v(\theta_t) + \lambda u(\theta_t) > 0$, $\alpha_t(\theta_t) = 0$ if $v(\theta_t) + \lambda u(\theta_t) < 0$, and any $\alpha_t(\theta_t) \in [0, 1]$ if $v(\theta_t) + \lambda u(\theta_t) = 0$. Clearly, it is ex ante efficient. In what follows, we consider the mechanism such that, for some $\kappa \in [0, 1]$, $\alpha_t(\theta_t) = \kappa$ if $v(\theta_t) + \lambda u(\theta_t) = 0$.

Denote the ex ante expected payoffs for Principal and Agent by $\hat{V}(\lambda, \kappa)$ and $\hat{U}(\lambda, \kappa)$. If $\hat{U}(0, 1) \geq 0$, the mechanism that is optimal for Principal satisfies Agent's *ex ante* participation constraint and is thus the solution of the relaxed problem. If $\hat{U}(0, 1) < 0$, let λ^* be the maximum $\lambda > 0$ such that $\hat{U}(\lambda, 1) < 0$. Such a maximum exists because $\hat{U}(\cdot, 1)$ is an increasing left-continuous step function of λ. A jump occurs when $\exists \theta$ s.t. $v(\theta) + \lambda u(\theta) = 0$, and at λ^* there is one of finitely many jumps. While $\hat{U}(\lambda^*, 1) < 0$, we must have $\hat{U}(\lambda^*, 0) \geq 0$ because, for small $\epsilon > 0$, $\hat{U}(\lambda^*, 0) = \hat{U}(\lambda^* + \epsilon, 0) = \hat{U}(\lambda^* + \epsilon, 1)$, which is non-negative by the definition of λ^*. Because $\hat{U}(\lambda^*, \kappa)$ is continuous in κ, there exists κ^* such that $\hat{U}(\lambda^*, \kappa^*) = 0$, and thus (λ^*, κ^*) determines a solution of the relaxed problem.

We proceed by showing that the mechanism α constructed above satisfies all the obedience constraints of the original problem (not only the one at $t = 0$). Since the state is iid, expected payoff from α is the same in all periods. This means that Agent's expected payoff in any period is non-negative (it is zero unless $\hat{U}(0, 1) > 0$). Moreover, since payoff from $a = 0$ is zero, expected payoff conditional on $a = 1$ is also non-negative. The mechanism is thus obedient and hence optimal.

The proof concludes with the observation that any inefficient mechanism $\tilde{\alpha}$ must be either strictly worse *ex ante* for Principal or strictly worse *ex ante* for Agent: $V(\tilde{\alpha}) < V(\alpha)$ or $U(\tilde{\alpha}) < 0$. □

A.3 Proof of Proposition 1 (c)

Proof. The proof is similar to that of Proposition 1 (b). Consider the same relaxed problem, the same candidate solution of the relaxed problem as a function of (λ, κ), and the same $\lambda^* > 0$ such that $\hat{U}(\lambda^*, 1) < 0$ and $\hat{U}(\lambda^*, 0) \geq 0$ (such λ^* exists unless $\hat{U}(0, 1) \geq 0$, in which case the mechanism that is optimal for Principal is the solution of the relaxed problem).

The difference from the iid case is that we no longer consider the mechanism

defined by (λ^*, κ^*) such that $\hat{U}(\lambda^*, \kappa^*) = 0$. Instead, we select the solution of the relaxed problem that is the mixture of the mechanism defined by $(\lambda^*, 1)$ and that defined by $(\lambda^*, 0)$ with probabilities such that $U(\alpha) = 0$.[31] With this mechanism, if θ_t is such that $v(\theta_t) + \lambda^* u(\theta_t) = 0$, the recommendation is drawn at $t = 0$ and repeated in every period. (By contrast, in the iid case, the recommendation is independently drawn at each t whenever θ_t is such that $v(\theta_t) + \lambda^* u(\theta_t) = 0$.)

We proceed by showing that α satisfies all the obedience constraints of the original problem (not only the one at $t = 0$). Since the state is perfectly persistent, Agent's expected payoff conditionally on $a_t = 1$ is the same in every period and thus non-negative. The mechanism is thus obedient and hence optimal.

The proof concludes with the observation that, as explained in the proof of Proposition 1 (b), any ex ante inefficient mechanism $\tilde{\alpha}$ is strictly suboptimal. □

B Proofs for Section 5

B.1 Proof of Proposition 2

Proof. In the P-optimal mechanism, when recommended action 1, Agent's continuation payoff of obedience depends on his belief about the common state. The worst case is if Agent knows that the common state in the previous period was 0. The condition in the statement guarantees that, even in this case, Agent's continuation payoff is non-negative. Therefore, the P-optimal mechanism is obedient.

Regarding the second statement, denote by $U_{fb}(\theta^c)$, with $\theta^c \in \{0, 1\}$, the continuation payoff of Agent in the P-optimal mechanism from period t on, conditionally on θ^c being the common state at $t - 1$. We have:

$$\begin{cases} U_{fb}(0) = (1-\beta)\left(\frac{u_0}{2} + \delta U_{fb}(0)\right) + \beta\left(u_1 + \delta U_{fb}(1)\right) \\ U_{fb}(1) = \beta\left(\frac{u_0}{2} + \delta U_{fb}(0)\right) + (1-\beta)\left(u_1 + \delta U_{fb}(1)\right), \end{cases}$$

[31] The probability of selecting $(\lambda^*, 1)$ is κ^*, and the probability of selecting $(\lambda^*, 0)$ is $1 - \kappa^*$.

which yields (recall the normalization $u_0 = -1$):

$$\begin{cases} U_{fb}(0) = \frac{-(1-\delta-\beta+2\beta\delta)+2\beta u_1}{2(1-\delta)[1-\delta(1-2\beta)]} \\ U_{fb}(1) = \frac{2u_1(1-\delta-\beta+2\beta\delta)-\beta}{2(1-\delta)[1-\delta(1-2\beta)]}. \end{cases}$$

Disobedience gives zero payoff to Agent, and, after being recommended $a = 0$, Agent learns that the state is 00 and thus continuation utility is $U_{fb}(0)$. Therefore, Agent will be willing to obey the next time $a = 1$ is recommended if and only if:

$$U_{fb}(0) \geq 0 \;\Leftrightarrow\; \delta \geq \frac{1-2u_1\beta-\beta}{1-2\beta}.$$

Under this condition, the P-optimal mechanism is obedient. □

B.2 Proof of Lemma 2

Proof. Let us solve it "backward". First, consider the case where the mechanism is already in Phase E.

(Phase E) Let $U^E(\theta^c)$ denote Agent's continuation payoff from t on (discounted to t), conditional on $\theta^c_{t-1} = \theta^c$. Because the mechanism is stationary, $U^E(\theta^c)$ does not depend on t, and hence is well-defined. Note that θ^c_{t-1} is not directly observed by Agent. We have, recursively:

$$\begin{cases} U^E(0) = (1-\beta)\left[\frac{q}{2}u_0 + \delta U^E(0)\right] + \beta\left[u_1 + \delta U^E(1)\right] \\ U^E(1) = \beta\left[\frac{q}{2}u_0 + \delta U^E(0)\right] + (1-\beta)\left[u_1 + \delta U^E(1)\right] \end{cases}$$

$$\Leftrightarrow \begin{cases} U^E(0) = \frac{2u_1\beta+qu_0(1-\beta-\delta+2\beta\delta)}{2(1-\delta)[1-\delta(1-2\beta)]} \\ U^E(1) = \frac{qu_0\beta+2u_1(1-\beta-\delta+2\beta\delta)}{2(1-\delta)[1-\delta(1-2\beta)]}. \end{cases}$$

Note that $U^E(1) > U^E(0)$ if $\beta < \frac{1}{2}$ and $\delta < 1$.

Whenever $(a_{t-1}, a_t) = (0, 1)$, Agent knows that $\theta^c_{t-1} = 0$, and thus Agent's belief about θ^c_t is $\psi \equiv \mathbb{P}(\theta^c_t = 1 | a_{t-1} = 0, a_t = 1) = \frac{\beta}{(1-\beta)\frac{q}{2}+\beta}$. Hence, Agent's

expectation about own continuation payoff is:

$$\psi\left[u_1+\delta U^E(1)\right]+(1-\psi)\left[u_0+\delta U^E(0)\right],$$

which is independent of time and of the remaining past history.

In the extreme case where $q=1$ (P-optimal policy), we obtain:

$$\mathbb{E}\left[U^E|a_{t-1}=0,a_t=1\right]=\frac{[2-\delta(1-\beta)][u_0(1-\delta)+2u_1\beta-u_0\beta(1-2\delta)]}{2(1+\beta)(1-\delta)[1-\delta(1-2\beta)]},$$

which (with $u_0=-1$) is strictly negative if and only if:

$$\delta<\frac{1-2u_1\beta-\beta}{1-2\beta},$$

which is exactly Assumption 2.

In the other extreme case, where $q=0$ (A-optimal policy), it is trivial that $\mathbb{E}\left[U^E|a_{t-1}=0,a_t=1\right]$ is strictly positive.

Therefore, by Bolzano's intermediate value theorem, there exists $q^*\in(0,1)$ such that Agent's belief about own continuation payoff conditionally on $(a_{t-1},a_t)=(0,1)$ is zero (Agent is thus indifferent between obeying and disobeying $a_t=1$).

Finally, we need to show that (for the same q^*) Agent's expectation about own continuation payoff is greater than zero for any public history such that $(a_{t-1},a_t)=(1,1)$. It is given by:

$$\phi\left[u_1+\delta U^E(1)\right]+(1-\phi)\left[u_0+\delta U^E(0)\right],$$

where ϕ denotes Agent's belief for $\theta_t^c=1$ conditional on $a_{\leq t}$. Clearly, we have $\phi\geq\psi$, that is, Agent is more optimistic given $(a_{t-1},a_t)=(1,1)$ than given $(a_{t-1},a_t)=(0,1)$.[32] Hence, $\phi\left[u_1+\delta U^E(1)\right]+(1-\phi)\left[u_0+\delta U^E(0)\right]\geq 0$.

(Phase P) Next, consider Phase P, where the past action recommendations have been all 1 (including the initial period).

First, consider the case of $t\geq 1$. Let $U^P(\theta^c)$ denote Agent's continuation payoff

[32] Recall that $a_{t-1}=0$ necessarily means that $\theta_{t-1}^c=0$, while $a_{t-1}=1$ does not.

from t on (discounted to t), conditional on $\theta^c_{t-1} = \theta^c$. Because the mechanism is stationary, $U^P(\theta^c)$ does not depend on t, and hence is well-defined. Note that θ^c_{t-1} is not directly observed by Agent. Since with q^* constructed as above, the expected payoff in Phase E is zero, we have:

$$\begin{cases} U^P(0) = \frac{1-\beta}{2}\left[u_0 + \delta U^P(0)\right] + \beta\left[u_1 + \delta U^P(1)\right] \\ U^P(1) = \frac{\beta}{2}\left[u_0 + \delta U^P(0)\right] + (1-\beta)\left[u_1 + \delta U^P(1)\right] \end{cases}$$
$$\Leftrightarrow \begin{cases} U^P(0) = \frac{\beta+\delta-\beta^2(1-\delta)(2u_1+1)-2\beta\delta(u_1+1)+4\beta u_1-1}{2-\delta[3-\delta-\beta(3-2\delta-\beta+\beta\delta)]} \\ U^P(1) = \frac{u_1(2-\delta-2\beta+2\beta\delta)-\beta}{2-\delta[3-\delta-\beta(3-2\delta-\beta+\beta\delta)]}. \end{cases}$$

Let ϕ denote Agent's belief at t given the history of recommendation $a_{\leq t} = (1, \ldots, 1)$. Clearly, $\phi \geq \frac{1}{2}$. Therefore, Agent's obedience condition is satisfied if:

$$\frac{1}{2}U^P(0) + \frac{1}{2}U^P(1) \geq 0,$$

which is indeed the case.[33]

With $t = 0$, Agent's obedience condition is satisfied under the same condition, $\frac{1}{2}U^P(0) + \frac{1}{2}U^P(1) \geq 0$, and therefore, the proof completes. □

B.3 Proof of Theorem 1

Proof. The proof is based on the Lagrange weak duality (Lemma 1). We construct λ^*, or equivalently, its cumulative version Λ^*, with which an efficient mechanism (as defined in Section 5.2.1, with q^* satisfying the property in Lemma 2) is optimal. Let us guess that: $\Lambda^*_t(h^A_t, a_t) = 0$ if h^A_t does not contain 0; else $\Lambda^*_t(h^A_t, a_t) = v_{01}$.

Intuitively, $\Lambda^*_t(h^A_t, a_t) = 0$ if h^A_t does not contain 0 is because Agent's obedience condition is slack in this phase (Phase P). Otherwise, we have $\Lambda^*_t(h^A_t, a_t) = v_{01}$, because with this (and only with this) multiplier structure, maximizing Lagrangian (linear in the probabilities) can be consistent with choosing an interior q^*.

[33] To see this, take the lower bound corresponding to $u_1 = 1$ to find: $\frac{1}{2}U^P(0) + \frac{1}{2}U^P(1) = \frac{1+\beta(2-3\beta)(1-\delta)}{4-2\delta[3-\delta-\beta(3-2\delta-\beta+\beta\delta)]} > 0$.

With this multiplier structure, the Lagrangian becomes:

$$L(\alpha, \Lambda^*) = \sum_t \sum_{h_t, a_t} \delta^t [v(\theta_t) + u(\theta_t)\Lambda_t^*(h_t^A, a_t)] a_t \mathbb{P}(h_t, a_t; \alpha).$$

Let $\alpha^* = (\alpha_t^*(h_t))_{t,h_t}$ be its maximizer (given Λ^*), and let $L_t(h_t)$ denote its continuation value from period t on given h_t:

$$\begin{aligned} L_t(h_t) &= \sum_{\tau \geq t} \sum_{(h_\tau, a_\tau) \succeq h_t} \delta^{\tau - t} [v(\theta_\tau) + u(\theta_\tau)\Lambda_\tau^*(h_\tau^A, a_\tau)] a_\tau \mathbb{P}(h_\tau, a_\tau; \alpha) \\ &= \max_{a_t} \{ v(\theta_t) + u(\theta_t)\Lambda_t^*(h_t^A, a_t) + \delta \mathbb{E}_{\theta_{t+1}} L_{t+1}((h_t, a_t, \theta_{t+1}) \}, \end{aligned}$$

where the last step is based on the one-shot deviation principle.

Suppose that the efficient mechanism with q^* is optimal. Then, this one-shot deviation principle must be satisfied. Recall that this mechanism has two phases: Phase P and Phase E.

Let us first examine Phase E. Analogous to $U^E(\theta^c)$ above, let $L^E(\theta^c)$ denote the continuation value of the Lagrangian (under the efficient mechanism with q^*) from period t on in Phase E, conditional on $\theta_{t-1}^c = \theta^c$:

$$\begin{aligned} &\begin{cases} L^E(0) = (1-\beta)\delta L^E(0) + \beta \left(\frac{v_{10}+v_{11}}{2} + v_{01}u_1 + \delta L^E(1) \right) \\ L^E(1) = \beta\delta L^E(0) + (1-\beta) \left(\frac{v_{10}+v_{11}}{2} + v_{01}u_1 + \delta L^E(1) \right), \end{cases} \\ \Leftrightarrow &\begin{cases} L^E(0) = \frac{\beta(2u_1 v_{01} + v_{10} + v_{11})}{2(1-\delta)[1-\delta(1-2\beta)]} \\ L^E(1) = \frac{(1-\beta-\delta+2\beta\delta)(2u_1 v_{01} + v_{10} + v_{11})}{2(1-\delta)[1-\delta(1-2\beta)]}. \end{cases} \end{aligned} \tag{1}$$

Hence, for each h_t such that $t \geq 2$ and $h_t^A \neq (1, \ldots, 1)$:

$$L_t(h_t) = \begin{cases} v(\theta_t) + v_{01}u(\theta_t^c) + \delta L^E(1) & \text{if} \quad \theta_t^c = 1 \\ \delta L^E(0) & \text{if} \quad \theta_t^c = 0, \end{cases}$$

noting that θ_t is a part of h_t.

The one-shot deviation principle requires that:

$$v(\theta_t) + v_{01}u(\theta_t^c) + \delta L^E(1) \geq \delta L^E(1)$$

for $\theta_t = 11, 10$, which is satisfied as $v_{11}, v_{10}, v_{01}, u_1 > 0$; and

$$v(\theta_t) + v_{01}u(\theta_t^c) + L^E(0) \leq \delta L^E(0)$$

for $\theta_t = 00$, which is satisfied as $v_{00} = u_0 = -1$; and

$$v(\theta_t) + v_{01}u(\theta_t^c) + \delta L^E(0) = \delta L^E(0),$$

for $\theta_t = 01$. This equality holds by construction, being necessary to avoid deviation from randomization between $a_t = 1$ and $a_t = 0$ when $\theta_t = 01$.

Now, we examine Phase P. Analogous to L^E, let $L^P(\theta^c)$ denote the continuation value of the Lagrangian (under the efficient mechanism with q^*) from period t on in Phase P, conditional on $\theta_{t-1}^c = \theta^c$:

$$\begin{cases} L^P(0) = \frac{1-\beta}{2}\delta L^E(0) + \frac{1-\beta}{2}\left(v_{01} + \delta L^P(0)\right) + \beta\left(\frac{v_{10}+v_{11}}{2} + \delta L^P(1)\right) \\ L^P(1) = \frac{\beta}{2}\delta L^E(0) + \frac{\beta}{2}\left(v_{01} + \delta L^P(0)\right) + (1-\beta)\left(\frac{v_{10}+v_{11}}{2} + \delta L^P(1)\right), \end{cases}$$

$$\Leftrightarrow \begin{cases} L^P(0) = \frac{\beta\delta 2u_1 v_{01}(1-\delta-\beta+2\beta\delta)+2v_{01}(1-\delta)[1-\delta(1-2\beta)](1-\delta-\beta+2\beta\delta)+\beta[2-\delta(3-\delta-3\beta+2\beta\delta)](v_{10}+v_{11})}{2(1-\delta)[1-\delta(1-2\beta)][2-\delta(3-\delta-3\beta+2\beta\delta)]} \\ L^P(1) = \frac{-u_1 v_{01}(2-\delta-2\beta+2\beta\delta)+\beta v_{01}}{2-\delta(3-\delta-3\beta+2\delta\beta)} + \frac{(1-2\beta)(2u_1 v_{01}+v_{10}+v_{11})}{4[1-\delta(1-2\beta)]} + \frac{2u_1 v_{01}+v_{10}+v_{11}}{4(1-\delta)}. \end{cases}$$

Hence, for each h_t such that either $t = 1$ or $h_t^A \neq (1, \ldots, 1)$:

$$L_t(h_t) = \begin{cases} v(\theta_t) + \delta L^P(1) & \text{if} \quad \theta_t = 11, 10 \\ v(\theta_t) + \delta L^P(0) & \text{if} \quad \theta_t = 01 \\ \delta L^E(0) & \text{if} \quad \theta_t = 00, \end{cases}$$

noting that θ_t is a part of h_t.

The one-shot deviation principle requires the following. If $\theta_t = 11, 10$, then:

$$v(\theta_t) + \delta L^P(1) \geq \delta L^E(1),$$

which boils down to:

$$v_{10} + \delta L^P(1) \geq \delta L^E(1) \;\Leftrightarrow\; \tfrac{v_{10}}{v_{01}} \geq \tfrac{u_1(2-\delta-2\beta+2\beta\delta)+\beta}{2-3\delta+\delta^2+\beta\delta(3-2\delta)}\delta,$$

which is the second condition in the statement of the Theorem. If $\theta_t = 01$, then:

$$v(\theta_t) + \delta L^P(0) \geq \delta L^E(0) \;\Leftrightarrow\; \delta \leq \tfrac{1}{1+\beta u_1-\beta},$$

which is implied by Assumption 2 and hence always satisfied. If $\theta_t = 00$, then:

$$\delta L^E(0) \geq v(00) + \delta L^P(0) \;\Leftrightarrow\; \tfrac{1}{v_{01}} \geq \tfrac{1-\beta-\delta+2\beta\delta-2\beta u_1}{2-3\delta+\delta^2+\beta\delta(3-2\delta)}\delta,$$

which is the first condition in the statement of the Theorem. □

B.4 Proof of Corollary 1

Proof. This is a corollary of Theorem 1. Part 1: The first and second conditions in Theorem 1 are implied by $\delta \leq \frac{2}{v_{01}+1}$ and $\delta \leq \frac{v_{10}}{u_1 v_{01}+v_{10}}$, respectively. Part 2: If $\frac{2}{v_{01}+1} < \delta \leq \frac{v_{10}}{u_1 v_{01}+v_{10}}$, there exists $\beta(\delta) \in \left(0, \frac{1-2\delta}{2u_1+1-2\delta}\right)$ above which the first condition in Theorem 1 holds. □

B.5 Proof of Theorem 2

Proof. Consider the Lagrangian with the same multiplier structure Λ^* as in the proof of Theorem 1: $\Lambda_t^*(h_t^A) = v_{01}$ if $t \geq 1$ and $h_t^A \neq (1, \ldots, 1)$; otherwise $\Lambda_t^*(h_t^A) = 0$. Of course, the event that $h_t^A \neq (1, \ldots, 1)$ never happens under the no-disclosure mechanism. However, it is important to set the evolution of the multiplier under such "off-path" events, in order to apply the one-shot deviation principle.[34]

[34]One might wonder what do we mean by the "off-path" events here, as usually (fully-committed) Principal is not supposed to deviate. We think the best way to understand it is to separate the Lagrangian maximization (a single-person decision problem) from the original two-person Principal-Agent relationship: As in the standard dynamic programming, one must consider all contingencies (including the cases where the decision maker has deviated) in order to apply the Bellman principle. This is exactly our exercise here.

In addition, consider the two-phase no-disclosure mechanism as defined in Section 5.2.1. The difference with respect to the two-phase efficient mechanism is that the recommendation of the no-disclosure mechanism in the initial phase (Phase I) is always action 1, which is different from the efficient mechanism (and which makes Phase E occur with probability 0).

By Assumption 1, the no-disclosure mechanism satisfies all the obedience conditions. Let us now check optimality using the one-shot deviation principle.

The continuation value of the Lagrangian in Phase E is the same as before and hence omitted. Recalling that it does not depend on what happens in Phase I, the one-shot deviation principle in Phase E continues to hold.

Analogous to L^P, let $L^I(\theta^c)$ denote the continuation value of the Lagrangian from period t on in Phase I, conditional on $\theta^c_{t-1} = \theta^c$:

$$\begin{cases} L^I(0) = (1-\beta)\left(\frac{v_{00}+v_{01}}{2} + \delta L^I(0)\right) + \beta\left(\frac{v_{10}+v_{11}}{2} + \delta L^I(1)\right) \\ L^I(1) = \beta\left(\frac{v_{00}+v_{01}}{2} + \delta L^I(0)\right) + (1-\beta)\left(\frac{v_{10}+v_{11}}{2} + \delta L^I(1)\right), \end{cases}$$
$$\Leftrightarrow \begin{cases} L^I(0) = \frac{(1-\beta-\delta+2\beta\delta)(v_{00}+v_{01})+\beta(v_{10}+v_{11})}{2(1-\delta)[1-\delta(1-2\beta)]} \\ L^I(1) = \frac{(1-\beta-\delta+2\beta\delta)(v_{10}+v_{11})+\beta(v_{00}+v_{01})}{2(1-\delta)[1-\delta(1-2\beta)]}. \end{cases} \tag{2}$$

The one-shot deviation conditions require:

$$v(\theta_t) + \delta L^I(\theta^c_t) \geq L^E(\theta^c_t),\ \forall \theta_t,$$

which is equivalent to the combination of the following two conditions:

$$v_{00} + \delta L^I(0) \geq \delta L^E(0) \ \Leftrightarrow\ \tfrac{1}{v_{01}} \leq \tfrac{1-\beta-\delta+2\beta\delta-2\beta u_1}{2-3\delta+\delta^2+\beta\delta(3-2\delta)}\delta,$$
$$v_{10} + \delta L^I(1) \geq \delta L^E(1) \ \Leftrightarrow\ v_{10} \geq \tfrac{\beta-\beta v_{01}+2u_1 v_{01}(1-\beta-\delta+2\beta\delta)}{2(1-\delta)[1-\delta(1-2\beta)]}\delta.$$

The no-disclosure mechanism is thus optimal under these two conditions.

These two conditions always hold if $\frac{2}{v_{01}+1} < \delta \leq \frac{v_{10}}{u_1 v_{01}+v_{10}}$ and $\beta \leq \beta(\delta)$. The first condition is actually $\beta \leq \beta(\delta)$, where $\beta(\delta) > 0$ is strictly positive if and only if $\delta > \frac{2}{v_{01}+1}$. The second condition is implied by $\delta \leq \frac{v_{10}}{u_1 v_{01}+v_{10}}$. □

B.6 Proof of Theorem 3

Proof. From Proposition 2, the P-optimal mechanism is optimal if and only if $\beta \geq \frac{1}{8}$. Below, we find the parameter regions where the obfuscation mechanism and the no-disclosure mechanism are optimal. Define:

$$\beta^* \equiv \frac{5v_{01}+3v_{10}+2-\sqrt{184+144v_{01}+9v_{01}^2+30v_{01}v_{10}-152v_{10}+25v_{10}^2}}{4(5+4v_{01}-4v_{10})} \in \left(0, \tfrac{1}{8}\right).$$

Part I: Obfuscation mechanism is optimal if $\beta \geq \beta^*$.
The proof has a similar structure as that of Theorems 1 and 2. Based on Lemma 1 (Lagrangian weak duality), we guess: $\Lambda_t^*(h_t^A, a_t) = 0$ if h_t^A does not contain 0; $\Lambda_t^*(h_t^A, a_t) = \Lambda^P$ if h_t^A contains a single 0; and $\Lambda_t^*(h_t^A, a_t) = v_{01}$ otherwise.

Intuitively, the first part of the guess is that Agent's obedience constraints do not bind in Phase O, while the third part of the guess delivers the necessary indifference condition for the optimality of randomizing between $a = 0$ and $a = 1$ in state 01 in Phase E.

With the obfuscation policy, there is also randomization in Phase O in state 10. Therefore, the Lagrangian maximizer must be indifferent between $a = 0$ and $a = 1$. Hence, we can compute the continuation value by setting $a = 1$ with full probability if $\theta = 10$. The continuation value of the Lagrangian in phase O can thus be written recursively as follows:

$$\begin{cases} L^O(0) = \frac{1-\beta}{2}\delta L^P(0) + \frac{1-\beta}{2}\left(v_{01} + \delta L^O(0)\right) + \beta\left(\frac{v_{10}+v_{11}}{2} + \delta L^O(1)\right) \\ L^O(1) = \frac{\beta}{2}\delta L^P(0) + \frac{\beta}{2}\left(v_{01} + \delta L^O(0)\right) + (1-\beta)\left(\frac{v_{10}+v_{11}}{2} + \delta L^O(1)\right). \end{cases} \tag{3}$$

Similarly, the continuation value in Phase P satisfies:

$$\begin{cases} L^P(0) = \frac{1-\beta}{2}\delta L^E(0) + \frac{1-\beta}{2}\left(v_{01} - \Lambda^P + \delta L^P(0)\right) + \beta\left(\frac{v_{10}+v_{11}}{2} + \Lambda^P u_1 + \delta L^P(1)\right) \\ L^P(1) = \frac{\beta}{2}\delta L^E(0) + \frac{\beta}{2}\left(v_{01} - \Lambda^P + \delta L^P(0)\right) + (1-\beta)\left(\frac{v_{10}+v_{11}}{2} + \Lambda^P u_1 + \delta L^P(1)\right), \end{cases} \tag{4}$$

where $L^E(0)$ and $L^E(1)$ are the continuation values in the Phase E, given in (1).

The indifference condition in Phase O and state 10 is given by:

$$v_{10} + \delta L^O(1) = \delta L^P(1). \tag{5}$$

To find $\left\{L^O(0), L^O(1), L^P(0), L^P(1), \Lambda^P\right\}$, we solve a system of five linear equations, given by (3), (4) and (5).[35] The solution is:

$$\begin{aligned}
L^O(0) &= \tfrac{(58\beta-44)v_{01}-8\beta v_{10}+v_{10}}{4\beta(5\beta-1)-9} - \tfrac{4v_{01}+v_{10}+v_{11}}{2\beta+1} + v_{10} + v_{11} && (6)\\
L^O(1) &= \tfrac{1}{5}\left[\tfrac{9(29-36\beta)v_{01}-6(\beta+6)v_{10}}{4\beta(5\beta-1)-9} + \tfrac{5(4v_{01}+v_{10}+v_{11})}{2\beta+1} + 9v_{01} - 4v_{10}\right] && (7)\\
L^P(0) &= \tfrac{(2\beta(2\beta-1)(16\beta+15)-6)v_{01}+(3-8\beta(\beta(3\beta+10)+4))v_{10}+2\beta(4\beta(5\beta-1)-9)v_{11}}{(2\beta+1)[4\beta(5\beta-1)-9]} && (8)\\
L^P(1) &= \tfrac{1}{5}\left[\tfrac{9(29-36\beta)v_{01}-6(\beta+6)v_{10}}{4\beta(5\beta-1)-9} + \tfrac{5(4v_{01}+v_{10}+v_{11})}{2\beta+1} + 9v_{01} + 6v_{10}\right] && (9)\\
\Lambda^P &= \tfrac{(3+4\beta)^2 v_{10}+2\beta(1-8\beta)v_{01}}{18+8\beta(1-5\beta)} && (10)
\end{aligned}$$

The cumulative Lagrangian multiplier must be nondecreasing. We must thus have $0 \leq \Lambda^P \leq v_{01}$. It is straightforward from (10) that $0 < \Lambda^P$ and that $\Lambda^P \leq v_{01} \Leftrightarrow \frac{v_{10}}{v_{01}} \leq \frac{6(1-\beta)}{3+4\beta}$. Since $\frac{6(1-\beta)}{3+4\beta}$ is decreasing in β, we replace $\beta = \frac{1}{8}$ to find the sufficient condition $\frac{v_{10}}{v_{01}} \leq \frac{3}{2}$, which is part of Assumption 4.

The solution in (6)-(10) is the actual solution of the dynamic information design problem if the following additional conditions are satisfied:

(i) Lagrangian-maximizer does not deviate (in Phases O and P)

(ii) Lagrangian-maximizer does not enjoy payoff from the Lagrangian slack.

The second condition always holds (with an appropriately chosen probability distribution between $a = 0$ and $a = 1$ in state 10). See Lemma 3 in the Appendix.

[35] This problem can be solved analytically for arbitrary parameter values. However, the resulting expressions are long. In this proof, we are restricting to $\delta = \frac{1}{2}$ and $u_1 = 2$.

Let us check the no-deviation constraints.[36] Define:

$$\begin{aligned}
IC^O(00) &= \delta L^P(0) - v_{00} - \delta L^O(0) \\
IC^O(01) &= v_{01} + \delta L^O(0) - \delta L^P(0) \\
IC^P(00) &= \delta L^E(0) - v_{00} + \Lambda^P - \delta L^P(0) \\
IC^P(01) &= v_{01} - \Lambda^P + \delta L^P(0) - \delta L^E(0) \\
IC^P(10) &= v_{10} + \Lambda^P u_1 + \delta L^P(1) - \delta L^E(1).
\end{aligned}$$

Replacing the expressions obtained above, we obtain:

$$\begin{aligned}
IC^O(00) \geq 0 &\Leftrightarrow \tfrac{1-10\beta+16\beta^2}{9+4\beta-20\beta^2} v_{01} + \tfrac{1-6\beta-16\beta^2}{9+4\beta-20\beta^2} v_{10} \leq 1 \qquad (11) \\
IC^O(01) \geq 0 &\Leftrightarrow (1+2\beta)(1-8\beta)v_{10} + 2(1-\beta)(5+2\beta)v_{01} \geq 0 \\
IC^P(00) \geq 0 &\Leftrightarrow \tfrac{3-28\beta+32\beta^2}{9+4\beta-20\beta^2} v_{01} + \tfrac{-6-2\beta+8\beta^2}{9+4\beta-20\beta^2} v_{10} \leq 1 \\
IC^P(01) \geq 0 &\Leftrightarrow \tfrac{v_{10}}{v_{01}} \leq \tfrac{6(1-\beta)}{3+4\beta} \\
IC^P(10) \geq 0 &\Leftrightarrow \tfrac{v_{10}}{v_{01}} \geq \tfrac{18-38\beta+34\beta^2}{27+31\beta-16\beta^2}.
\end{aligned}$$

Observe that $IC^O(01) \geq 0$, and that $IC^P(00) \geq 0$ is implied by $IC^O(00) \geq 0$. In addition: $\frac{v_{10}}{v_{01}} \leq \frac{3}{2}$ implies $IC^P(01) \geq 0$; and $\frac{v_{10}}{v_{01}} \geq \frac{2}{3}$ implies $IC^P(10) \geq 0$. Hence, the only condition left to check is $IC^O(00) \geq 0$, i.e., condition (11).

Note that condition (11) fails for $\beta = 0$ if and only if $v_{01} + v_{10} > 9$, which we assume. By contrast, it always holds for $\beta = \frac{1}{8}$. Since condition (11) is quadratic in β, there is one and only one root in $\left(0, \frac{1}{8}\right)$, which is the threshold above which $IC^O(00) \geq 0$. This root turns out to be exactly β^*, and thus the proof that obfuscation is optimal if $\beta \geq \beta^*$ is complete.

Part II: No-disclosure is optimal if $\beta \leq \beta^*$.

To establish optimality of *no disclosure*, we consider the following policy:

- Phase I ($\Lambda^I = 0$): $a = 1$ in all states; if $a = 0$ (off-path), move to next phase

[36] Only these 5 constraints are non-trivial. In both Phases O and P, no deviation from $a = 1$ in state 10 implies no deviation from $a = 1$ in state 11.

- Phase P ($\Lambda^P \in [0, v_{01}]$): $a = 0$ in state 00; $a = 1$ in states 01, 10 and 11; if $a = 0$, move to next phase
- Phase E ($\Lambda^E = v_{01}$): mix between P-optimal and A-optimal.

The continuation values in Phases I and E remain given by L^I and L^E in (2) and (1). In Phase P, the recursive structure of the continuation values is as in (4), but with a possibly different multiplier:

$$\begin{cases} L^P(0) = \frac{1-\beta}{2}\delta L^E(0) + \frac{1-\beta}{2}\left(v_{01} - \Lambda^P + \delta L^P(0)\right) + \beta\left(\frac{v_{10}+v_{11}}{2} + \Lambda^P u_1 + \delta L^P(1)\right) \\ L^P(1) = \frac{\beta}{2}\delta L^E(0) + \frac{\beta}{2}\left(v_{01} - \Lambda^P + \delta L^P(0)\right) + (1-\beta)\left(\frac{v_{10}+v_{11}}{2} + \Lambda^P u_1 + \delta L^P(1)\right). \end{cases}$$

To determine the value of Λ^P (in addition to $L^P(0)$ and $L^P(1)$), we need an additional equation (besides the recursive structure). We guess that the multiplier is such that the Lagrangian-maximizer is indifferent between $a = 0$ and $a = 1$ in Phase I and state 10: $v_{10} + \delta L^I(1) = \delta L^P(1)$.

This condition, together with the recursive structure (4), yields:[37]

$$\begin{aligned} L^P(0) &= \frac{2\beta^2(8v_{01}-5v_{10}+3v_{11}+4)-\beta(7v_{01}+12v_{10}+6v_{11}+1)-2v_{01}+v_{10}}{3(\beta-1)(2\beta+1)} \\ L^P(1) &= \frac{2\beta(v_{01}+2v_{10}-1)+3v_{10}+v_{11}}{2\beta+1} \\ \Lambda^P &= \frac{(1+2\beta)(3+4\beta)v_{10}-\beta[3+4\beta-(1-8\beta)v_{01}]}{6(1-\beta)(2\beta+1)} \end{aligned}$$

The multiplier Λ^P is clearly positive, and $v_{01} \geq \frac{2}{3}v_{10}$ (which holds by assumption) implies $\Lambda^P \leq v_{01}$. Hence, this candidate is the actual solution if the Lagrangian maximizer does not deviate.[38] The relevant no-deviation constraints are in state 00 in Phase I, and in states 00, 01 and 10 in Phase P. Let:

$$\begin{aligned} IC^I(00) &= v_{00} + \delta L^I(0) - \delta L^P(0) \\ IC^P(00) &= \delta L^E(0) - v_{00} + \Lambda^P - \delta L^P(0) \\ IC^P(01) &= v_{01} - \Lambda^P + \delta L^P(0) - \delta L^E(0) \\ IC^P(10) &= v_{10} + \Lambda^P u_1 + \delta L^P(1) - \delta L^E(1). \end{aligned}$$

[37] Again, this system of linear equations can be solved for arbitrary parameter values.
[38] Note that the payoff from the slack is zero because only Phase I is on-path.

Using the expressions for $L^P(0)$, $L^P(1)$ and Λ^P obtained above, we obtain:

$$\begin{aligned}
IC^I(00) \geq 0 \;&\Leftrightarrow\; \tfrac{4(5v_{01}-1)}{2\beta+1} + \tfrac{7(v_{01}-3v_{10}+1)}{1-\beta} - 6(4v_{01} - 4v_{10} + 5) \geq 0 \qquad (12)\\
IC^P(00) \geq 0 \;&\Leftrightarrow\; 4\beta(2v_{01} + v_{10} + 1) - v_{01} + 2v_{10} + 3 \geq 0\\
IC^P(01) \geq 0 \;&\Leftrightarrow\; -\beta(v_{01} + 2v_{10} - 1) + 2v_{01} - v_{10} \geq 0\\
IC^P(10) \geq 0 \;&\Leftrightarrow\; -\beta(\beta + 6) - [\beta(11\beta - 10) + 6]v_{01} + (9 - 2\beta)(2\beta + 1)v_{10} \geq 0
\end{aligned}$$

Note that, with $\beta \in \left[0, \frac{1}{8}\right]$: $\frac{v_{10}}{v_{01}} \geq \frac{2}{3}$ implies $IC^P(00) \geq 0$ and $IC^P(10) \geq 0$; while $\frac{v_{10}}{v_{01}} \leq \frac{3}{2}$ implies $IC^P(01) \geq 0$.

If $\beta = 0$, condition (12) holds if and only if $v_{01} + v_{10} > 9$, which we assume. If $\beta = \frac{1}{8}$, the condition fails. Since condition (12) is quadratic in β, it has a single root in $\left(0, \frac{1}{8}\right)$, which is the threshold value for β below which $IC^I(00) \geq 0$ holds. Again, this root turns out to be exactly β^*, which concludes the proof. □

The following Lemma, invoked in the proof of Theorem 3, establishes complementary slackness. The obedience constraints are satisfied and strictly positive Lagrangian multipliers are associated with obedience constraints satisfied in equality (the Lagrangian-maximizer does not enjoy payoff from the slack).

Lemma 3. Consider the following class of policies (n-shot obfuscation):

- Phases O_1 to O_n: $a = 0$ if $\theta = 00$; $a = 1$ if $\theta \in \{01, 11\}$; mix between $a = 0$ and $a = 1$ if $\theta = 10$; if $a = 0$, move to next phase
- Phase P: P-optimal policy; if $a = 0$, move to next phase;
- Phase E: mix between the P-optimal and the A-optimal policies.

Under Assumptions 1-2, there exists a policy in this class such that Agent's obedience constraint for $a = 1$ is binding at each t conditionally on having been recommended $a = 0$ at $t-1$ (and all the other obedience constraints are satisfied).

Proof. Agent's expected payoff in each phase except Phase O_1 must be zero, as this means that Agent's obedience constraint for $a_t = 1$ is binding if $a_{t-1} = 0$.

In Phase E, while the A-optimal policy gives a strictly positive payoff to Agent, the P-optimal policy gives payoff $U_{fb}(0)$ to Agent (note that $\theta^c = 0$ in the last period of phase P), which is strictly negative under Assumption 2 (non-obedience of the first-best policy). Therefore, as shown in Lemma 2, there exists a mix between the P-optimal and the A-optimal policies (a history-independent probability of recommending $a_t = 1$ if $\theta_t = 01$) that gives zero expected payoff to Agent in Phase E, as desired, and which is such that all other obedience constraints in Phase E are satisfied.

In Phase P, Agent's expected payoff depends on the common state θ^c in the last period of phase O_n, being determined by:

$$\begin{cases} U^P(0) = \frac{1-\beta}{2}\left(u_0 + \delta U^P(0)\right) + \beta\left(u_1 + \delta U^P(1)\right) \\ U^P(1) = \frac{\beta}{2}\left(u_0 + \delta U^P(0)\right) + (1-\beta)\left(u_1 + \delta U^P(1)\right), \end{cases}$$

which yields:

$$\begin{cases} U^P(0) = \frac{-1+\delta+\beta(1+2u_1-2\delta)}{2-3\delta+\delta^2+3\beta\delta-2\beta\delta^2} & (< 0 \text{ under Assumption 2}) \\ U^P(1) = \frac{u_1(2-\delta-2\beta+2\beta\delta)-\beta}{2-3\delta+\delta^2+3\beta\delta-2\beta\delta^2} & (> 0 \text{ under Assumption 1}). \end{cases}$$

For the obedience constraint entering Phase P to hold in equality, Agent's posterior belief about θ^c when $a = 0$ is recommended in phase O_n (causing transition to Phase P), which we denote by $\psi^{O_n} \equiv Prob(\theta^c_t = 1 | a_t = 0 \text{ in Phase } O_n)$ must be equal to $\frac{-U^P(0)}{U^P(1)-U^P(0)}$. Note that $\psi^{O_n} \in \left(0, \frac{1}{2}\right)$, because $U^P(0) < 0$ (under Assumption 2) and $\frac{1}{2}U^P(0) + \frac{1}{2}U^P(1) > 0$ (under Assumption 1). The randomization between $a = 0$ and $a = 1$ if $\theta = 10$ in phase O_n must generate this precise belief (we verify later, using Lemma 4, that such randomization exists).

In Phase O_i, with $i \in \{2, n\}$, Agent's payoff is between the payoff of the policy that always recommends $a = 1$ in state 10 (P-optimal policy) and the payoff of the policy that always recommends $a = 0$ in state 10. The latter policy is fully uninformative about the common state (recommendation only depends on the i.i.d. state) and gives Agent half of the expected payoff of the no-disclosure policy. These two extreme policies (and thus any obfuscation policy) yield to

Agent: (i) strictly negative payoff in Phase O_i if $\theta^c = 0$ in the last period of phase O_{i-1} (under Assumption 2); (ii) strictly positive payoff in Phase O_i if $Prob(\theta^c = 1) \geq \frac{1}{2}$ in the last period of phase O_{i-1} (under Assumption 1). Therefore, there exists an intermediate belief about θ^c formed when $a = 0$ is recommended in phase O_{i-1} (causing transition to phase O_i), denoted $\psi^{O_{i-1}} \equiv Prob(\theta^c_t = 1 | a_t = 0 \text{ in Phase } O_{i-1}) \in \left(0, \frac{1}{2}\right)$, such that Agent's expected payoff in Phase O_i is zero.

We now apply Lemma 5 to show that the end-of-phase beliefs $\left\{\psi^{O_i}\right\}_{i=1}^n$ must be such that $\frac{1}{2} > \psi^{O_1} > ... > \psi^{O_{n-1}} > \psi^{O_n} > 0$, for Agent's expected payoff to be zero in all phases (except Phase O_1). First, notice that the P-optimal policy can be seen as a degenerate obfuscation policy that induces end-of-phase belief $\psi^P = 0$. Since Phase O_n has a greater end-of-phase belief than Phase P ($\psi^{O_n} > 0$), it must also have a greater end-of-previous-phase belief: $\psi^{O_{n-1}} > \psi^{O_n}$. Otherwise, from Lemma 5, Agent's payoff in Phase O_n would be strictly negative. For the same reason, since Phase O_{n-1} has a greater end-of-phase belief than Phase O_n ($\psi^{O_{n-1}} > \psi^{O_n}$), it must also have a greater end-of-previous-phase belief: $\psi^{O_{n-2}} > \psi^{O_{n-1}}$. By induction, we conclude that $\frac{1}{2} > \psi^{O_1} > ... > \psi^{O_{n-1}} > \psi^{O_n} > 0$.

Hence, we can apply Lemma 4 to establish existence of an n-shot obfuscation policy that generates the end-of-phase beliefs $\left\{\psi^{O_i}\right\}_{i=1}^n$ which are such that Agent's payoff in each phase (except phase O_1) is zero, so that Agent's obedience constraint for $a_t = 1$ is binding if $a_{t-1} = 0$.

Observing that (with this class of policies) recommendation $a_{t-1} = 1$ can only make Agent more optimistic, we conclude that any obedience constraint for $a_t = 1$ conditionally on $a_{t-1} = 1$ is also satisfied. □

Lemma 4. Let $\left\{\psi^{O_i}\right\}_{i=1}^n$ be such that $\frac{1}{2} \geq \psi^{O_1} \geq ... \geq \psi^{O_{n-1}} \geq \psi^{O_n}$. There exists an obfuscation policy (of the class defined in Lemma 3) that generates these end-of-phase beliefs.

Proof. The proof is based on the relations between beliefs and obfuscation probabilities given in equations (13)-(14) below.

Let $\psi_{t-1} \equiv Prob(\theta^c_{t-1} = 1 | h^A_{t-1})$ denote Agent's posterior at $t-1$, and let $\psi^0_t \equiv Prob(\theta^c_t = 1 | h^A_{t-1})$ denote Agent's "prior" at t. The two beliefs differ due to

Markovian transition (if $\alpha \neq 0$). Note that $\psi_t^0 = (1-\alpha)\psi_{t-1} + \alpha(1-\psi_{t-1})$, and thus: (i) ψ_t^0 is increasing in ψ_{t-1}; (ii) $\psi_{t-1} = \frac{1}{2} \Rightarrow \psi_t^0 = \frac{1}{2}$; (iii) $\psi_{t-1} < \frac{1}{2} \Rightarrow \psi_t^0 \in \left[\psi_{t-1}, \frac{1}{2}\right]$; (iv) $\psi_{t-1} > \frac{1}{2} \Rightarrow \psi_t^0 \in \left[\frac{1}{2}, \psi_{t-1}\right]$.

To generate the end-of-phase belief ψ^{O_i} (Agent's posterior at t if $a_t = 0$), the probability z_t of recommending $a_t = 0$ in state 10 must be such that:

$$\frac{\psi^{O_i}}{1-\psi^{O_i}} = \frac{\psi_t^0 \frac{z_t}{2}}{(1-\psi_t^0)\frac{1}{2}} \Leftrightarrow \frac{\psi^{O_i}}{1-\psi^{O_i}} = \frac{\psi_t^0}{1-\psi_t^0} z_t. \tag{13}$$

Notice also that if $a_t = 1$ is drawn, Agent's posterior at t is such that:

$$\frac{\psi_t}{1-\psi_t} = \frac{\psi_t^0(1-\frac{z_t}{2})}{(1-\psi_t^0)\frac{1}{2}} \Leftrightarrow \frac{\psi_t}{1-\psi_t} = \frac{\psi_t^0}{1-\psi_t^0}(2-z_t). \tag{14}$$

In Phase O_i, to generate end-of-phase belief ψ^{O_i}, it is necessary to set z_t such that $\frac{\psi^{O_i}}{1-\psi^{O_i}} = \frac{\psi_t^0}{1-\psi_t^0} z_t$. This is possible as long as $\psi_t^0 \geq \psi^{O_i}$ (because $z_t \leq 1$). The rest of the proof is to check that $\psi_t^0 \geq \psi^{O_i}$ holds for all t and all $i \in \{1, ..., n\}$.

To check that $\psi_t^0 \geq \psi^{O_1}$ in Phase O_1, note that, at $t = 1$: $\psi_1^0 = \frac{1}{2} \geq \psi^{O_1}$. At $t = 2$, since $\psi_1 \geq \frac{1}{2}$ (posterior conditional on $a_1 = 1$), we also have $\psi_2^0 \geq \frac{1}{2} \geq \psi^{O_1}$. For the same reason, $\psi_t^0 \geq \frac{1}{2} \geq \psi^{O_1}$ for all t (in Phase O_1).

Let us now consider Phase O_2. Recall that $\psi^{O_2} \leq \psi^{O_1} \leq \frac{1}{2}$. In the first period of Phase O_2, $\psi_t^0 \in \left[\psi^{O_1}, \frac{1}{2}\right]$, and thus $\psi_t^0 \geq \psi^{O_2}$. Again, since observing $a_t = 1$ makes Agent more optimistic, $\psi_{t+1}^0 \geq \psi_t^0$, and thus $\psi_{t+1}^0 \geq \psi^{O_2}$. The same is true for all t (in Phase O_2).

Exactly the same logic applies in all other phases. Hence, $\psi_t^0 \geq \psi^{O_i}$ for all t and all $i \in \{1, ..., n\}$. □

Lemma 5. With an obfuscation policy (of the class defined in Lemma 3) that generates end-of-phase beliefs $\left\{\psi^{O_i}\right\}_{i=1}^n$ such that $\frac{1}{2} \geq \psi^{O_1} \geq ... \geq \psi^{O_{n-1}} \geq \psi^{O_n}$, Agent's expected payoff in Phase O_i is increasing in $\psi^{O_{i-1}}$ and decreasing in ψ^{O_i}.

Proof. Equations (13)-(14) allow us to observe that, for a given end-of-previous-phase belief $\psi^{O_{i-1}}$, and thus a given ψ_t^0 in the first period of Phase O_i: (i) greater ψ^{O_i} implies greater z_t in the first period of Phase O_i; (ii) greater z_t implies smaller

ψ_t and thus smaller ψ^0_{t+1}; (iii) smaller ψ^0_{t+1} implies greater z_{t+1}. By induction, increasing ψ^{O_i} requires increasing z_t for all t in Phase O_i. Agent's payoff in Phase O_i is, therefore, decreasing in the end-of-phase belief, ψ^{O_i}.

Similarly, for a given end-of-phase belief ψ^{O_i}: (i) greater end-of-previous-phase belief $\psi^{O_{i-1}}$ (and thus greater ψ^0_t in the first period of Phase O_i) implies smaller z_t in the first period of Phase O_i; (ii) smaller z_t (and greater ψ^0_t) implies greater ψ_t and thus greater ψ^0_{t+1}; (iii) greater ψ^0_{t+1} implies smaller z_{t+1}. By induction, increasing $\psi^{O_{i-1}}$ requires decreasing z_t for all t in Phase O_i. Agent's payoff in Phase O_i is, therefore, increasing in the end-of-phase belief, $\psi^{O_{i-1}}$. □

C More General Parameter Spaces

C.1 No-disclosure (random transition after a deviation)

Consider now a different continuation policy after a deviation (by the fictitious Lagrangian-maximizer) from the *no disclosure* policy. Let the transition to Phase E (efficient policy) occur with probability $1-x$. With probability $x \in (0,1)$, there is no transition: the mechanism remains in Phase I and thus continues to recommend $a=1$ for all state histories as if there was no deviation.

The expressions for $L^I(\theta^c)$ and $L^E(\theta^c)$ remain the same (as well as $\Lambda^I = 0$ and $\Lambda^E = v_{01}$). They are given in (2) and (1), respectively. The no-deviation conditions are $IC^I_{rt}(00) \geq 0$ and $IC^I_{rt}(10) \geq 0$ where:

$$IC^I_{rt}(00) \equiv v_{00} + \delta(1-x)L^I(0) - \delta(1-x)L^E(0),$$
$$IC^I_{rt}(10) \equiv v_{10} + \delta(1-x)L^I(1) - \delta(1-x)L^E(1).$$

Equivalently:

$$\begin{cases} \frac{v_{00}}{1-x} + \delta L^I(0) - \delta L^E(0) \geq 0 \\ \frac{v_{10}}{1-x} + \delta L^I(1) - \delta L^E(1) \geq 0, \end{cases}$$

which means that the no-transition probability, x, increases the weight of the

current-period payoff (equivalently, decreases the weight of future payoffs), thereby relaxing $IC^I_{rt}(10) \geq 0$ and tightening $IC^I_{rt}(00) \geq 0$.

Under the following assumption, this continuation policy allows us to establish optimality of *no disclosure* in a larger parameter region.

Assumption 5. $v_{10} > \frac{2u_1 v_{01}}{v_{01}-1}$.

Assumption 5 is also used implicitly in Corollary 1, because it is exactly the condition under which $\frac{2}{v_{01}+1} < \frac{v_{10}}{v_{10}+v_{01}u_1}$.

Proposition 3. Under Assumptions 1-2 and 5, the *no disclosure* policy is optimal if the following no-deviation constraints hold:[39]

$$\begin{cases} \frac{1}{v_{01}} \leq \frac{1-\beta-\delta+2\beta\delta-2\beta u_1}{2-3\delta+\delta^2+\beta\delta(3-2\delta)}\delta \\ \delta \leq \frac{\beta(v_{10}-1)(2u_1v_{01}+v_{01}-1)+2u_1v_{01}-v_{01}v_{10}+v_{10}}{(1-2\beta)(2u_1v_{01}-v_{01}v_{10}+v_{10})} \end{cases} \tag{15}$$

Proof. The first condition in (15) corresponds to $IC^I_{rt}(00) \geq 0$ with $x = 0$. If $IC^I_{rt}(10) \geq 0$ also holds with $x = 0$, the result holds.

Suppose that $IC^I_{rt}(10) \geq 0$ is not satisfied with $x = 0$. In that case, there exists $x^* \in (0, 1)$ such that $IC^I_{rt}(10) = 0$, given by:

$$x^* = 1 - \frac{2\frac{1-\delta}{\delta}(1-\delta+2\beta\delta)v_{10}}{2u_1v_{01}(1-\delta-\beta+2\beta\delta)-\beta(v_{01}-1)}.$$

Replacing x^* in $IC^I_{rt}(00) \geq 0$, we obtain:

$$\beta + v_{10}(1-\delta-\beta+2\beta\delta) - \frac{2u_1v_{01}}{v_{01}-1}(1-\delta-\beta+2\beta\delta+\beta v_{10}) \geq 0.$$

If $\beta = 0$, this condition holds strictly if and only if Assumption 5 holds. Moreover, under Assumption 5, the constraint becomes tighter as δ increases, being satisfied if and only if:

$$\delta \leq \frac{\beta(v_{10}-1)(2u_1v_{01}+v_{01}-1)+2u_1v_{01}-v_{01}v_{10}+v_{10}}{(1-2\beta)(2u_1v_{01}-v_{01}v_{10}+v_{10})},$$

[39] The upper bound for δ in the second part of condition (15) is decreasing in β. With less persistence, players need to be more impatient for this condition to hold.

which (as we already know) is always true if $\beta = 0$. □

If $IC^I_{rt}(00) \geq 0$ fails with $x = 0$ (i.e., with transition to phase E occurring with certainty), then $IC^I_{rt}(00) \geq 0$ also fails with $x > 0$. In this case, random transition does not help. But if $IC^I_{rt}(00) > 0$ holds with $x = 0$, then (as we have shown) it also holds if x is such that $IC^I_{rt}(10) \geq 0$ is satisfied in equality and the second part of condition (15) is satisfied. In this case, random transition helps.

Figure 3 (gray region) plots the values of (β, δ) for which this random multiplier structure allows to establish optimality of no-disclosure, given by condition (15).

In particular, random transition supports optimality of no-disclosure if δ is high and the common state is perfectly persistent. With $\beta = 0$, $IC^I_{rt}(00) \geq 0$ is satisfied with $x = 0$ if and only if $\delta \geq \frac{2}{v_{01}+1}$. Since the second part of condition (15) and Assumption 2 are always satisfied if $\beta = 0$, we have the following corollary.

Corollary 2. Let $\beta = 0$. Under Assumptions 1 and 5: *no disclosure* is optimal if and only if $\delta \geq \frac{2}{v_{01}+1}$; the *best efficient* policy is optimal if and only if $\delta \leq \frac{2}{v_{01}+1}$.

Proof. Optimality of the best efficient policy is delivered by Theorem 1 because Assumption 5 means that the binding condition is $\delta \leq \frac{2}{v_{01}+1}$.

Optimality of no-disclosure is obtained by replacing $\beta = 0$ in Proposition 3, noting that Assumption 2 and the second part of (15) always hold if $\beta = 0$. □

C.2 Multiple obfuscation phases

Figure 2 (yellow region) shows the parameter region where the obfuscation mechanism (with a single obfuscation phase) defined in Section 5.2.2 is optimal. The borders are defined by the following constraints: (i) with greater β, the P-optimal policy is feasible and thus optimal; (ii) with smaller β, IC00 in Phase O fails and no-disclosure is optimal; (iii) with smaller δ, we have $\Lambda^P_{obf} > v_{01}$ and the best efficient is optimal; (iv) with greater δ, IC10 in Phase P fails.

In the latter case (greater δ), what is the optimal policy? Failure of IC10 in Phase P (deviation to $a = 0$ in state 10) leads us to guess that it may be optimal to introduce an additional obfuscation phase before Phase P. Let us then consider the following class of obfuscation policies with multiple obfuscation phases:

- Phases O_1 to O_n: $a = 0$ if $\theta = 00$; $a = 1$ if $\theta \in \{01, 11\}$; mix between $a = 0$ and $a = 1$ if $\theta = 10$; if $a = 0$, move to next phase;[40]
- Phase P: P-optimal policy; if $a = 0$, move to next phase;
- Phase E: mix between P-optimal and A-optimal ($\Lambda^E = v_{01}$).

In Phases O_1 to O_n, the Lagrangian-maximizer must be indifferent between $a = 0$ and $a = 1$ when $\theta = 10$. Hence, the continuation value of the Lagrangian can be computed by setting $a = 1$ with full probability. For $z \in \{O_1, ..., O_n\}$:

$$\begin{cases} L^z_{obf}(0) = & \frac{1-\beta}{2}\delta L^{z+1}_{obf}(0) + \frac{1-\beta}{2}\left(v_{01} - \Lambda^z_{obf} + \delta L^z_{obf}(0)\right) \\ & +\beta\left(\frac{v_{10}+v_{11}}{2} + \Lambda^z_{obf}u_1 + \delta L^z_{obf}(1)\right) \\ L^z_{obf}(1) = & \frac{\beta}{2}\delta L^{z+1}_{obf}(0) + \frac{\beta}{2}\left(v_{01} - \Lambda^z_{obf} + \delta L^z_{obf}(0)\right) \\ & +(1-\beta)\left(\frac{v_{10}+v_{11}}{2} + \Lambda^z_{obf}u_1 + \delta L^z_{obf}(1)\right), \end{cases} \tag{16}$$

where $O_{n+1} \equiv P$. Similarly, the continuation value in Phase P satisfies:

$$\begin{cases} L^P_{obf}(0) = & \frac{1-\beta}{2}\delta L^E(0) + \frac{1-\beta}{2}\left(v_{01} - \Lambda^P_{obf} + \delta L^P_{obf}(0)\right) \\ & +\beta\left(\frac{v_{10}+v_{11}}{2} + \Lambda^P_{obf}u_1 + \delta L^P_{obf}(1)\right) \\ L^P_{obf}(1) = & \frac{\beta}{2}\delta L^E(0) + \frac{\beta}{2}\left(v_{01} - \Lambda^P_{obf} + \delta L^P_{obf}(0)\right) \\ & +(1-\beta)\left(\frac{v_{10}+v_{11}}{2} + \Lambda^P_{obf}u_1 + \delta L^P_{obf}(1)\right). \end{cases} \tag{17}$$

In Phase E, the continuation values are $L^E(0)$ and $L^E(1)$ obtained in (1).

With a single obfuscation phase ($n = 1$), to find the continuation values in phases O_1 and P, and the multiplier in phase P, we solve a linear system of five equations: the recursive structure of the continuation values in phases O_1 and P, given by (3) and (4), plus the binding condition IC10[41] in Phase O_1, given by $v_{10} + \delta L^{O_1}_{obf}(1) = \delta L^P_{obf}(1)$; and five unknowns: $L^{O_1}_{obf}(0)$, $L^{O_1}_{obf}(1)$, $L^P_{obf}(0)$, $L^P_{obf}(1)$ and Λ^P_{obf}. This system of linear equations can be solved analytically for arbitrary $v(\cdot)$ and $u(\cdot)$. However, the resulting expressions are long.

[40] In Phase O_z, with $z < n$, move to Phase O_{z+1}. In Phase O_n, move to Phase P.

[41] For simplicity, we write ICθ to refer to the no-deviation constraint in state θ.

With multiple obfuscation phases ($n > 1$), the solution can still be computed (at least if n is not too large) although this becomes more complicated as n increases because the number of linear equations that need to be solved is $3n + 2$.

Finally, we need to check: that no-deviation constraints are satisfied, and that the cumulative multiplier is nonincreasing.[42] The parameter region where all these conditions are satisfied is the region of optimality of this mechanism.

In Figure 3 (blue regions), we present the optimality regions for obfuscation policies with up to nine obfuscation phases.

C.3 No-disclosure (obfuscation off-path)

Similarly, we are able to establish optimality of *no disclosure* for greater values of δ by considering the following class of *obfuscation* policies off-path:

- Phase I ($\Lambda^I = 0$): $a = 1$ in all states; if $a = 0$, move to next phase
- Phases O_1 to O_n: $a = 0$ in state 00; $a = 1$ in states 01 and 11; mix in state 10; if $a = 0$, move to next phase
- Phase P: P-optimal; if $a = 0$, move to next phase
- Phase E ($\Lambda^E = v_{01}$): mix P-optimal and A-optimal.

The continuation values in Phases I and E remain given by (2) and (1). In phases O_1 to P, the recursive structure of the continuation values is as in (16)-(17), but with possibly different values for the multipliers.

The candidate solution is obtained by solving a linear system with $3n+3$ equations: the recursive structure in Phases O_1 to P, and binding IC10 in Phases I to O_n; and $3n+3$ unknowns: the continuation values and the cumulative Lagrangian multipliers in Phases O_1 to P. For example, with $n = 1$, there are two equations besides the recursive structure, corresponding to binding IC10 in Phases I and O_1:

$$v_{10} + \delta L^I_{nd}(1) \geq \delta L^{O_1}_{nd}(1)$$
$$v_{10} + \Lambda^{O_1}_{nd} u_1 + \delta L^{O_1}_{nd}(1) \geq \delta L^P_{nd}(1),$$

[42] From Lemma 3, we know that the Lagrangian-maximizer does not get payoff from the slack.

and two multipliers to determine: $\Lambda_{nd}^{O_1}$ and Λ_{nd}^{P}.

Again, the linear system can be solved analytically for arbitrary $v(\cdot)$ and $u(\cdot)$, but the resulting expressions are long (and become longer as n increases).

The resulting candidate is the optimal mechanism if: the no-deviation constraints are satisfied, and the cumulative multiplier is nonincreasing. These conditions yield the parameter region where the mechanism is optimal.

Figure 3 (yellow regions) puts together the optimality regions of the *no disclosure* policy with up to eight obfuscation phases off-path.

References

Abreu, D., Pearce, D., & Stacchetti, E. (1990). Toward a theory of discounted repeated games with imperfect monitoring. Econometrica, 58(5), 1041-1063

Au, P.H. (2015). Dynamic information disclosure. RAND Journal of Economics, 46(4), 791-823

Ball, I. (2023). Dynamic information provision: Rewarding the past and guiding the future. Econometrica, 91(4), 1363-1391

Bergemann, D., & Morris, S. (2019). Information design: A unified perspective. Journal of Economic Literature, 57(1), 44-95

Carrasco, V., Fuchs, W., & Fukuda, S. (2019). From equals to despots: The dynamics of repeated decision making in partnerships with private information. Journal of Economic Theory, 182, 402-432

Che, Y.-K., & Mierendorff, K. (2019). Optimal dynamic allocation of attention. American Economic Review, 109(8), 2993-3029

Crawford, V.P., & Sobel, J. (1982). Strategic information transmission. Econometrica, 50(6), 1431-1451

Ely, J.C. (2017). Beeps. American Economic Review, 107(1), 31-53

Ely, J. C., & Szydlowski, M. (2020). Moving the goalposts. Journal of Political Economy, 128(2), 468-506

Ichihashi, S. (2019). Limiting Sender's information in Bayesian persuasion. Games and Economic Behavior, 117, 276-288

Kamenica, E. (2019). Bayesian persuasion and information design. Annual Review of Economics, 11, 249-272

Lyu, C. (2023). Information Design for Social Learning on a Recommendation Platform. Available at SSRN. http://dx.doi.org/10.2139/ssrn.4275515

Marcet, A., & Marimon, R. (2019). Recursive contracts. Econometrica, 87(5), 1589-1631

Nikandrova, A., & Pancs, R. (2018). Dynamic project selection. Theoretical Economics, 13(1), 115-143

Orlov, D., Skrzypacz, A., & Zryumov, P. (2020). Persuading the principal to wait. Journal of Political Economy, 128(7), 2542-2578

Renault, J., Solan, E., & Vieille, N. (2017). Optimal dynamic information provision. Games and Economic Behavior, 104, 329-349

Spear, S.E., & Srivastava, S. (1987). On repeated moral hazard with discounting. The Review of Economic Studies, 54(4), 599-617

Zhao, W. (2022). Essays in Information Design and network theory (Ph.D. dissertation, Chapter 4). Jouy-en Josas, HEC.

Zhao, W., Mezzetti, C., Renou, L., & Tomala, T. (2024). Contracting over Persistent Information. Theoretical Economics, 19, 917-974