

# Marginal Reputation

Daniel Luo and Alexander Wolitzky

MIT\*

December 17, 2024

## Abstract

We study reputation formation where a long-run player repeatedly observes private signals and takes actions. Short-run players observe the long-run player's past actions but not her past signals. The long-run player can thus develop a reputation for playing a distribution over actions, but not necessarily for playing a particular mapping from signals to actions. Nonetheless, we show that the long-run player can secure her Stackelberg payoff if distinct commitment types are statistically distinguishable and the Stackelberg strategy is *confound-defeating*. This property holds if and only if the Stackelberg strategy is the unique solution to an optimal transport problem. If the long-run player's payoff is supermodular in one-dimensional signals and actions, she secures the Stackelberg payoff if and only if the Stackelberg strategy is monotone. An application of our results provides a reputational foundation for a class of Bayesian persuasion solutions when the sender has a small lying cost. Our results extend to the case where distinct commitment types may be indistinguishable but the Stackelberg type is *salient* under the prior.

**Keywords.** Reputation, repeated games, confound-defeating, optimal transport, cyclical monotonicity, signaling, Bayesian persuasion.

**JEL Codes.** C73, D83

---

\*For helpful comments, we thank Sandeep Baliga, Ian Ball, V. Bhaskar, Drew Fudenberg, Eric Gao, Navin Kartik, Anton Kolotilin, Stephen Morris, David Pearce, Harry Pei, and Tom Wiseman.

# 1 Introduction

This paper considers reputation formation in settings where one desires a reputation not only for taking certain actions, but for acting in the right circumstances. Our main applications are to deterrence, trust, and communication games, where the importance of establishing a reputation for *conditional* action has long been accepted in the informal literature. For example, Schelling (1966) writes,

“any coercive threat requires corresponding *assurances*; the object of a threat is to give somebody a choice. To say, “One more step and I shoot,” can be a deterrent threat only if accompanied by the implicit assurance, “And if you stop I won’t.” Giving notice of *unconditional* intent to shoot gives him no choice.”

Similarly, when an informed sender asks a receiver to take an action that the sender prefers, the request is persuasive only if the receiver believes that the sender tends to make it only when compliance is in the receiver’s interest.

To study reputation formation in these settings, we consider a model where a long-run player facing a sequence of short-run opponents repeatedly observes private signals and takes actions. For example, in the deterrence context, the signal is whether the long-run player detects an attack by a short-run player who moves first, and the action is whether she fights back. In the communication context, the long-run player is the first-mover, the signal is a payoff-relevant state variable, and the action is a signal or message to a short-run player who moves second. The long-run player is either rational or is one of a number of possible *commitment types* that play a fixed mapping from signals to actions in each period. The set of possible commitment types includes the *Stackelberg type* that plays the long-run player’s most-preferred commitment strategy.

In this setup, if short-run players observe the history of the long-run player’s past actions *and signals*, standard results imply a patient long-run player is assured at least her Stackelberg (best commitment) payoff in every Nash equilibrium (Fudenberg and Levine 1989; 1992). We instead consider the case where the long-run player’s actions are observed, but her signals are not. In the deterrence context, this says that potential attackers know when the long-run player has fought in the past, but not whether this fighting was a response to detected attacks. In the communication context, it says that receivers observe the history of messages sent by the long-run player, but not the history of states.

Existing results say little about the outcomes of these games. The key issue is that the long run player’s strategy—how she maps signals to actions—is not identified by the observed

marginal distribution over her actions. This implies that existing payoff bounds for reputation games with imperfect monitoring (Fudenberg and Levine, 1992; Gossner, 2011) are extremely loose and often trivial in our setting. For example, suppose that in the deterrence context the long-run player follows her (pure) Stackelberg strategy of fighting if and only if she detects an attack. This strategy results in the long-run player fighting a certain fraction of the time, say 50%. However, after seeing her fight half the time, potential attackers need not come to believe that she is playing the Stackelberg strategy—they might instead believe that she is playing a different strategy with the same marginal over actions, such as fighting half the time independent of her signal. Which inference potential attackers draw is critical for the long-run player, as they will be deterred if they believe she is playing the Stackelberg strategy of fighting when she detects an attack, but not if they believe she is randomly fighting half the time. As existing results do not restrict the short-run players’ inferences in this situation, they make no non-trivial predictions about the long-run player’s payoff. Formally, in this example attacking and not attacking are both “0-confirmed best responses” to the Stackelberg strategy, so Fudenberg and Levine’s (1992) lower bound is vacuous.

Our main result provides conditions for a patient long-run player to secure her Stackelberg payoff when only the marginal over actions is identified. The key sufficient condition is that the Stackelberg strategy is *confound-defeating*: against any 0-confirmed best response, the Stackelberg strategy is uniquely optimal among strategies that induce the same marginal over actions. Intuitively, if the Stackelberg strategy is confound-defeating then the rational long-run player never plays a different strategy that induces the same marginal in any Nash equilibrium. Therefore, establishing a reputation for playing the “Stackelberg marginal” suffices to establish a reputation for playing the Stackelberg strategy.

A strategy is confound-defeating if and only if the induced joint distribution over the long-run player’s signal and action is uniquely optimal among all distributions with the same marginals: that is, if and only if it is the unique solution to the optimal transport problem of maximizing the long-run player’s payoff subject to given marginals over her signal and action. Adapting standard optimal transport results, we show that this holds if and only if the support of the induced joint distribution satisfies a strict version of cyclical monotonicity (Rochet, 1987). We also use cyclical monotonicity to provide a converse to our main result: if the long-run player is rational with high probability, her payoff in any equilibrium cannot exceed that from some cyclically monotone strategy.

The cyclical monotonicity characterization makes confound-defeatingness easy to check in many games. In particular, if the long-run player’s payoff is supermodular in one-dimensional

signals and actions, a strategy is confound-defeating if and only if every selection from its support is monotone. Applied to the deterrence game, this says that the long-run player can secure her Stackelberg payoff if fighting is relatively more appealing when an attack is detected. Conversely, if the long-run player is rational with high probability, her payoff cannot exceed that from a monotone strategy in any Nash equilibrium. For example, in the deterrence game, the long-run player obtains close to her minmax payoff in every Nash equilibrium if fighting is more appealing when an attack is not detected.

Our results have strong implications for repeated communication games. An immediate implication is that, in repeated signaling games where the sender’s payoff is additively separable in the receiver’s action and supermodular in her own type and action, a patient sender can secure her best commitment payoff from any monotone signaling strategy.<sup>1</sup> We then consider repeated cheap talk games with state-independent sender preferences (which coincide with Bayesian persuasion when the sender has commitment power). Here, we show that perturbing the sender’s payoff by adding a small submodular “lying cost” yields a reputational foundation for any communication mechanism that is monotone with respect to some order on states and receiver actions. While not fully general, this class includes all partitions (deterministic communication mechanisms) and all linear partitions with randomization at the boundaries. We are thus able to provide a reputational foundation for a general class of communication mechanisms, even when the history of realized states is unobserved.

Our exposition assumes that distinct commitment types in the support of the short-run players’ prior are statistically distinguishable. Without this assumption, non-Stackelberg play by commitment types with the same marginal as the Stackelberg strategy can hinder reputation formation. Nonetheless, we show in Appendix A that our results extend without this assumption, so long as the Stackelberg type is sufficiently *salient* under the prior. Roughly, this condition says that the Stackelberg type has sufficiently high prior weight relative to other commitment types that induce the same marginal but different best responses. In the deterrence context, this says that short-run players believe that, conditional on the long-run player being irrational, she is more much likely to play the strategy “fight if and only if an attack is detected” than the strategy “fight half the time independent of the signal.”

---

<sup>1</sup>The results of Fudenberg and Levine (1992) imply the sender can secure her Stackelberg payoff in repeated signaling games where actions *and states* are observed at the end of each period. Our results imply the same conclusion holds when only actions are observed, if the Stackelberg strategy is monotone and the sender’s payoff is additively separable in the receiver’s action and supermodular in her own type and action.

**Related Literature.** We contribute to the literature on reputation formation with imperfect monitoring, introduced by Fudenberg and Levine (1992). They show that a patient long-run player can ensure her commitment payoff against the least favorable of her opponent’s 0-confirmed best responses. In our partially identified setting, the set of 0-confirmed best responses is typically large, so this payoff lower bound is weak and often vacuous. Gossner (2011) gives a different proof—which we build on—of a similar lower bound, which is also too weak in our setting for the same reason. Fudenberg and Levine and Gossner also give upper bounds on a patient long-run player’s payoffs, which Ely and Valimaki (2003) show is much too loose in a class of games where short-run players have outside options. In contrast, we show that their lower bounds are much too loose in a class of games where the long-run player observes private signals and the Stackelberg strategy is confound-defeating.<sup>2</sup>

Pei (2020) studies a reputation model with interdependent values, where a possibly committed long-run player privately observes a perfectly persistent, payoff-relevant state. Our model instead covers (as a special case) the case where the state is i.i.d.: see Section 2.2. In both papers, supermodularity-type conditions are important for securing the Stackelberg payoff, but the precise conditions and arguments are very different.<sup>3</sup> Other papers in the reputation literature where super/submodularity conditions play key roles in deriving payoff bounds include Liu (2011), Liu and Skrzypacz (2014), and Pei (2024).

We also relate to a diverse literature on games and mechanisms where strategies are partially identified, so certain deviations are undetectable. The connection between incentive compatibility and optimal transport in such settings dates to Rochet (1987); Rahman (2024) gives an alternative interpretation and proof. Applications include quota mechanisms (Jackson and Sonnenschein, 2007; Matsushima, Miyazaki, and Yagi, 2010; Escobar and Toikka, 2013; Frankel, 2014; Ball and Kattwinkel, 2024), multidimensional or repeated cheap talk (Chakraborty and Harbaugh, 2010; Renault, Solan, and Vieille, 2013; Margaria and Smolin, 2018; Meng, 2021), and repeated random matching games (Takahashi, 2010; Heller and Mohlin, 2018; Clark, Fudenberg, and Wolitzky, 2021). A particularly related paper is Lin and Liu (2024), who study optimal information disclosure when only the marginal distribution over signals is observed. In their setting, a joint distribution over states and receiver

---

<sup>2</sup>Ely, Fudenberg, and Levine (2008) add “good commitment types” to Ely and Valimaki (2003) and show that this does not restore the long-run player’s commitment payoff. The reason for the difference from our results is that we assume that the long-run player’s action is always identified, whereas in Ely and Valimaki (2003) and Ely, Fudenberg, and Levine (2008) it is not identified when the short-run player exits.

<sup>3</sup>For example, Pei’s Stackelberg payoff theorem requires binary actions for the short-run player and a condition on the prior, while we require no such conditions.

actions is implementable if it maximizes the sender’s payoff over all joint distributions with the same marginals. In contrast, our confound-defeating property requires the joint distribution over states and *sender* actions to uniquely maximize the sender’s payoff over all joint distributions with the same marginals, for any receiver best response. The two conditions are thus related, but involve different objects (distributions over states and receiver actions vs. sender actions) and come from different strategic considerations (static information disclosure subject to a “credibility” constraint vs. long-run reputation formation). They also apply to different classes of communication games: Lin and Liu’s “credibility” constraint has bite only if the sender’s utility is *state-dependent*, while our commitment payoff theorem applies only if the sender’s utility is *state-independent*.

We are also not the first to discuss reputational foundations for the Bayesian persuasion commitment assumption. The commitment assumption has been controversial ever since its introduction by Rayo and Segal (2010) and Kamenica and Gentzkow (2011). Rayo and Segal (2010) suggested reputation effects as a possible foundation. Mathevet, Pearce, and Stacchetti (2024) observe that reputation effects provide a valid foundation for the commitment payoff when a long-run sender faces a sequence of short-run receivers who observe the history of messages *and states*. In this setting, they study whether the sender’s *behavior* likewise coincides with the commitment solution. We instead ask when reputation effects yield the commitment payoff when receivers observe past messages but not states. Farther afield, Pei (2023) and Best and Quigley (2024) show that the commitment payoff arises as one of many equilibrium payoffs in repeated communication game with lying costs and coarse information on histories, respectively; Kuvalekar, Lipnowski, and Ramos (2022) study repeated sender-receiver games with two long-run players and unobserved past states, without commitment types; and Fudenberg, Gao, and Pei (2022) provide a foundation for the commitment payoff in a model where the long-run player sends messages before taking actions and can develop a reputation for “honesty” about the action they are about to take.

**Organization.** Section 2 analyzes the deterrence game discussed above, as well as repeated trust and signaling games. Section 3 develops the general model. Section 4 presents the confound-defeating property and our main result: the long-run player can secure her Stackelberg payoff if the Stackelberg strategy is confound-defeating. Section 5 characterizes confound-defeatingness via cyclical monotonicity, gives a converse to the main result, and applies our results to one-dimensional supermodular games. Section 6 considers communication games. Section 7 summarizes the paper and discusses some extensions and directions for

future work. Appendix A extends our results to allow indistinguishable commitment types by introducing our salience notion. Appendix B contains omitted proofs.

## 2 Motivating Examples

We begin with three examples: a deterrence game, a trust (or “product choice”) game, and a signaling game. These examples illustrate our main results in simple settings. Among other simplifications, we assume here that the long-run player has a single commitment type, which plays a pure strategy. The general analysis relaxes these assumptions.

### 2.1 Deterrence

We start with a deterrence game.<sup>4</sup> A long-run player with discount factor  $\delta$  faces a sequence of short-run opponents. Each period, the short-run player first chooses whether to *Cooperate* ( $C$ ) or *Defect* ( $D$ ). The long-run player then observes a private signal,  $c$  or  $d$ , generated with conditional probability  $\Pr(c|C) = \Pr(d|D) = p \in (0, 1)$ , before choosing whether to *Accommodate* ( $A$ ) or *Fight* ( $F$ ). The short-run player observes the history of the long-run player’s past actions, but not her signals. The short-run player’s payoff is specified as a function of both players’ actions while the long-run player’s payoff is specified as a function of her action and private signal as follows.<sup>5</sup>

	$C \quad D$		$c \quad d$
$A$	$1 \quad 1 + g$	$A$	$1 \quad y$
$F$	$-l \quad 0$	$F$	$x \quad 0$
Short-Run Player Payoff		Long-Run Player Payoff	

Assume that  $g, l > 0$  and  $x, y \in (0, 1)$ , so that  $D$  is dominant for the short-run player and  $A$  is dominant for the long-run player. Assume also that  $p$  is sufficiently large that the short-run player strictly prefers to take  $C$  if the long-run player plays the “deterrence” strategy,

---

<sup>4</sup>The stage game in this subsection is an example of an *inspection game* (Avenhaus, von Stengel, and Zamir (2002)). Acemoglu and Wolitzky (2024) survey deterrence and related games in economics and political science, emphasizing the role of private signals.

<sup>5</sup>An interpretation of the dependence of the long-run player’s payoff on her signal rather than the short-run player’s action is that each period’s short-run player is “crazy” with probability  $2(1 - p)$ , independently across periods, in which case she mixed 50-50 between her actions, and this is the only source of noise. In this case, the long-run player’s “signal” is just the short-run player’s action, accounting for the crazy types.

( $A$  after  $c$ ;  $F$  after  $d$ ), which we denote as  $(A, F)$ ; and that the long-run player would rather play  $(A, F)$  against  $C$  than play  $(A, A)$  against  $D$ , so that  $(A, F)$  is the long-run player's pure Stackelberg strategy.<sup>6</sup> Finally, assume that the long-run player is committed to  $(A, F)$  with probability  $\mu_0$  and is rational otherwise.

What can be said about the equilibria of this game when the discount factor  $\delta$  is close to 1? A first observation is that, in Fudenberg and Levine's (1992) terminology, both  $C$  and  $D$  are "0-confirmed best responses" to the Stackelberg strategy  $(A, F)$ . (We review this definition in Section 3.) For example, if the short-run player takes  $D$ , the long-run player ends up fighting with probability  $p$  when she plays  $(A, F)$ , but also when she fights with probability  $p$  after each signal. Since  $D$  is a best response to the latter strategy, it is also a 0-confirmed best response to  $(A, F)$ .

The results of Fudenberg and Levine (1992) (Theorem 3.1) and Gossner (2011) (Corollary 1) state that, as  $\delta \rightarrow 1$ , the long-run player's payoff in any Nash equilibrium is at least her payoff from playing the Stackelberg strategy against the least favorable 0-confirmed best response. (See Theorem 0 in Section 3.2.) Since  $D$  is a 0-confirmed best response to  $(A, F)$ , these results say only that the long-run player's payoff is above  $1 - p$ . When noise is small, so that  $p$  is close to 1, this just says that the long-run player's payoff is feasible.<sup>7</sup>

In contrast, we have the following result. Note that the long-run player's pure Stackelberg payoff is  $p$ , while her minmax payoff is  $1 - p + py$ .

**Proposition 1.** *Let  $\underline{U}_1(\delta)$  and  $\bar{U}_1(\delta)$  be the infimum and supremum of the long-run player's payoff in any Nash equilibrium. The following hold:*

- (1) *If  $x + y < 1$  then  $\liminf_{\delta \rightarrow 1} \underline{U}_1(\delta) \geq p$  for all  $\mu_0 > 0$ .*
- (2) *If  $x + y > 1$  then  $\lim_{\mu_0 \rightarrow 0} \bar{U}_1(\delta) = 1 - p + py$  for all  $\delta < 1$ .*

That is, if  $x + y < 1$ , a patient long-run player is assured at least her Stackelberg payoff in any Nash equilibrium; while if  $x + y > 1$ , the long-run player obtains close to her minmax payoff whenever the prior commitment probability is small. In particular, when noise is small, the long-run player is assured her best feasible payoff in the first case and her minmax payoff in the second.

To see why it matters whether  $x + y$  is below or above 1, note that the long-run player's payoff is strictly supermodular in her signal and action in the first case (with the order

<sup>6</sup>These two conditions holds iff  $p > \max\{\frac{1+g+l}{2+g+l}, \frac{1}{2-y}\}$ .

<sup>7</sup>Fudenberg and Levine and Gossner also give payoff upper bounds in terms of the mixed Stackelberg action. Our general results similarly allow mixed commitment types.



$A \succ F, c \succ d$ ) and is strictly submodular in the second. In the strictly supermodular case,  $F$  is relatively more appealing when  $d$  is observed. This implies that the Stackelberg strategy  $(A, F)$  strictly outperforms any other strategy that induces the same marginal over actions, and hence is “confound-defeating.” As we show in Theorem 1, this implies that establishing a reputation for playing the Stackelberg marginal over actions suffices to establish a reputation for playing the Stackelberg strategy, and hence secures the Stackelberg payoff.

Conversely, in the strictly submodular case,  $F$  is relatively more appealing when  $c$  is observed. This implies that the rational long-run player always plays  $F$  with (weakly) higher probability when  $c$  is observed, and plays  $A$  with higher probability when  $d$  is observed. Such a strategy *encourages* the short-run player to take  $D$  rather than deterring him, so the short-run player must take  $D$  whenever he believes that the long-run player is rational with high probability. Finally, since beliefs are martingale, it follows that the short-run player usually takes  $D$  when the ex ante commitment probability is small.<sup>8</sup>

## 2.2 Trust

Now suppose that the stage game is the following “trust game” (or “product choice game”) adapted from Pei (2020). There is a *state*  $\theta \in \{G(\text{ood}), B(\text{ad})\}$ , drawn i.i.d. across periods with equal probability on each state. In each period, the long-run player observes  $\theta$  before taking an action  $a_1 \in \{H(\text{igh}) \text{ Effort}, L(\text{ow}) \text{ Effort}\}$ . Simultaneously (and having observed the history of past actions, but not states), the short-run player takes an action  $a_2 \in \{T(\text{rust}), N(\text{ot Trust})\}$ . Payoffs in each state are given by the following matrices, with the long-run player’s payoff listed first in each entry.

	$T$	$N$		$T$	$N$
$H$	1, 2	−1, 0	$H$	$1 - w, -1$	$-1 - z, 0$
$L$	2, −1	0, 0	$L$	2, −1	0, 0
Payoffs in State $\theta = G$			Payoffs in State $\theta = B$		

Assume that  $w, z > -1$ , so that  $L$  is dominant for the long-run player, and the unique stage-game Nash equilibrium outcome is  $(L, N)$  in both states. Note that  $T$  is optimal for the short-run player only if he believes that, with high enough probability, *both*  $\theta = G$  and  $a_1 = H$ . For example, suppose that player 1 is the chef of a seafood restaurant,  $\theta$  is the quality of the day’s catch, and player 2 is a customer who wants to eat only fish that is

---

<sup>8</sup>The submodular case of Proposition 1 follows from Corollary 3 in Section 5.3.

both of high quality and carefully cooked. Note that the long-run player's pure Stackelberg strategy is  $(H, L)$  (as this strategy lets the long-run player enjoy taking  $L$  in state  $B$  while inducing the short-run player to take  $T$ ), and assume that she is committed to this strategy with probability  $\mu_0$ . Note also that the long-run player's pure Stackelberg payoff is  $3/2$ , while her minmax payoff is  $0$ .

**Proposition 2.** *The following hold:*

- (1) *If  $\min\{w, z\} > 0$  then  $\liminf_{\delta \rightarrow 1} \underline{U}_1(\delta) \geq 3/2$  for all  $\mu_0 > 0$ .*
- (2) *If  $\max\{w, z\} < 0$  then  $\lim_{\mu_0 \rightarrow 0} \overline{U}_1(\delta) = 0$  for all  $\delta < 1$ .*

While the timing of the deterrence and trust games are different (e.g., the short-run player moves first in the former and simultaneously with the long-run player in the latter), the logic of Proposition 2 is similar to that of Proposition 1. If  $\min\{w, z\} > 0$  then the long-run player's payoff is strictly supermodular in  $(\theta, a_1)$  for any  $a_2$ , which we show lets her secure her Stackelberg payoff when she is patient. If instead  $\max\{w, z\} < 0$  then the long-run player's payoff is strictly submodular in  $(\theta, a_1)$  for any  $a_2$ , which we show limits her to her minmax payoff when the prior commitment probability is small.<sup>9</sup>

## 2.3 Signaling

Finally, suppose the stage game is a signaling game. The state  $\theta \in \Theta \subset \mathbb{R}$  is drawn i.i.d. across periods. In each period, the long-run player observes  $\theta$  before taking an action  $a_1 \in A_1 \subset \mathbb{R}$ . The short-run player observes the current action  $a_1$  (but not  $\theta$ ) and the history of past actions (but not past states) and then takes an action  $r \in R$  in response. Assume  $\Theta$ ,  $A_1$ , and  $R$  are finite, the long-run player's payoff is given by  $(1 - \lambda)v(r) - \lambda w(a_1, \theta)$  for some functions  $v$  and  $w$  and some  $\lambda \in (0, 1)$ , and  $w$  is strictly submodular. Thus, the long-run player's preferences over the short-run player's action  $r$  are independent of the state  $\theta$ , and the parameter  $\lambda$  measures the weight on the long-run player's signaling cost  $w(a_1, \theta)$  relative to her payoff from the short-run player's action  $v(r)$ . Recall that submodularity of  $w$  is a standard assumption in signaling theory: for example, in Spence (1973),  $w(a_1, \theta) = a_1/\theta$ .<sup>10</sup>

<sup>9</sup>If one of  $\{w, z\}$  is positive and the other negative, the long-run player's payoff is supermodular in  $(\theta, a_1)$  for one  $a_2 \in \{T, N\}$  and submodular for the other. Our results do not cover this case. We also note that Proposition 2 is roughly consistent with Pei's (2020) results for the case where  $\theta$  is perfectly persistent: Pei shows that the long-run player can fail to secure her Stackelberg payoff in the submodular product choice game, but does secure it in the supermodular case (under an additional condition on the prior).

<sup>10</sup>Our analysis of signaling games remains valid for more general preferences of the form  $(1 - \lambda)v(a_1, r) - \lambda w(a_1, \theta)$ . That is,  $v$  can depend on  $a_1$  as well as  $r$ .

A strategy for the long-run player is now a mapping  $s_1 : \Theta \rightarrow \Delta(A_1)$ . Assume that the long-run player is committed to some strategy  $\hat{s}_1$  with probability  $\mu_0 > 0$  and is rational otherwise. Let  $\hat{v}_1$  be the long-run player's payoff when she takes  $\hat{s}_1$  and the short-run player takes his least favorable best response. Finally, say that a strategy  $s_1$  is *monotone* if any selection from  $\text{supp}(s_1(\theta))$  is monotone in  $\theta$ .

Just as in the deterrence game, standard results say little here. In particular, they cannot imply that a patient long-run player benefits from signaling, because any best response to the “babbling” strategy that takes actions  $a_1$  with the same marginal as  $\hat{s}$  but does so independently of  $\theta$  is a 0-confirmed best response to  $\hat{s}$ . However, we have the following.

**Proposition 3.** *If  $\hat{s}_1$  is monotone then  $\liminf_{\delta \rightarrow 1} \underline{U}_1(\delta) \geq \hat{v}_1$ .*

The logic is similar to the supermodular cases of Propositions 1 and 2: strict submodularity of  $w$  implies that the rational long-run player prefers  $\hat{s}_1$  to any strategy that induces the same marginal over actions, so establishing a reputation for the Stackelberg marginal suffices to establish a reputation for the Stackelberg strategy. In particular, if  $\hat{s}_1$  is the long-run player's Stackelberg strategy for the limiting payoffs where  $\lambda = 0$  (i.e., the Bayesian persuasion solution) and this strategy is monotone, then Proposition 3 implies that the long-run player secures a payoff only slightly below her Stackelberg payoff when  $\lambda$  is small. This result provides a foundation for the Bayesian persuasion solution when the sender's payoff is state-independent and the persuasion solution is monotone. In Section 6, we strengthen this result by characterizing all communication mechanisms that are monotone with respect to some order on states and receiver actions.

Finally, Corollary 3 in Section 5.3 implies a partial converse to Proposition 3: if all Bayesian persuasion solutions are non-monotone, then even the long-run player's best equilibrium payoff is bounded away from her commitment payoff for all  $\lambda > 0$ . The logic is that strict submodularity of  $w$  implies the long-run player must play a monotone strategy in any equilibrium and thus cannot establish a reputation for playing a non-monotone strategy.

### 3 Model

We consider repeated games where a possibly-committed long-run player faces a sequence of short-run opponents. To cover both applications where a short-run player moves first (like deterrence games) and those where a short-run player moves simultaneously with or after the long-run player (like trust or communication games), we consider repeated three-player

games, where one short-run player (player 0) moves first, and then the long-run player (player 1) and another short-run player (player 2) move simultaneously. In deterrence games, the second short-run player is absent. In trust games, the first short-run player is absent. In communication games, the first short-run player is Nature, and the second short-run player's action is a mapping from the long-run player's action to a set of possible responses.

We first describe the stage game, followed by the repeated game.

### 3.1 The Stage Game

There are three players,  $i \in \{0, 1, 2\}$ . Player 1 is the long-run player; players 0 and 2 are short-run players. Each player  $i$  has a finite action set  $A_i$  with generic element  $a_i$ . Each action  $a_i$  generates a signal in a finite set  $Y_i$ , with independent conditional probabilities  $\rho(\cdot|a_i) \in \Delta(Y_i)$ . The signals satisfy the following assumption.

**Assumption 1.**

- (1) *The signal  $y_0$  has a full support distribution:  $\rho(y_0|a_0) > 0$  for all  $y_0 \in Y_0, a_0 \in A_0$ .*
- (2) *The signal  $y_1$  statistically identifies the long-run player's action: the  $|A_1|$  vectors  $\rho(\cdot|a_1)_{a_1 \in A_1}$  are linearly independent in  $\mathbb{R}^{|Y_0|}$ .*

The stage game timing is as follows.

- (1) Player 0 takes an action  $a_0$ . This generates a signal  $y_0$ , drawn from  $\rho(\cdot|a_0)$ , which is observed by player 1 only.
- (2) Players 1 and 2 simultaneously take actions  $a_1$  and  $a_2$ . This generates signals  $y_1$  and  $y_2$ , drawn independently from  $\rho(\cdot|a_1)$  and  $\rho(\cdot|a_2)$ , respectively, and publicly observed by all players.

Thus, stage game strategies for players 0 and 2 are simply mixed actions  $\alpha_0 \in \Delta(A_0)$  and  $\alpha_2 \in \Delta(A_2)$ , respectively, while a stage game strategy for player 1 is a function  $s_1 : Y_0 \rightarrow \Delta(A_1)$ . Note that a strategy profile  $(\alpha_0, s_1, \alpha_2)$  induces a joint distribution  $\gamma(\alpha_0, s_1) \in \Delta(Y_0 \times A_1)$  (independent of  $\alpha_2$ ) over player 1's private signal  $y_0$  and action  $a_1$  according to

$$\gamma(\alpha_0, s_1)[y_0, a_1] = \sum_{a_0 \in A_0} \alpha_0(a_0) \rho(y_0|a_0) s_1(y_0)[a_1],$$

and it induces a joint distribution  $p(\alpha_0, s_1, \alpha_2) \in \Delta(Y_1 \times Y_2)$  over the public signals  $y_1$  and  $y_2$  according to

$$p(\alpha_0, s_1, \alpha_2)[y_1, y_2] = \sum_{a_0 \in A_0} \sum_{y_0 \in Y_0} \sum_{a_1 \in A_1} \sum_{a_2 \in A_2} \alpha_0(a_0) \rho(y_0|a_0) s_1(y_0)[a_1] \rho(y_1|a_1) \alpha_2(a_2) \rho(y_2|a_2).$$

We emphasize a key point: while the public signals  $(y_1, y_2)$  identify player 1's action  $a_1$  by Assumption 1, they do not identify her strategy  $s_1$  (whenever  $|Y_0| \geq 2$ ), because  $y_0$  is player 1's private information.

Throughout the paper, for any joint distribution  $\chi \in \Delta(X_1 \times X_2)$  over a product set  $X_1 \times X_2$ ,  $\pi_{X_i}(\chi)$  denotes its marginal on  $X_i$ . With slight abuse of notation, we also denote the marginal of  $\gamma(\alpha_0, s_1)$  over  $Y_0$  (which depends only on  $\alpha_0$ ) by  $\rho(\alpha_0) = \pi_{Y_0}(\gamma(\alpha_0, s_1))$ , and we denote its marginal over  $A_1$  by  $\phi(\alpha_0, s_1) = \pi_{A_1}(\gamma(\alpha_0, s_1))$ .

The players' payoff functions are given by  $u_0 : A_0 \times A_1 \rightarrow \mathbb{R}$  for player 0 and  $u_i : Y_0 \times A_1 \times A_2 \rightarrow \mathbb{R}$  for players  $i \in \{1, 2\}$ . Thus, player 0's payoff depends on his own action and player 1's action, while the payoffs of players 1 and 2 depend on their actions and the signal  $y_0$ . The assumption that player 0's payoff does not depend on player 2's action simplifies the analysis and is satisfied in our applications.<sup>11</sup> Finally, we write  $u_i(\alpha_0, s_1, \alpha_2)$  for player  $i$ 's expected payoff at stage-game strategy profile  $(\alpha_0, s_1, \alpha_2)$ , and we let  $\underline{u}_1 = \min_{a_0, s_1, a_2} u_1(a_0, s_1, a_2)$  and  $\bar{u}_1 = \max_{a_0, s_1, a_2} u_1(a_0, s_1, a_2)$ .

Deterrence games fit this framework by making  $A_2$  a singleton, which effectively drops player 2 from the model. Trust games fit by making  $A_0$  a singleton (i.e., making player 0 Nature). Communication games fit by making  $A_0$  a singleton; viewing  $\rho(y_0)$  as the prior distribution of a payoff-relevant state  $y_0$ ; letting  $\rho(y_1|a_1) = \mathbf{1}(\{y_1 = a_1\})$  (so  $a_1$  is perfectly monitored); viewing  $a_2$  as a mapping from  $a_1$  to a finite set of responses  $R$ ; and assuming that  $u_1$  and  $u_2$  depend on  $a_2$  only through the induced response  $a_2(a_1) \in R$ .

We conclude this subsection by adapting some definitions from Fudenberg and Levine (1992). For any strategy  $s_1$ , let  $B(s_1) \subset \Delta(A_0) \times \Delta(A_2)$  be the set of short-run player strategies  $(\alpha_0, \alpha_2)$  satisfying

$$\text{supp}(\alpha_0) \subset \underset{a_0 \in A_0}{\text{argmax}} u_0(a_0, s_1) \quad \text{and} \quad \text{supp}(\alpha_2) \subset \underset{a_2 \in A_2}{\text{argmax}} u_2(\alpha_0, s_1, a_2),$$

so that player 0 best responds to  $s_1$ , and player 2 best responds to  $\alpha_0$  and  $s_1$ . With this

---

<sup>11</sup>However, none of our results requires this assumption, with the exception of Theorem 2 in Appendix A.

notation, the long-run player's (*lower*) *Stackelberg payoff* is

$$v_1^* = \sup_{s_1 \in \Delta(A_1)^{Y_0}} \inf_{(\alpha_0, \alpha_2) \in B(s_1)} u_1(\alpha_0, s_1, \alpha_2).$$

We refer to a strategy that attains this supremum as a *Stackelberg strategy*. More generally, for any strategy  $s_1$ , we denote the corresponding lower commitment payoff by

$$V(s_1) = \inf_{(\alpha_0, \alpha_2) \in B(s_1)} u_1(\alpha_0, s_1, \alpha_2).^{12}$$

Finally, we employ the following definition<sup>13</sup>.

**Definition 1.** For any long-run player strategy  $s_1$  and any  $\eta \geq 0$ , a short-run player strategy  $(\alpha_0, \alpha_2) \in \Delta(A_0) \times \Delta(A_2)$  is an  $\eta$ -*confirmed best response* to  $s_1$  if there exists  $s'_1$  such that

- (1)  $(\alpha_0, \alpha_2) \in B(s'_1)$ , and
- (2)  $\|p(\alpha_0, s_1, \alpha_2) - p(\alpha_0, s'_1, \alpha_2)\| \leq \eta$ .

For any  $s'_1$  that satisfies these conditions, we say that it  $\eta$ -*confirms*  $(\alpha_0, \alpha_2)$  *against*  $s_1$ .

Let  $B_\eta(s_1)$  be the set of  $\eta$ -confirmed best responses to  $s_1$ . Note that  $B_1(s_1) \supset B_{\eta'}(s_1) \supset B_\eta(s_1) \supset B(s_1)$  for all  $\eta' \geq \eta \geq 0$ , where  $B_0(s_1) = B(s_1)$  if  $s_1$  is identified (which, again, is not the case in our model whenever  $|Y_0| \geq 2$ ), and  $B_1(s_1)$  is the set of all short-run player strategies that best respond to *some* long-run player strategy. In addition,  $B_\eta(s_1) \downarrow B_0(s_1)$  as  $\eta \rightarrow 0$ , by upper hemi-continuity of the short-run players' best-response correspondences. Finally, for any strategy  $s_1$ , we denote the lower commitment payoff when the short-run players take a 0-confirmed best response by

$$V_0(s_1) = \inf_{(\alpha_0, \alpha_2) \in B_0(s_1)} u_1(\alpha_0, s_1, \alpha_2).$$

Note that  $V(s_1) \geq V_0(s_1)$  for each  $s_1$ , since  $B(s_1) \subset B_0(s_1)$ .

## 3.2 The Repeated Game

The stage game is repeated in each period  $t = 0, 1, 2, \dots$ . Player 1 is a long-lived player with discount factor  $\delta \in (0, 1)$ , while players 0 and 2 are short-lived and take myopic best

---

<sup>12</sup>Recall that the *lower* (resp., *upper*) commitment payoff results from the *least favorable* (resp., *most favorable*) short-run best response.

<sup>13</sup>Here and throughout the paper,  $\|\cdot\|$  denotes the sup norm.

replies after observing only the public history of signals. A period- $t$  public history is denoted  $h^t = (y_{1,t'}, y_{2,t'})_{t'=0}^{t-1} \in (Y_1 \times Y_2)^t$ . Let  $H^t$  be the set of period- $t$  (public) histories,  $H = \bigcup_t H^t$  the set of all finite histories, and  $H^\infty = (Y_1 \times Y_2)^\infty$  the set of infinite histories. A repeated game strategy  $\sigma_i$  for player  $i$  maps public histories to stage game strategies: formally,  $\sigma_i$  is a function from  $H$  to  $\Delta(A_i)$  for  $i \in \{0, 2\}$ , and is a function from  $H$  to  $\Delta(A_1)^{Y_0}$  for  $i = 1$ .<sup>14</sup>

The long-run player's *type*, denoted  $\omega \in \Omega$ , is either *rational* ( $\omega = \omega_R$ ) or is one of a countable number of *commitment types* indexed by a (potentially mixed) stage game strategy  $s_1 \in (\Delta(A_1))^{Y_0}$ , where type  $\omega_{s_1}$  plays  $s_1$  in every period.<sup>15</sup> The type  $\omega$  is drawn according to a full-support prior  $\mu_0 \in \Delta(\Omega)$  at the start of the game and is perfectly persistent. We study  $\underline{U}_1(\delta)$  and  $\bar{U}_1(\delta)$ , the infimum and supremum of the long-run player's payoff in any Nash equilibrium  $(\sigma_0^*, \sigma_1^*, \sigma_2^*)$  of this incomplete-information repeated game. (Here and throughout the paper,  $\sigma_1^*$  denotes player 1's strategy unconditional on her type, with the rational type's strategy denoted  $\sigma_1^*(\omega_R)$ .)

The key prior result in this context is the following.

**Theorem 0** (Fudenberg and Levine, 1992). *If  $\omega_{s_1^*} \in \Omega$ , then*

$$\liminf_{\delta \rightarrow 1} \underline{U}_1(\delta) \geq V_0(s_1^*).$$

Our main contribution is providing conditions on  $s_1^*$  under which this bound can be improved to  $V(s_1^*)$ . As we saw in Section 2, this improvement can mean the difference between securing the minmax payoff and the Stackelberg payoff.

### 3.3 Behavioral Confounding

We make use of the following definition.

**Definition 2.** A strategy  $s_1$  is *not behaviorally confounded* if, for any  $\omega_{s'_1} \in \Omega$  such that  $s'_1 \neq s_1$  and any  $(\alpha_0, \alpha_2) \in B_1(s_1)$ , we have  $p(\alpha_0, s_1, \alpha_2) \neq p(\alpha_0, s'_1, \alpha_2)$ .

Thus, a strategy is not behaviorally confounded if the public signals distinguish it from any other commitment type, whenever the short-run players take actions that best respond

---

<sup>14</sup>In principle, the long-run player could condition on her past private signal realizations and actions in addition to the public history, but allowing this does not affect the set of equilibrium payoffs, because short-run player strategies are measurable with respect to the public history and long-run player payoffs are independent of past signals and actions. We also note that the signal  $y_2$  is irrelevant for the analysis and is included only for generality.

<sup>15</sup>We thus assume that there is only one rational type (unlike Fudenberg and Levine (1992), who allow multiple rational types). We discuss the extension to multiple rational types in Section 7.

to some long-run player strategy. Note that the definition allows the possibility that  $s_1$  is indistinguishable from a mixture of two commitment types  $\omega_{s'_1}, \omega_{s'_2} \in \Omega$ .<sup>16</sup> Note also that if there is only one commitment type  $\omega_{s_1}$  then  $s_1$  is not behaviorally confounded.

Our results in the body of the paper assume that desired commitment strategies are not behaviorally confounded. In games without a player 0 (like trust and communication games), this is fairly innocuous, as  $\alpha_0$  is exogenous, so the identification condition  $p(\alpha_0, s_1, \alpha_2) \neq p(\alpha_0, s'_1, \alpha_2)$  for all  $\alpha_2 \in B_0(s_1)$  holds for generic  $s_1 \neq s'_1$ . In games with a player 0, it is much more restrictive, because  $\alpha_0$  is endogenous, so the identification condition need not hold generically. For example, in the deterrence game in Section 2.1, the pure Stackelberg strategy  $(A, F)$  is not behaviorally confounded iff each other type  $\omega_{s'_1} \in \Omega$  satisfies either  $ps'_1(A|c) + (1-p)s'_1(A|d) > p$  and  $(1-p)s'_1(A|c) + ps'_1(A|d) > 1-p$ , or  $ps'_1(A|c) + (1-p)s'_1(A|d) < p$  and  $(1-p)s'_1(A|c) + ps'_1(A|d) < 1-p$ .

Nonetheless, in Appendix A we show that our results extend even if the Stackelberg type is behaviorally confounded, so long as it has sufficiently high prior weight relative to any behavioral confound that induces an  $\eta$ -confirmed best response that is not also a best response to  $s_1^*$ . In fact, we establish a payoff lower bound that linearly interpolates between  $V(s_1^*)$  and  $V_0(s_1^*)$  as a function of this prior weight.

## 4 The Commitment Payoff Theorem

This section presents the confound-defeating property and our main result: a patient long-run player can secure the commitment payoff  $V(s_1^*)$  corresponding to any strategy  $s_1^* \in \Omega$  that is confounding-defeating and not behaviorally confounded.

### 4.1 The Confound-Defeating Property

We first introduce the confound-defeating property. We give two equivalent definitions of confound-defeatingness. The first definition is more useful for proving our main result. The second definition is more elegant and is easier to characterize in applications (as we do in Sections 5 and 6). The second definition is stated in optimal transport terms: for any two distributions  $\rho \in \Delta(Y_0)$  and  $\phi \in \Delta(A_1)$ , and any strategy for player 2  $\alpha_2 \in \Delta(A_2)$ , define

---

<sup>16</sup>Ruling out this possibility would simplify the proof of Lemma 3 and would also let us replace  $B_1(s_1)$  with  $B_0(s_1)$  in Definition 2. That is, all of our results go through with the alternative definition that  $s_1$  is not behaviorally confounded if, for any  $(\alpha_0, \alpha_2) \in B_0(s_1)$ ,  $p(\alpha_0, s_1, \alpha_2)$  lies outside the convex hull of the set  $\bigcup_{s'_1 \neq s_1: \omega_{s'_1} \in \Omega} p(\alpha_0, s'_1, \alpha_2)$ .



the optimal transport problem

$$\text{OT}(\rho, \phi; \alpha_2) : \max_{\gamma \in \Delta(Y_0 \times A_1)} \int u_1(y_0, a_1, \alpha_2) d\gamma \quad \text{s.t.} \quad \pi_{Y_0}(\gamma) = \rho \text{ and } \pi_{A_1}(\gamma) = \phi.$$

**Definition 3.** Strategy  $s_1^*$  is *confound-defeating* if it satisfies one of the following conditions:

- (1) For all  $\varepsilon > 0$ , there exists  $\eta > 0$  such that for any  $(\alpha_0, \alpha_2) \in B_\eta(s_1^*)$  and any  $s_1'$  satisfying  $\|s_1' - s_1^*\| > \varepsilon$  but  $\|p(\alpha_0, s_1', \alpha_2) - p(\alpha_0, s_1^*, \alpha_2)\| < \eta$ , there exists  $\tilde{s}_1$  satisfying  $p(\alpha_0, \tilde{s}_1, \alpha_2) = p(\alpha_0, s_1', \alpha_2)$  and  $u_1(\alpha_0, \tilde{s}_1, \alpha_2) > u_1(\alpha_0, s_1', \alpha_2)$ .
- (2) For any  $(\alpha_0, \alpha_2) \in B_0(s_1^*)$ ,  $\gamma(\alpha_0, s_1^*)$  is the unique solution to  $\text{OT}(\rho(\alpha_0), \phi(\alpha_0, s_1^*); \alpha_2)$ .

The first definition says that a strategy  $s_1^*$  is confound-defeating if any strategy  $s_1'$  that is a possible confound—in that it differs significantly from  $s_1^*$  but induces a similar marginal over signals against some  $\eta$ -confirmed best response—is undetectably dominated—in that the long-run player is strictly better-off under a different strategy  $\tilde{s}$  that induces the same marginal. The second definition says that  $s_1^*$  itself undetectably dominates any strategy  $s_1'$  that induces the same marginal over signals against some 0-confirmed best response. The two definitions are equivalent, as we now show.<sup>17</sup>

**Proposition 4.** *The two definitions of confound-defeatingness are equivalent.*

*Proof.* If the second definition fails, there exist  $(\alpha_0, \alpha_2) \in B_0(s_1^*)$ ,  $\varepsilon > 0$ , and a strategy  $s_1'$  satisfying  $\|s_1' - s_1^*\| > \varepsilon$  such that  $\gamma(\alpha_0, s_1')$  solves  $\text{OT}(\rho(\alpha_0), \phi(\alpha_0, s_1^*); \alpha_2)$ . Since  $B_0(s_1^*) \subset B_\eta(s_1^*)$  for all  $\eta > 0$ , this implies that the first definition fails.

Conversely, if the first definition fails, there exists  $\varepsilon > 0$  such that for all  $\eta > 0$ , there exist  $s_1^\eta$  and  $(\alpha_0^\eta, \alpha_2^\eta) \in B_\eta(s_1^*)$  where  $\|s_1^\eta - s_1^*\| > \varepsilon$ ,  $\|p(\alpha_0^\eta, s_1^\eta, \alpha_2^\eta) - p(\alpha_0^\eta, s_1^*, \alpha_2^\eta)\| < \eta$ , and  $s_1^\eta$  is not undetectably dominated:  $u_1(\alpha_0^\eta, s_1^\eta, \alpha_2^\eta) \geq u_1(\alpha_0, \tilde{s}_1, \alpha_2^\eta)$  for all  $\tilde{s}_1$  such that  $p(\alpha_0^\eta, s_1^\eta, \alpha_2^\eta) = p(\alpha_0^\eta, \tilde{s}_1, \alpha_2^\eta)$ . Since  $a_1$  is identified,  $p(\alpha_0^\eta, s_1^\eta, \alpha_2^\eta) = p(\alpha_0^\eta, \tilde{s}_1, \alpha_2^\eta)$  implies  $\phi(\alpha_0^\eta, s_1^\eta) = \phi(\alpha_0^\eta, \tilde{s}_1)$ , and hence  $s_1^\eta$  solves  $\text{OT}(\rho(\alpha_0^\eta), \phi(\alpha_0^\eta, s_1^\eta); \alpha_2^\eta)$ . Now, since  $B_\eta(s_1) \downarrow B_0(s_1)$ ,  $\Delta(A_0) \times (\Delta(A_1))^{Y_0} \times \Delta(A_2)$  is compact, and  $\text{OT}(\rho(\alpha_0), \phi(\alpha_0, s_1); \alpha_2)$  is jointly upper hemi-continuous in  $(\alpha_0, s_1, \alpha_2)$ , passing to the limit yields  $s_1^0$  and  $(\alpha_0, \alpha_2) \in B_0(s_1^*)$  such that  $\|s_1^0 - s_1^*\| \geq \varepsilon$  and  $s_1^0$  solves  $\text{OT}(\rho(\alpha_0), \phi(\alpha_0, s_1^*); \alpha_2)$ . Thus, the second definition fails.  $\square$

<sup>17</sup>The converse direction of Proposition 4 is the only point where we use Assumption 1(2) (player 1's action is identified). Thus, without identification, the first definition of confound-defeatingness still gives a commitment payoff theorem, but it is more difficult to check and needs not reduce to monotonicity in one-dimensional supermodular games.

The following lemma gives the key implication of confound-defeatingness: in any Nash equilibrium, at any history where the marginal over signals is close to that induced by a confound-defeating strategy  $s_1^*$ —both unconditionally and conditional on the event the long-run player is rational—the rational long-run player must play a strategy close to  $s_1^*$  itself. Here and throughout the paper, given a repeated game strategy profile  $(\sigma_0, \sigma_1, \sigma_2)$  and a period- $t$  history  $h^t$ , we abbreviate  $p(\sigma_0(h^t), \sigma_1(h^t), \sigma_2(h^t))$  to  $p(\sigma_0, \sigma_1, \sigma_2|h^t)$ .

**Lemma 1.** *Fix a Nash equilibrium  $(\sigma_0^*, \sigma_1^*, \sigma_2^*)$  and suppose that  $s_1^*$  is confound-defeating. Then for all  $\varepsilon > 0$ , there exists  $\eta > 0$  such that, for any history  $h^t$  where*

$$(1) \quad \|p(\sigma_0^*, \sigma_1^*, \sigma_2^*|h^t) - p(\sigma_0^*, s_1^*, \sigma_2^*|h^t)\| < \eta, \text{ and}$$

$$(2) \quad \|p(\sigma_0^*, \sigma_1^*(\omega_R), \sigma_2^*|h^t) - p(\sigma_0^*, s_1^*, \sigma_2^*|h^t)\| < \eta,$$

we have  $\|\sigma_1^*(h^t, \omega_R) - s_1^*\| \leq \varepsilon$ .

*Proof.* Suppose not, so there exists a history  $h^t$  where conditions (1) and (2) hold, but  $\|\sigma_1^*(h^t, \omega_R) - s_1^*\| > \varepsilon$ . Condition (1) and the fact  $(\sigma_0^*, \sigma_1^*, \sigma_2^*)$  is an equilibrium imply that  $(\sigma_0^*(h^t), \sigma_2^*(h^t))$  is an  $\eta$ -confirmed best reply to  $s_1^*$ , as  $\sigma_1^*(h^t)$   $\eta$ -confirms it against  $s_1^*$ . Hence, condition (2) along with  $\|\sigma_1^*(h^t, \omega_R) - s_1^*\| > \varepsilon$  and confound-defeatingness imply that there exists some strategy  $\tilde{s}_1$  such that  $p(\sigma_0^*, \tilde{s}_1, \sigma_2^*|h^t) = p(\sigma_0^*, \sigma_1^*(\omega_R), \sigma_2^*|h^t)$  and  $u_1(\sigma_0^*, \tilde{s}_1, \sigma_2^*|h^t) > u_1(\sigma_0^*, \sigma_1^*(\omega_R), \sigma_2^*|h^t)$ . But this implies that if the long-run player deviates from  $\sigma_1^*(h^t, \omega_R)$  to  $\tilde{s}_1$  at  $h^t$ , her continuation payoff is unchanged while her stage game payoff increases. So, this deviation is profitable, contradicting the assumption that  $(\sigma_0^*, \sigma_1^*, \sigma_2^*)$  is an equilibrium.  $\square$

## 4.2 Payoff Lower Bound

We are now prepared to state our main result.

**Theorem 1.** *If  $\omega_{s_1^*} \in \Omega$  and  $s_1^*$  is confound-defeating and not behaviorally confounded, then*

$$\liminf_{\delta \rightarrow 1} \underline{U}_1(\delta) \geq V(s_1^*).$$

In particular, if  $s_1^*$  is a Stackelberg strategy, Theorem 1 implies that a patient long-run player can secure her Stackelberg payoff  $v_1^*$ .

The logic of Theorem 1 is as follows. Fix any equilibrium, and suppose player 1 deviates by taking  $s_1^*$  in every period. By standard arguments (Fudenberg and Levine, 1992; Sorin, 1999; Gossner, 2011), the short-run players eventually come to expect the signal distribution

$p(\sigma_0^*(h^t), s_1^*, \sigma_2^*(h^t))$  at public history  $h^t$ . Since  $s_1^*$  is confound-defeating, by Lemma 1, the short-run players additionally come to expect that if player 1 is rational, she plays a stage game strategy close to  $s_1^*$ . Since  $s_1^*$  is not behaviorally confounded, the short-run players also eventually learn that player 1 is not some commitment type other than  $\omega_{s_1^*}$ .<sup>18</sup> In total, the short-run players come to believe that player 1 is either the commitment type  $\omega_{s_1^*}$  or is rational and plays a stage game strategy close to  $s_1^*$ . This leads the short-run players to best respond to  $s_1^*$ , which ensures the long-run player a payoff of at least  $V(s_1^*)$ .

*Proof.* Fix any  $\varepsilon > 0$ . We show that there exists  $\bar{\delta} < 1$  such that for all  $\delta > \bar{\delta}$ , we have  $\underline{U}_1(\delta) \geq V(s_1^*) - \varepsilon$ . To do so, we fix any Nash equilibrium  $(\sigma_0^*, \sigma_1^*, \sigma_2^*)$  and show that player 1's payoff from deviating by always taking  $s_1^*$  is at least  $V(s_1^*) - \varepsilon$ . Let  $\mathbb{P}$  denote the equilibrium probability measure over infinite histories  $H^\infty$ , and let  $\mathbb{Q}$  denote the corresponding measure under this deviation.

For any  $\eta > 0$ , define the set of period- $t$  histories where the equilibrium signal distribution is within  $\eta$  of that under the deviation by

$$H_\eta^t = \left\{ h^t : \left| p(\sigma_0^*, s_1^*, \sigma_2^* | h^t) - p(\sigma_0^*, \sigma_1^*, \sigma_2^* | h^t) \right| \leq \eta \right\}.$$

We first recall a standard bound (essentially due to Gossner (2011)) on the expected number of periods  $t$  where  $h^t \notin H_\eta^t$ . We include a proof in Appendix B.1.

**Lemma 2.** *We have*

$$\mathbb{E}^{\mathbb{Q}} \left[ \# \left\{ t : h^t \notin H_\eta^t \right\} \right] < \bar{T}(\eta, \mu_0) := -\frac{2 \log \mu_0(\omega_{s_1^*})}{\eta^2}.$$

Next, for any  $\zeta > 0$ , denote the set of beliefs with at most  $\zeta$  weight on commitment types other than  $s_1^*$  by

$$M_\zeta = \left\{ \mu \in \Delta(\Omega) : \mu \left( \{ \omega_R, \omega_{s_1^*} \} \right) \geq 1 - \zeta \right\}.$$

The next lemma shows that beliefs under  $\mathbb{Q}$  concentrate on  $M_0$  with high probability, uniformly in  $\delta$ . The proof, which relies on the martingale convergence theorem and the assumption  $y_0$  has a full support distribution for any  $a_0$ , is deferred to Appendix B.2. In what follows, given a history  $h$ ,  $\mu_t(\cdot | h) \in \Delta(\Omega)$  denotes the posterior belief over  $\Omega$  conditional on the period  $t$  truncation  $h^t$  of  $h$ . (We also write  $\mu_t(\cdot | h^t)$  for beliefs at period- $t$  history  $h^t$ .)

---

<sup>18</sup>This step is not trivial, for example because we allow  $s_1^*$  to be indistinguishable from a mixture of commitment types.

**Lemma 3.** For all  $\zeta > 0$ , there exists a set of infinite histories  $G(\zeta) \subset H^\infty$  satisfying  $\mathbb{Q}(G(\zeta)) > 1 - \zeta$  and a period  $\hat{T}(\zeta)$  (independent of  $\delta$  and the choice of equilibrium) such that, for any  $h \in G(\zeta)$  and any  $t \geq \hat{T}(\zeta)$ , we have  $\mu_t(\cdot|h) \in M_\zeta$ .

Now, for any  $\xi > 0$ , we say that a short-run player strategy  $(\alpha_0, \alpha_2)$  is a  $\xi$ -close best response to  $s_1^*$  (denoted  $(\alpha_0, \alpha_2) \in \hat{B}_\xi(s_1^*)$ ) if  $(\alpha_0, \alpha_2) \in B(s_1)$  for some  $s_1$  such that  $\|s_1 - s_1^*\| < \xi$ . Since  $\hat{B}_\xi(s_1^*)$  is upper hemi-continuous and  $\hat{B}_0(s_1^*) = B(s_1^*)$ , we have

$$\liminf_{\xi \rightarrow 0} \inf_{(\alpha_0, \alpha_2) \in \hat{B}_\xi(s_1^*)} u_1(\alpha_0, s_1^*, \alpha_2) = \inf_{(\alpha_0, \alpha_2) \in B(s_1^*)} u_1(\alpha_0, s_1^*, \alpha_2) = V(s_1^*).$$

The next lemma shows that if the short-run players expect the marginal induced by  $s_1^*$  and believe that player 1 is either the  $s_1^*$  commitment type or rational, they will take a  $\xi$ -close best response to  $s_1^*$ . Its proof (deferred to Appendix B.3) crucially relies on Lemma 1.

**Lemma 4.** There exist strictly positive functions  $\zeta(\eta)$  and  $\xi(\eta)$ , satisfying  $\lim_{\eta \rightarrow 0} \zeta(\eta) = \lim_{\eta \rightarrow 0} \xi(\eta) = 0$ , such that if  $h^t \in H_\eta^t$  and  $\mu_t(\cdot|h^t) \in M_{\zeta(\eta)}$  then  $(\sigma_0^*(h^t), \sigma_2^*(h^t)) \in \hat{B}_{\xi(\eta)}(s_1^*)$ .

We can now complete the proof of Theorem 1. By Lemmas 2 and 3, conditional on the (at least) probability  $1 - \zeta(\eta)$  event that  $h \in G(\zeta(\eta))$ , the expected number of periods where either  $h^t \notin H_\eta^t$  or  $\mu_t(\cdot|h^t) \notin M_{\zeta(\eta)}$  is at most  $\bar{T}(\eta, \mu_0) + \hat{T}(\zeta(\eta))$ . By Lemma 4, in any period where  $h^t \in H_\eta^t$  and  $\mu_t(\cdot|h^t) \in M_{\zeta(\eta)}$ , we have  $(\sigma_0^*(h^t), \sigma_2^*(h^t)) \in \hat{B}_{\xi(\eta)}(s_1^*)$ , and hence, for any sufficiently small  $\eta$ ,

$$u_1(\sigma_0^*(h^t), s_1^*, \sigma_2^*(h^t)) \geq \liminf_{\eta \rightarrow 0} \inf_{(\alpha_0, \alpha_2) \in \hat{B}_{\xi(\eta)}(s_1^*)} u_1(\alpha_0, s_1^*, \alpha_2) - \frac{\varepsilon}{3} = V(s_1^*) - \frac{\varepsilon}{3}.$$

Front-loading the expected periods where  $h^t \notin H_\eta^t$  or  $\mu_t(\cdot|h^t) \notin M_{\zeta(\eta)}$  (and positing the minimum payoff of  $\underline{u}_1$  in these periods) gives a lower bound for player 1's payoff of

$$\left(1 - \delta^{\bar{T}(\eta, \mu_0) + \hat{T}(\zeta(\eta))}\right) \underline{u}_1 + \delta^{\bar{T}(\eta, \mu_0) + \hat{T}(\zeta(\eta))} \left( (1 - \zeta(\eta)) \bar{V}(s_1^*) + \zeta(\eta) \underline{u}_1 - \frac{\varepsilon}{3} \right)$$

As  $\delta \rightarrow 1$ , this lower bound converges to

$$(1 - \zeta(\eta)) \bar{V}(s_1^*) + \zeta(\eta) \underline{u}_1 - \frac{\varepsilon}{3}.$$

By continuity, at a cost of  $\varepsilon/3$ , this bound stays valid for all large enough  $\delta < 1$ . Finally, taking  $\eta$  to be small enough so  $\zeta(\eta) |\underline{u}_1 - \bar{V}(s_1^*)| < \varepsilon/3$  gives the desired bound of  $V(s_1^*) - \varepsilon$ .  $\square$

In Appendix A, we generalize Theorem 1 to the case where  $s_1^*$  is behaviorally confounded. There, we show that if  $\omega_{s_1^*} \in \Omega$  and  $s_1^*$  is confound-defeating, then

$$\liminf_{\delta \rightarrow 1} \underline{U}_1(\delta) \geq \beta(s_1^*; \mu_0)V(s_1^*) + (1 - \beta(s_1^*; \mu_0))V_0(s_1^*),$$

for some function  $\beta(s_1^*; \mu_0)$  (constructed explicitly in Appendix A), where  $\beta(s_1^*; \mu_0) = 1$  if  $s_1^*$  is not behaviorally confounded, and  $\beta(s_1^*; \mu_0) \rightarrow 1$  as the prior weight on  $s_1^*$  relative to that on confounding commitment types that induce best responses outside  $B(s_1^*)$  converges to 1.

## 5 Cyclical Monotonicity and Supermodular Games

This section characterizes confound-defeating strategies in terms of the support of the joint distributions over  $Y_0 \times A_1$  they induce, and uses this characterization to develop a partial converse to Theorem 1. We then explain the implications of these results for one-dimensional supermodular games. These include the proofs of the claimed results for our motivating examples, Propositions 1–3.

### 5.1 Cyclical Monotonicity

Our characterization is based on the following strict version of the familiar notion of cyclical monotonicity (Rochet, 1987). The definition and subsequent characterization are elementary, but we are not aware of a reference.<sup>19</sup>

**Definition 4.** Fix finite sets  $X, Y$  and a function  $u : X \times Y \rightarrow \mathbb{R}$ . A set  $S \subset X \times Y$  is *strictly  $u$ -cyclically monotone* if for any finite collection of pairs  $\{(x_i, y_i)_{i=1}^N\} \subset S$  such that  $\{(x_i, y_i)_{i=1}^N\} \neq \{(x_i, y_{i+1})_{i=1}^N\}$  (with convention  $y_{N+1} = y_1$ ),

$$\sum_{i=1}^N u(x_i, y_i) > \sum_{i=1}^N u(x_i, y_{i+1}).$$

**Proposition 5.** *A joint distribution  $\gamma \in \Delta(X \times Y)$  satisfying  $\pi_X(\gamma) = \rho$  and  $\pi_Y(\gamma) = \phi$  is the unique solution to the optimal transport problem*

$$OT(\rho, \phi) : \max_{\gamma' \in \Delta(X \times Y)} \int u(x, y) d\gamma' \quad s.t. \quad \pi_X(\gamma') = \rho \text{ and } \pi_Y(\gamma') = \phi$$

---

<sup>19</sup>The closest argument we are aware of is the proof of Lemma 2 of Ball and Kattwinkel (2024). We thank Ian Ball for pointing out this connection and suggesting the proof of Proposition 5.

if and only if its support  $\text{supp}(\gamma) \subset X \times Y$  is strictly  $u$ -cyclically monotone.

We apply Proposition 5 to our setting with  $X = Y_0$  and  $Y = A_1$  and use the OT definition of confound-defeatingness to characterize confound-defeatingness in terms of the support of  $\gamma(\alpha_0, s_1^*)$ . Since we have assumed that the distribution of  $y_0$  has full support, this set depends only on  $s_1^*$ , so we write it as

$$\text{supp}(s_1^*) := \{(y_0, a_1) \in Y_0 \times A_1 : a_1 \in \text{supp}(s_1^*(y_0))\}.$$

In addition, letting  $u_1(\cdot, \alpha_2)$  denote player 1's utility  $u_1(y_0, a_1, \alpha_2)$  as a function of  $(y_0, a_1)$  for a fixed player 2 strategy  $\alpha_2$ , we say that  $u_1$  is *strictly cyclically separable* if whenever a set  $S \subset Y_0 \times A_1$  is strictly  $u_1(\cdot, \alpha_2)$ -cyclically monotone for some  $\alpha_2$ , it is strictly  $u_1(\cdot, \alpha_2)$ -cyclically monotone for all  $\alpha_2$ . In this case, the strict  $u_1$ -cyclical monotonicity of a set  $S \subset Y_0 \times A_1$  is well-defined independent of  $\alpha_2$ .<sup>20</sup> Finally, we say that a strategy  $s_1^*$  is *strictly  $u_1(\cdot, \alpha_2)$  (resp.,  $u_1$ )-cyclically monotone* if  $\text{supp}(s_1^*) \subset Y_0 \times A_1$  is strictly  $u_1(\cdot, \alpha_2)$  (resp.,  $u_1$ )-cyclically monotone. We obtain the following characterization.

**Corollary 1.** *A strategy  $s_1^*$  is confound-defeating if and only if it is strictly  $u_1(\cdot, \alpha_2)$ -cyclically monotone for all  $(\alpha_0, \alpha_2) \in B_0(s_1^*)$ .*

*Moreover, if  $u_1$  is strictly cyclically separable, a strategy  $s_1^*$  is confound-defeating if and only if it is strictly  $u_1$ -cyclically monotone.*

Together with Theorem 1, we obtain the following corollary, where (when  $u_1$  is cyclically separable) we denote the long-run player's greatest lower commitment payoff from any strictly  $u_1$ -cyclically monotone strategy by

$$\underline{v}_1^{\text{CM}} := \sup_{s_1: \omega_{s_1} \in \Omega \text{ and } s_1 \text{ is strictly } u_1\text{-cyclically monotone and not behaviorally confounded}} \min_{(\alpha_0, \alpha_2) \in B(s_1)} u_1(\alpha_0, s_1, \alpha_2).$$

**Corollary 2.** *If  $u_1$  is strictly cyclically separable,  $\omega_{s_1} \in \Omega$ , and  $s_1$  is strictly  $u_1$ -cyclically monotone and not behaviorally confounded, then*

$$\liminf_{\delta \rightarrow 1} \underline{U}_1(\delta) \geq V(s_1^*).$$

*In particular,*

$$\liminf_{\delta \rightarrow 1} \underline{U}_1(\delta) \geq \underline{v}_1^{\text{CM}}.$$

---

<sup>20</sup>Note that this case always applies in games without a player 2, such as deterrence games.

## 5.2 Payoff Upper Bound

We now give a partial converse to Corollary 2: if the long-run player is rational with high probability, her payoff is bounded above by her greatest upper commitment payoff from any  $u_1$ -cyclically monotone strategy, defined by

$$\bar{v}_1^{\text{CM}} := \sup_{s_1 : s_1 \text{ is } u_1\text{-cyclically monotone}} \max_{(\alpha_0, \alpha_2) \in B(s_1)} u_1(\alpha_0, s_1, \alpha_2),$$

where  $s_1$  is  $u_1$ -cyclically monotone if  $\text{supp}(s_1)$  satisfies the usual definition of  $u_1$ -cyclically monotonicity (i.e., Definition 4 with the strict inequality replaced by a weak one). The idea is that if a strategy  $s_1$  is not  $u_1$ -cyclically monotone strategy, then the rational long-run player has a profitable and undetectable deviation from  $s_1^*$ , so the short-run players cannot expect to face a strategy close to  $s_1$  with high probability when the long-run player is rational with high probability.

In the following statement,  $u_1$  is *cyclically separable* if whenever a set  $S \subset Y_0 \times A_1$  is  $u_1(\cdot, \alpha_2)$ -cyclically monotone for some  $\alpha_2$ , it is  $u_1(\cdot, \alpha_2)$ -cyclically monotone for all  $\alpha_2$ .

**Proposition 6.** *Suppose  $u_1$  is cyclically separable. Then for all  $\varepsilon > 0$ , there exists  $\kappa > 0$  such that, for any prior  $\mu_0$  satisfying  $\mu_0(\omega_R) > 1 - \kappa$  and any  $\delta < 1$ ,*

$$\bar{U}_1(\delta) < \bar{v}_1^{\text{CM}} + \varepsilon.$$

The key step in the proof of Proposition 6 is the following lemma, which says that a rational long-run player with a cyclically separable utility must play a cyclically monotone stage game strategy at every history in any repeated game Nash equilibrium. We record the lemma and its (short) proof, as it applies equally to any repeated game, with or without multiple long-run players and incomplete information.

**Lemma 5.** *For any Nash equilibrium  $(\sigma_0^*, \sigma_1^*, \sigma_2^*)$  and any history  $h^t$ ,  $\sigma_1^*(h^t, \omega_R)$  is  $u_1$ -cyclically monotone.*

*Proof.* Note that  $\sigma_1^*(h^t, \omega_R)$  must solve

$$\text{OT}(\sigma_0^*(h^t), \phi(\sigma_0^*(h^t), \sigma_1^*(h^t, \omega_R)), \sigma_2^*(h^t)).$$

This holds because, otherwise, there exists a strategy  $s_1$  that gives player 1 a strictly higher payoff at history  $h^t$  than  $\sigma_1^*(h^t, \omega_R)$  does, but gives the same signal distribution, and hence

the same continuation payoff. By a standard optimal transport result (e.g., Theorem 1.38 in Santambrogio, 2015), this implies that  $\text{supp}(\gamma(\sigma_0^*(h^t), \sigma_1^*(h^t, \omega_R)))$  is  $u_1(\cdot, \sigma_2^*(h^t))$ -cyclically monotone. Hence, by cyclical separability,  $\sigma_1^*(h^t, \omega_R)$  is  $u_1$ -cyclically monotone.  $\square$

Combining Corollary 2 and Proposition 6 gives a fairly tight characterization of a patient long-run player's payoff when  $u_1$  is cyclically and strictly cyclically separable: it is at least her lower commitment payoff from any non-behaviorally confounded, strictly  $u_1$ -cyclically monotone commitment type strategy, and at most her greatest upper commitment payoff from any  $u_1$ -cyclically monotone strategy.

### 5.3 One-Dimensional Supermodular Games

We now show how our results apply in games where the long-run player's utility is supermodular in a one-dimensional signal and action. This class of games includes the deterrence, trust, and signaling examples of Section 2.

The relevant supermodularity notion is strict supermodularity in  $(y_0, a_1)$  for all  $a_2$ .

**Definition 5.** The long-run player's payoff  $u_1$  is *strictly supermodular* if there exist total orders  $(\succ_{Y_0}, \succ_{A_1})$  such that

$$u_1(y_0, a_1, a_2) - u_1(y_0, a'_1, a_2) > u_1(y'_0, a_1, a_2) - u_1(y'_0, a'_1, a_2) \quad \text{for all } y_0 \succ y'_0, a_1 \succ a'_1, a_2.$$

We say that a (possibly mixed) strategy  $s_1 : Y_0 \rightarrow \Delta(A_1)$  is monotone if any selection from its graph is monotone. Formally,

**Definition 6.** A long-run player strategy  $s_1$  is *monotone* if, for any  $y_0 \succ y'_0$ ,  $a_1 \in \text{supp}(s_1(y_0))$ , and  $a'_1 \in \text{supp}(s_1(y'_0))$ , we have  $a_1 \succ a'_1$ .

The following is the key result of this subsection.

**Proposition 7.** *Suppose  $u_1$  is strictly supermodular. Then, for any strategy  $s_1^*$ , the following are equivalent:*

- (1)  $s_1^*$  is confound-defeating.
- (2)  $s_1^*$  is monotone.
- (3)  $s_1^*$  is  $u_1$ -cyclically monotone.



*Proof.* The key step is the following standard result from optimal transport.

**Lemma 6.** *Suppose  $u_1$  is strictly supermodular. Then, for any  $(\alpha_0, \alpha_2)$ ,  $s_1^*$  is monotone iff  $\gamma(\alpha_0, s_1^*)$  is the unique solution to  $OT(\rho(\alpha_0), \phi(\alpha_0, s_1^*); \alpha_2)$ .*

*Proof.* By Lemma 2.8 in Santambrogio (2015), if  $s_1^*$  is monotone then  $\gamma(\alpha_0, s_1^*)$  is the unique co-monotone transport plan between  $\rho(\alpha_0)$  and  $\phi(\alpha_0, s_1^*)$ —that is, the unique joint distribution  $\gamma \in \Delta(Y_0 \times A_1)$  with marginals  $\rho(\alpha_0)$  and  $\phi(\alpha_0, s_1^*)$  such that, according to  $\gamma$ ,  $y_0$  and  $a_1$  are co-monotone random variables. Conversely, since  $\rho(\alpha_0)$  has full support, if  $s_1^*$  is not monotone then  $\gamma(\alpha_0, s_1^*)$  is not co-monotone. Finally, by Theorem 2.9 and Exercise 10 in Santambrogio (2015), when  $u_1$  is strictly supermodular, the co-monotone transport plan is the unique solution to  $OT(\rho(\alpha_0), \phi(\alpha_0, s_1^*); \alpha_2)$ .  $\square$

By Lemma 6, if  $s_1^*$  is confound-defeating then  $\gamma(\alpha_0, s_1^*)$  is the unique solution to  $OT(\rho(\alpha_0), \phi(\alpha_0, s_1^*); \alpha_2)$  for any  $(\alpha_0, \alpha_2) \in B_0(s_1^*)$ , and hence is monotone; and, conversely, if  $s_1^*$  is monotone then it is the unique solution to  $OT(\rho(\alpha_0), \phi(\alpha_0, s_1^*); \alpha_2)$  for any  $(\alpha_0, \alpha_2)$ , and hence is confound-defeating. Moreover,  $s_1^*$  is monotone if and only if  $\gamma(\alpha_0, s_1^*)$  is co-monotone, as shown in the proof of Lemma 6, and  $\gamma(\alpha_0, s_1^*)$  is co-monotone if and only if it is  $u_1$ -cyclically monotone, by Lemma 1 of Lin and Liu (2024) (see also Proposition 1 of Rochet (1987)). This establishes the desired three-way equivalence.  $\square$

Denote the long-run player's lower commitment payoff from any non-behaviorally confounded, monotone commitment type strategy and her upper commitment payoff from any monotone strategy, respectively, by

$$\begin{aligned} \underline{v}_1^{\text{mon}} &= \sup_{s_1: \substack{\omega_{s_1} \in \Omega \text{ and } s_1 \text{ is monotone} \\ \text{and not behaviorally confounded}}} \min_{(\alpha_0, \alpha_2) \in B(s_1)} u_1(\alpha_0, s_1, \alpha_2), \\ \bar{v}_1^{\text{mon}} &= \sup_{s_1: s_1 \text{ is monotone}} \max_{(\alpha_0, \alpha_2) \in B(s_1)} u_1(\alpha_0, s_1, \alpha_2). \end{aligned}$$

Combining Theorem 1 and Propositions 6 and 7, we obtain the following corollary.

**Corollary 3.** *Suppose  $u_1$  is strictly supermodular. Then*

$$\liminf_{\delta \rightarrow 1} \underline{U}_1(\delta) \geq \underline{v}_1^{\text{mon}}.$$

*Conversely, for all  $\varepsilon > 0$ , there exists  $\kappa > 0$  such that for any prior  $\mu_0$  satisfying  $\mu_0(\omega_R) > 1 - \kappa$  and any  $\delta < 1$ ,*

$$\bar{U}_1(\delta) < \bar{v}_1^{\text{mon}} + \varepsilon.$$

Corollary 3 is a main conclusion of this paper: a patient long-run player is assured at least her commitment payoff from any monotone, non-behaviorally confounded strategy  $s_1$  such that  $\mu_0(\omega_{s_1}) > 0$ . This implies the supermodular cases of Proposition 1 (as in the deterrence game  $u_1$  is strictly supermodular with the order  $A \succ F$  and  $C \succ D$  when  $x + y < 1$ ) and Proposition 2 (as in the trust game  $u_1$  is strictly supermodular with the order  $H \succ L$  and  $T \succ N$  when  $\min\{w, z\} > 0$ ), as well as Proposition 3 (as in a signaling game  $u_1 = (1 - \lambda)v(a_2(a_1)) - \lambda w(a_1, \theta)$  is strictly supermodular in  $(a_1, \theta)$  for any function  $a_2 : a_1 \rightarrow R$  if  $w$  is strictly submodular). Moreover, it does not require the assumption in Section 2 that there is only one commitment type that plays a pure strategy. Thus, Corollary 3 implies that, in general repeated signaling games with strictly submodular signaling costs and state-independent sender preferences, a patient long-run sender is assured at least her commitment payoff from any monotone, non-behaviorally confounded strategy  $s_1$  such that  $\mu_0(\omega_{s_1}) > 0$ , even if short-run players observe only past actions and not past state realizations.<sup>21</sup>

At the same time, the converse direction of Corollary 3 implies the submodular cases of Propositions 1 and 2. For example, in the deterrence game, if  $x + y > 1$  (the “submodular case”) then  $u_1$  is strictly supermodular with the order  $F \succ A$  and  $C \succ D$ . Since any long-run player strategy that is monotone with this order takes  $F$  with higher probability after  $c$ , the unique short-run player best response to any such strategy is  $D$ , implying that  $\bar{v}_1^{\text{mon}} = 1 - p + py$ . The argument for the submodular case of Proposition 2 is similar.

Finally, in games where  $\underline{v}_1^{\text{mon}} = \bar{v}_1^{\text{mon}}$  (which holds if the short-run players have a unique best response to any monotone strategy  $s_1$ ), Corollary 3 gives a unique payoff prediction as  $\mu_0(\omega_R)$  and  $\delta$  both approach 1. For example, this holds in the deterrence and trust games.

## 6 Communication Games

We now consider implications of our results for repeated communication games. Recall that the model covers communication games by dropping player 0; viewing  $\rho(y_0)$  as the prior distribution of a payoff-relevant state  $y_0$ ; letting  $\rho(y_1|a_1) = \mathbf{1}(\{y_1 = a_1\})$  (so  $a_1$  is perfectly monitored); viewing  $a_2$  as a mapping from  $a_1$  to a finite set of responses  $R$ ; and assuming that  $u_1$  and  $u_2$  depend on  $a_2$  only through the induced response  $a_2(a_1) \in R$ . In this section, we refer to player 1 as the sender and player 2 as the receiver, and we relabel  $y_0$  as  $\theta$ .

We have already observed that if the sender’s preferences are state-independent, meaning that  $u_1(\theta, a_1, r) = (1 - \lambda)v(a_1, r) - \lambda w(a_1, \theta)$  for some functions  $v$  and  $w$  and some  $\lambda \in (0, 1)$ ,

---

<sup>21</sup>Moreover, this result extends to preferences of the form  $u_1(a_1, r, \theta) = (1 - \lambda)v(a_1, r) - \lambda w(a_1, \theta)$ .

and if  $w$  is strictly submodular, then  $\limsup_{\delta \rightarrow 1} \bar{U}_1(\delta) \geq \underline{v}^{\text{mon}}$ . Thus, a patient sender with state-independent preferences over the receiver's action and a strictly submodular signaling cost can secure her best commitment payoff from any monotone signaling strategy.

We now turn to the following question. Consider a cheap talk game with a state-independent utility  $v(r)$  for the sender and a utility  $u_2(\theta, r)$  for the receiver, so the sender's action  $a_1$  is payoff irrelevant. In this game, the sender's utility  $u_1(\theta, a_1, r)$  is independent of  $a_1$  and hence cannot be strictly supermodular in  $(\tilde{r}, \theta)$ . However, suppose that the sender has a “grain of commitment power,” in that she can publicly adjust her preferences at the beginning of the game by committing to pay a small communication cost  $w(a_1, \theta)$  whenever she takes action  $a_1$  in state  $\theta$ . We ask what commitment payoffs  $V(s_1)$  can be secured by leveraging such a grain of commitment.

By a standard revelation principle argument, for any set of sender actions  $A_1$  and any strategy  $\hat{s}_1 : \Theta \rightarrow \Delta(A_1)$ , there exists a direct *communication mechanism*  $s_1 : \Theta \rightarrow \Delta(R)$  such that  $V(s_1) = V(\hat{s}_1)$ . We thus assume for the rest of this section that  $A_1 = R$ , so the sender's message  $a_1 = \tilde{r}$  can be interpreted as a recommended action for the receiver. The following result provides a general sufficient condition for the commitment payoff  $V(s_1)$  from a communication mechanism  $s_1 : \Theta \rightarrow \Delta(R)$  to be approximately securable.

**Proposition 8.** *If a communication mechanism  $s_1 : \Theta \rightarrow \Delta(R)$  is monotone with respect to some order  $(\succsim_\Theta, \succsim_R)$  and is such that  $\omega_{s_1} \in \Omega$  and  $s_1$  is not behaviorally confounded, then for any  $\varepsilon > 0$  and any strictly submodular cost function  $w : R \times \Theta \rightarrow [0, 1]$ ,*

$$\liminf_{\delta \rightarrow 1} \underline{U}^w(\delta) \geq (1 - \varepsilon)V(s_1) - \varepsilon,$$

where  $\underline{U}^w(\delta)$  is the infimum of the long-run player's payoff in any Nash equilibrium in the repeated game where her utility is given by

$$u_1(\theta, \tilde{r}, r) = (1 - \varepsilon)v(r) - \varepsilon w(\tilde{r}, \theta).$$

*Proof.* Immediately from Corollary 3. □

An interpretation of the communication cost  $w(\tilde{r}, \theta)$  is that this represents a “lying cost” (Chen, Kartik, and Sobel, 2008; Kartik, 2009) incurred by a sender who recommends action  $\tilde{r}$  in state  $\theta$ . In particular, if  $R = \Theta$  and the receiver's optimal action in state  $\theta$  is  $r = \theta$ , we can interpret the sender's message  $\tilde{r} \in \Theta$  as a report of the state, and we can interpret  $w(\tilde{r}, \theta)$  as the lying cost associated with misreporting state  $\theta$  as  $\tilde{r}$ . This example matches the

main example in Kartik (2009), where it is likewise assumed that the lying cost  $w(\tilde{r}, \theta)$  is strictly submodular. Proposition 8 thus implies that augmenting repeated cheap talk with a small lying cost provides a reputational foundation for any communication mechanism that is monotone with respect to some order over states and actions.<sup>22</sup>

To operationalize Proposition 8, it remains to characterize what mechanisms  $s_1 : \Theta \rightarrow \Delta(R)$  are monotone with respect to some order  $(\succ_{\Theta}, \succ_R)$ . Our last result provides this characterization. To state it, let  $G(s_1)$  be the bipartite graph with vertices  $\Theta$  and  $R$ , where a state  $\theta$  and an action  $r$  are linked if  $r \in \text{supp}(s_1(\theta))$ . We will see that if  $s_1$  is monotone then  $G(s_1)$  is acyclic and also does not contain what we call a “forbidden triple.”

**Definition 7.** A *forbidden triple* for a mechanism  $s_1 : \Theta \rightarrow \Delta(R)$  is either

- (1) A set of three distinct actions  $\{r_1, r_2, r_3\}$  and four distinct state  $\{\theta_1, \theta_2, \theta_3, \theta_4\}$  where  $r_k \in \text{supp}(s_1(\theta_k))$  for  $k \in \{1, 2, 3\}$  and  $\{r_1, r_2, r_3\} \subset \text{supp}(s_1(\theta_4))$ ; or
- (2) A set of three distinct states  $\{\theta_1, \theta_2, \theta_3\}$  and four distinct actions  $\{r_1, r_2, r_3, r_4\}$  where  $\{r_k, r_4\} \in \text{supp}(s_1(\theta_k))$  for  $k \in \{1, 2, 3\}$ .

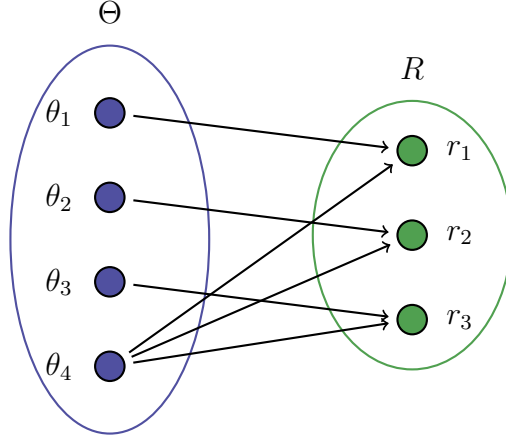


Figure 1: A Type (1) Forbidden Triple

*Notes.* Monotonicity is violated for any placement of  $\theta_4$  in the order  $\succ_{\Theta}$ .

For example, if a Type (1) forbidden triple exists, where without loss  $r_1 \prec r_2 \prec r_3$  and  $\theta_1 \prec \theta_2 \prec \theta_3$ , then  $s_1$  cannot be monotone with respect to any order, as if  $\theta_4 \prec \theta_2$  then  $s_1$

<sup>22</sup>For example, adding a Stackelberg commitment type and a small lying cost in the infinitely-repeated “political correctness” game of Morris (2001) suffices to secure the Stackelberg payoff, in contrast to Morris’s negative result.

is non-monotone because  $r_3 \in \text{supp}(s_1(\theta_4))$  but  $r_2 \in \text{supp}(s_1(\theta_2))$ ; and if  $\theta_4 \succ \theta_2$  then  $s_1$  is non-monotone because  $r_1 \in \text{supp}(s_1(\theta_4))$  but  $r_2 \in \text{supp}(s_1(\theta_2))$ . See Figure 1.

Our final result is that, conversely, if  $G(s_1)$  is acyclic and does not contain a forbidden triple, then  $s_1 : \Theta \rightarrow \Delta(R)$  is monotone with respect to some order.

**Proposition 9.** *A communication mechanism  $s_1 : \Theta \rightarrow \Delta(R)$  is monotone with respect to some order  $(\succsim_\Theta, \succsim_R)$  if and only if  $G(s_1)$  is acyclic and does not contain a forbidden triple.*

Proposition 9 implies that the set of mechanisms that are monotone with respect to some order includes, for example, all partitions (i.e., deterministic mechanisms  $s_1 : \Theta \rightarrow R$ ) and all linear partitions with randomization at the endpoints. This set includes the set of all *monotone partitions*, which are shown to be optimal in certain persuasion problems by Kolotilin (2018) and Dworzak and Martini (2019). However, it does not always include the optimal mechanism. For example, if the receiver’s optimal action is  $r = \mathbb{E}[\theta]$  and  $\theta \in \{0, 1\}$  with equal probability, and the sender’s utility is  $\mathbf{1}(\{r \in \{1/3, 2/3\}\})$ , then the unique optimal mechanism induces  $r \in \{1/3, 2/3\}$  with equal probability, but this mechanism is not monotone with respect to any order because the corresponding graph  $G(s_1)$  contains the cycle  $(0, 1/3), (1, 1/3), (1, 2/3), (0, 2/3)$ .<sup>23</sup>

Proposition 9 is a general mathematical result that could have other applications (and that may have been previously noted in other contexts, although we have not found a reference). It characterizes when a Markov transition matrix  $f : X \rightarrow Y$  is consistent with a joint distribution over  $X \times Y$  that is co-monotone with respect to some order  $(\succsim_X, \succsim_Y)$ . Alternatively, it characterizes when the vertices of a bipartite graph can be drawn on two straight lines so that no edges cross.

## 7 Discussion

This paper has studied reputation-formation when a player desires a reputation for conditional action. The main result is that if a strategy is *confound-defeating* and either *not behaviorally confounded* or *salient*, a patient long-run player can secure the corresponding commitment payoff. A strategy is confound-defeating if and only if it is the unique solution to an optimal transport problem. In one-dimensional supermodular games, a strategy

---

<sup>23</sup>This example is a discrete version of a “bi-pooling” policy. With a continuous state, Kleiner, Moldovanu, and Strack (2021) and Arieli et al. (2023) show that the bi-pooling policies are those that are uniquely optimal in some persuasion problem.

is confound-defeating if and only if it is monotone. In repeated signaling games with state-independent sender preferences and a strictly submodular signaling cost, a patient sender can secure her commitment payoff from any monotone strategy. Finally, we characterized the communication strategies that are monotone with respect to some order, and that are thus implementable with a small “lying cost.”

We mention some possible extensions of our results. First, the connection between unobserved deviations and optimal transport is not specific to the long-run/short-run model we study and could also be applied to repeated games with multiple long-run players, with or without incomplete information. Second, extending the model to allow multiple short-run players and to allow  $u_0$  to depend on  $a_2$  would encompass reputation-formation by a long-run mediator who coordinates play among multiple short-run players. This extension can potentially provide a reputational foundation for a general static mediation solution, as in Myerson (1982). Third, our results extend to the case with multiple rational types with different preferences, so long as they all have the same Stackelberg strategy and it is confound-defeating for all of them. A possible extension to the case with multiple rational types with different Stackelberg strategies that are confound-defeating only for some types would require additional analysis and qualifications. Fourth, cyclical monotonicity can be explored in multidimensional games, for example communication games with multidimensional states or actions. In particular, we are not sure if Proposition 8 has a useful multidimensional analogue. Fifth, in signaling games, we gave conditions under which the sender can secure her commitment payoff from any monotone strategy. An open question is when the Stackelberg signaling strategy is monotone. Finally, Watson (1993) and Battigalli and Watson (1997) show that the classic reputation results of Fudenberg and Levine (1989; 1992) require only two rounds of iterated deletion of dominated strategies, rather than the full force of Nash equilibrium. Our results instead require three rounds of deletion: under our conditions, the long-run player secures the Stackelberg payoff by best responding to any short-run player best response to any long-run player strategy that is not undetectably dominated.

## A Appendix: Salience

This appendix generalizes Theorem 1 to the case where  $s_1^*$  is behaviorally confounded. As noted in the text, this extension is particularly important for games where  $\alpha_0$  is endogenous, like the deterrence game in Section 2.1, as in these games  $s_1^*$  is often behaviorally confounded when there are multiple commitment types.

Our approach to extending Theorem 1 is as follows. If  $s_1^*$  is behaviorally confounded, we calculate the minimum weight on  $s_1^*$  that ensures that once short-run players learn the desired signal distribution, they best respond to  $s_1^*$  (rather than a confounding strategy). Then we calculate the minimum probability  $\beta$  that the long-run weight on  $s_1^*$  exceeds this level under the deviation measure  $\mathbb{Q}$ . We call  $\beta$  the “salience” of  $s_1^*$ , and we establish a lower bound for a patient long-run player’s payoff as a function of  $\beta$ . If  $s_1^*$  is not behaviorally confounded then its salience is 1, in which case we recover Theorem 1.

Before defining salience, we require a preliminary definition. In what follows, given a belief  $\mu \in \Delta(\Omega)$  and a subset  $\Omega' \subset \Omega$ , we denote the conditional distribution of  $\mu$  over  $\Omega'$  by  $\mu(\cdot|\Omega')$ . In addition, given a belief  $\mu \in \Delta(\Omega \setminus \{\omega_R\})$ , we slightly abuse notation by also denoting by  $\mu$  the strategy  $\tilde{s}_1$  given by  $\tilde{s}_1(y_0)[a_1] = \sum_{\omega_{s_1} \in \Omega \setminus \{\omega_R\}} \mu(\omega_{s_1})s_1(y_0)[a_1]$ .

**Definition 8.** For any  $\eta, \varsigma > 0$ , a number  $c \in [0, 1]$  is an  $(\eta, \varsigma)$ -confounding weight if there exists a belief  $\mu \in \Delta(\Omega)$  satisfying the following conditions.

- (1)  $\mu(\omega_R) < 1$ .
- (2) There exists  $s_1 \in \Delta(A_1)^{Y_0}$  such that  $\|\mu(\cdot|\Omega \setminus \{\omega_R\}) - s_1\| < \varsigma$  and  $B(s_1) \setminus B(s_1^*) \neq \emptyset$ .
- (3)  $\mu(\Omega_\eta(s_1^*)) > 1 - \eta$ , where

$$\Omega_\eta(s_1^*) = \{\omega_{s_1} \in \Omega \setminus \{\omega_R\} : \|p(\alpha_0, s_1, \alpha_2) - p(\alpha_0, s_1^*, \alpha_2)\| < \eta \text{ for some } (\alpha_0, \alpha_2) \in B_1(s_1^*) \cup \{\omega_R\}\}.$$

- (4)  $\mu(s_1^*|\Omega \setminus \{\omega_R\}) = c$ .

Let  $c_{\eta, \varsigma} = \sup\{c : c \text{ is } (\eta, \varsigma)\text{-confounding}\}$ , with the convention that if no  $(\eta, \varsigma)$ -confounding weight exists, then  $c_{\eta, \varsigma} = -\infty$ . Finally, let  $c_0 = \lim_{\varsigma \rightarrow 0} \lim_{\eta \rightarrow 0} c_{\varsigma, \eta}$ .

Condition (1) implies that  $\mu(\cdot|\Omega \setminus \{\omega_R\})$  is well-defined. Condition (2) says that  $\mu(\cdot|\Omega \setminus \{\omega_R\})$  is within  $\varsigma$  of a belief to which the short-run players have a best response outside  $B(s_1^*)$ . Condition (3) says that  $\mu$  puts at most  $\eta$  weight on commitment types that induce signals that are not  $\eta$  close to those induced by  $s_1^*$ . Condition (4) says that  $\mu(\cdot|\Omega \setminus \{\omega_R\})$  assigns probability  $c$  to  $s_1^*$ . Note that  $c_0 < 1$ , by upper hemi-continuity of the best-response correspondence  $B(\cdot)$ . In addition, the sets  $\Omega_\eta(s_1^*)$  are nested and all contain  $s_1^*$ , which implies that the set

$$\Omega_0(s_1^*) := \bigcap_{\eta > 0} \Omega_\eta(s_1^*)$$

is well-defined and contains  $\{\omega_{s_1^*}, \omega_R\}$ .

**Definition 9.** The *salience* of a strategy  $s_1^* \in \Omega$  is

$$\beta = \max \left\{ \frac{\mu_0(s_1^* | \Omega_0(s_1^*) \setminus \{\omega_R\}) - c_0}{1 - c_0}, 0 \right\},$$

with the convention that if  $c_0 = -\infty$  then  $\beta = 1$ .

The logic of this definition is that conditional on the long-run player being irrational,  $c_0$  is the minimum weight on  $s_1^*$  that ensures short-run players best respond to  $s_1^*$ , and (by Bayes' rule)  $\beta$  is the minimum probability the long-run weight on  $s_1^*$  strictly exceeds  $c_0$ .

The general version of our main result is as follows.

**Theorem 2.** *If  $\omega_{s_1^*} \in \Omega$  and  $s_1^*$  is confound-defeating and has salience  $\beta$ , then*

$$\liminf_{\delta \rightarrow 1} \underline{U}_1(\delta) \geq \beta V(s_1^*) + (1 - \beta)V_0(s_1^*).$$

Note that if  $s_1^*$  is not behaviorally confounded then  $\Omega_0(s_1^*) = \{\omega_{s_1^*}, \omega_R\}$  for sufficiently small  $\eta$ . By upper hemi-continuity of  $B(\cdot)$ , this implies that  $c_{\eta, \varsigma} = -\infty$  for sufficiently small  $\eta$  and  $\varsigma$ , so  $\beta = 1$ . Theorem 2 therefore generalizes Theorem 1.

Moreover, since  $c_0 < 1$ , we have  $\beta \rightarrow 1$  whenever  $\mu_0|_{\Omega_0(s_1^*) \setminus \{\omega_R\}}(s_1^*) \rightarrow 1$ , in which case Theorem 2 recovers the conclusion of Theorem 1 even if  $s_1^*$  is behaviorally confounded. In fact, Theorem 1 delivers much more than continuity of the payoff lower bound at  $\mu_0|_{\Omega_0(s_1^*) \setminus \{\omega_R\}}(s_1^*) = 1$ —it gives an explicit lower bound that declines linearly with  $\mu_0|_{\Omega_0(s_1^*) \setminus \{\omega_R\}}(s_1^*)$ .<sup>24</sup>

For example, consider the deterrence game from Section 2.1 with two commitment types, the pure Stackelberg type  $(A, F)$  and the type  $s_1$  that takes  $A$  with probability  $p$  for each signal. In this example,  $c_0$  is the probability such that the short-run player is indifferent between  $C$  and  $D$  when the long-run player plays  $(A, F)$  with probability  $c_0$  and plays  $s_1$  with probability  $1 - c_0$ , which is given by  $c_0 = \frac{pg + (1-p)l}{(2p-1)(1+g)}$ . The salience of type  $(A, F)$  is then  $\beta = \max \left\{ \frac{\mu_0|_{\Omega_0(s_1^*) \setminus \{\omega_R\}}((A, F)) - c_0}{1 - c_0}, 0 \right\}$ , and Theorem 2 implies that as  $\delta \rightarrow 1$  the long-run player is assured a payoff of at least  $\beta p + (1 - \beta)(1 - p)$ . In particular, whenever  $C$  is the unique best response to  $(A, F)$ , we have  $c_0 < 1$ , so that as the prior weight on  $(A, F)$  relative to  $s_1$  increases,  $\beta$  converges to 1 and the long-run player is assured her pure Stackelberg payoff  $p$ .

<sup>24</sup>This feature contrasts with the approach of Ely, Fudenberg, and Levine (2008), who show that introducing a sufficiently high conditional probability of the Stackelberg type overturns Ely and Valimaki's 2003 bad reputation result, but require an unbounded likelihood ratio between the Stackelberg type and the "bad commitment type."



The proof of Theorem 2 follows from the proof of Theorem 1 and the fact that  $(\alpha_0, \alpha_2) \in B_\eta(s_1^*)$  (and hence  $u_1(\sigma_0^*, s_1^*, \sigma_2^*) \geq \inf_{(\alpha_0, \alpha_2) \in B_\eta(s_1^*)} u_1(\alpha_0, s_1^*, \alpha_2)$ ) for all  $h^t \in H_\eta^t$ , once we replace Lemma 4 with the following lemma.

**Lemma 7.** *For any  $\eta$  sufficiently small, any  $t > \bar{T}(\eta)$  (chosen as in Theorem 1), and any Nash equilibrium  $(\sigma_0^*, \sigma_1^*, \sigma_2^*)$ , we have*

$$\liminf_{\xi \rightarrow 0} \mathbb{Q}(h \in H^\infty : (\sigma_0^*(h^t), \sigma_2^*(h^t)) \in \hat{B}_\xi(s_1^*)) \geq \beta.$$

The proof of Lemma 7 in turn relies on the following technical lemma, which will be used to show that if  $B(\mu) \subset B(s_1^*)$  and  $d(\tilde{\mu}, \text{conv}(\{\mu, s_1^*\})) < \vartheta$  (where  $d(\cdot, \cdot)$  denotes distance from a point to a set, and  $\text{conv}(\cdot)$  denotes convex hull), then  $B(\tilde{\mu}) \subset B(s_1^*)$ , where  $\vartheta$  can be chosen uniformly over a certain set of beliefs  $\mu$ .

**Lemma 8.** *For any  $s_1 \in \Delta(A_1)^{Y_0}$  and  $\varsigma > 0$ , let*

$$C_\varsigma(s_1) = \{s'_1 : B(s''_1) \subset B(s_1) \text{ for all } s''_1 \text{ s.t. } \|s'_1 - s''_1\| \leq \varsigma\}.$$

*Then, there exists  $\vartheta(\varsigma, s_1) > 0$ , vanishing as  $\varsigma \rightarrow 0$ , such that for all  $\tilde{s}_1, s'_1$  such that  $s'_1 \in C_\varsigma(s_1)$  and  $d(\tilde{s}_1, \text{conv}(\{s_1, s'_1\})) \leq \vartheta(\varsigma, s_1)$ , we have  $B(\tilde{s}_1) \subset B(s_1)$ .*

*Proof.* We first show that  $B(\tilde{s}_1) \subset B(s_1)$  if  $\vartheta(\varsigma, s_1) = 0$ , so that  $\tilde{s}_1 \in \text{conv}(\{s_1, s'_1\})$ . To see this, note that since  $B(s'_1) \subset B(s_1)$ , the set of player 0 best responses at any  $\tilde{s}_1 \in \text{conv}(\{s_1, s'_1\})$  (other than  $s'_1$ ) is the same as that at  $s'_1$ , by the sure-thing principle. This then implies the same conclusion for player 2, so  $B(\tilde{s}_1) = B(s'_1) \subset B(s_1)$ .

Next, we show there exists some  $\vartheta(\varsigma, s_1) > 0$ , which can be chosen independently of  $s'_1 \in C_\varsigma(s_1)$ , such that if  $d(\tilde{s}_1, \text{conv}(\{s_1, s'_1\})) \leq \vartheta(\varsigma, s_1)$  then  $B(\tilde{s}_1) \subset B(s_1)$ . For any  $\varsigma$ , note that if  $s''_1 \in \overline{C_\varsigma(s_1)}$  (the closure of  $C_\varsigma(s_1)$ ), then  $B(s''_1) \subset B(s_1)$ . From here, suppose no such  $\vartheta > 0$  has the desired property for all  $s'_1 \in C_\varsigma(s_1)$ . Then there exists a sequence  $(s_1^n, \tilde{s}_1^n)$  such that  $s_1^n \in C_\varsigma(s_1)$ ,  $d(\tilde{s}_1^n, \text{conv}(s_1^n, s_1)) < \frac{1}{n}$ , and  $B(\tilde{s}_1^n) \setminus B(s_1) \neq \emptyset$  for large enough  $n$ . Taking a subsequence, this implies there exists  $(s_1^{n_k}, \tilde{s}_1^{n_k}) \rightarrow (s'_1, \tilde{s}_1)$  such that  $s'_1 \in \overline{C_\varsigma(s_1)}$  (since this set is closed) and  $d(\tilde{s}_1, \text{conv}(s'_1, s_1)) = 0$ , but  $B(\tilde{s}_1) \setminus B(s_1) \neq \emptyset$ . But this contradicts the fact, established above, that  $B(\tilde{s}_1) \subset B(s_1)$  when  $\tilde{s}_1 \in \text{conv}(\{s_1, s'_1\})$ , completing the proof.  $\square$

*Proof of Lemma 7.* We show there exists strictly positive functions  $\zeta(\eta)$  and  $\xi(\eta)$ , vanishing

as  $\eta \rightarrow 0$ , and  $\bar{T}(\eta)$  such that, for all  $t > \bar{T}(\eta)$ ,

$$\mathbb{Q}(h \in H^\infty : (\sigma_0^*(h^t), \sigma_2^*(h^t)) \in \hat{B}_{\xi(\eta)}(s_1^*)) \geq (1 - \zeta(\eta))\beta_{\zeta, \eta}, \quad \text{where}$$

$$\beta_{\zeta, \eta} = \frac{(1 - \eta)\mu_0|_{\Omega_0(s_1^*) \setminus \{\omega_R\}}(s_1^*) - c_{\eta, \varsigma}}{1 - \eta - c_{\eta, \varsigma}}.$$

Lemma 2 and an appropriate modification of Lemma 3 (with  $\Omega_\eta(s_1^*)$  in place of  $\{\omega_{s_1^*}\}$ ) imply that, on a set of histories  $G(\zeta(\eta))$  satisfying  $\mathbb{Q}(G(\zeta(\eta))) > 1 - \zeta(\eta)$ , both  $h^t \in H_\eta^t$  and  $\mu_t(\Omega_\eta(s_1^*) \setminus \{\omega_R\} | h^t) > 1 - \eta$  for all  $t > \bar{T}(\eta)$ , independent of the choice of the equilibrium strategy and discount factor. Suppose these two conditions are satisfied and  $t > \bar{T}(\eta)$ . We consider three possible cases, and show for sufficiently small  $\eta$ , that in the first two cases  $(\sigma_0^*(h^t), \sigma_2^*(h^t)) \in \hat{B}_{\xi(\eta)}(s_1^*)$  and the third arises with probability at most  $1 - \beta_{\zeta, \eta}$ . This then implies, in total,  $\mathbb{Q}(h \in H^\infty : (\sigma_0^*(h^t), \sigma_2^*(h^t)) \in \hat{B}_{\xi(\eta)}(s_1^*))$  is no less than  $(1 - \zeta(\eta))\beta_{\zeta, \eta}$ , completing the proof.

First, suppose that  $\mu_t(\{\omega_R, \omega_{s_1^*}\} | h^t) > 1 - \zeta(\eta)$ . Then, for  $\zeta(\eta)$  and  $\xi(\eta)$  chosen as in Lemma 4, we have that  $(\sigma_0^*(h^t), \sigma_2^*(h^t)) \in \hat{B}_{\xi(\eta)}(s_1^*)$ .

Second, suppose that  $\mu_t(\{\omega_R, \omega_{s_1^*}\} | h^t) \leq 1 - \zeta(\eta)$  but  $\mu_t(\cdot | h^t, \Omega \setminus \{\omega_R\}) \in C_\varsigma(s_1^*)$ , for some  $\varsigma > 0$  fixed independent of  $\eta$ . Since  $h^t \in H_\eta^t$  and  $\mu_t(\Omega_\eta(s_1^*) | h^t) > \mu_t(\Omega_\eta(s_1^*) \setminus \{\omega_R\}) > 1 - \eta$ , an argument identical to the proof of Lemma 4 implies that the minimum of  $\mu_t(\omega_R | h^t)$  and  $\|\sigma_1^*(h^t) - s_1^*\|$  is bounded above by a function that vanishes as  $\eta \rightarrow 0$ . Thus, as  $\eta$  vanishes,  $d(\sigma_1^*(h^t), \text{conv}(\{\mu_t(\cdot | h^t, \Omega_\eta(s_1^*) \setminus \{\omega_R\}), s_1^*\})) \leq \vartheta(\varsigma, s_1^*)$  where  $\vartheta(\cdot, \cdot)$  is defined in Lemma 8. Since  $\mu_t(\cdot | h^t, \Omega \setminus \{\omega_R\}) \in C_\varsigma(s_1^*)$ , Lemma 8 then implies  $(\sigma_0^*(h^t), \sigma_2^*(h^t)) \in B(s_1^*) \subset \hat{B}_{\xi(\eta)}(s_1^*)$ .

Third, suppose that  $\mu_t(\{\omega_R, \omega_{s_1^*}\} | h^t) \leq 1 - \zeta(\eta)$  and  $\mu_t(\cdot | h^t, \Omega \setminus \{\omega_R\}) \notin C_\varsigma(s_1^*)$ . The former condition implies that  $\mu_t(\cdot | h^t)$  satisfies Condition (1) of Definition 8, while the latter condition implies that it also satisfies Condition (2). Moreover, since  $\mu_t(\Omega_\eta(s_1^*) | h^t) \geq 1 - \eta$  (as  $t > \bar{T}(\eta)$ ), it also satisfies Condition (3). Thus, by the definition of  $(\eta, \varsigma)$ -confounding weights,  $\mu_t(\omega_{s_1^*} | h^t, \Omega \setminus \{\omega_R\}) \leq c_{\varsigma, \eta}$ . Now, since  $\omega_{s_1^*} \in \Omega_\eta(s_1^*) \setminus \{\omega_R\} \subset \Omega \setminus \{\omega_R\}$ , we have

$$\mu_t(\omega_{s_1^*} | h^t, \Omega \setminus \{\omega_R\}) = \mu_t(\omega_{s_1^*} | h^t, \Omega_\eta(s_1^*) \setminus \{\omega_R\}) \mu_t(\Omega_\eta(s_1^*) \setminus \{\omega_R\} | h^t).$$

Since  $\mu_t(\omega_{s_1^*} | h^t, \Omega \setminus \{\omega_R\}) \leq c_{\varsigma, \eta}$  and  $\mu_t(\Omega_\eta(s_1^*) \setminus \{\omega_R\} | h^t) \geq 1 - \eta$ , we have

$$\mu_t(\omega_{s_1^*} | h^t, \Omega_\eta(s_1^*) \setminus \{\omega_R\}) \leq \frac{c_{\varsigma, \eta}}{1 - \eta}.$$

Hence, the probability that  $h$  lies in this third case is at most

$$q_{\eta,\varsigma} := \mathbb{Q} \left( h \in H^\infty : \mu_t(\omega_{s_1^*} | h^t, \Omega_\eta(s_1^*) \setminus \{\omega_R\}) \leq \frac{c_{\eta,\varsigma}}{1-\eta} \right).$$

Thus, because  $\mu_t(s_1^* | h^t, \Omega_\eta(s_1^*) \setminus \{\omega_R\})$  is a  $\mathbb{Q}$ -submartingale, we have

$$\begin{aligned} q_{\eta,\varsigma} \left( \frac{c_{\eta,\varsigma}}{1-\eta} \right) + (1 - q_{\eta,\varsigma})(1) &\geq \mu_0(s_1^* | \Omega_\eta(s_1^*) \setminus \{\omega_R\}) && \iff \\ q_{\eta,\varsigma} &\leq \min \left\{ \frac{1 - \mu_0(s_1^* | \Omega_\eta(s_1^*) \setminus \{\omega_R\})}{1 - \frac{c_{\eta,\varsigma}}{1-\eta}}, 1 \right\} = 1 - \beta_{\eta,\varsigma}, \end{aligned}$$

completing the proof. □

## B Appendix: Omitted Proofs

### B.1 Proof of Lemma 2

For any two signal distributions  $p$  and  $q$ , let  $d(p||q) = \int \log(p/q) dp$  denote the relative entropy from  $q$  to  $p$ . By Lemma 4 of Gossner (2011) and a standard application of the chain rule for relative entropy,

$$\sum_t \mathbb{E}^{\mathbb{Q}} \left[ d(p(\sigma_0^*, s_1^*, \sigma_2^* | h^t) || p(\sigma_0^*, \sigma_1^*, \sigma_2^* | h^t)) \right] \leq -\log \mu_0(\omega_{s_1^*}).$$

Hence, by Markov's inequality,

$$\mathbb{E}^{\mathbb{Q}} \left[ \# \left\{ t : d(p(\sigma_0^*, s_1^*, \sigma_2^* | h^t) || p(\sigma_0^*, \sigma_1^*, \sigma_2^* | h^t)) > \frac{\eta^2}{2} \right\} \right] < -\frac{2 \log \mu_0(\omega_{s_1^*})}{\eta^2}.$$

On the other hand, by Pinsker's inequality,

$$d \left( p(\sigma_0^*, s_1^*, \sigma_2^* | h^t) \middle| \middle| p(\sigma_0^*, \sigma_1^*, \sigma_2^* | h^t) \right) \leq \frac{\eta^2}{2} \implies h^t \in H_\eta^t.$$

This gives the desired bound.

## B.2 Proof of Lemma 3

We first show that the desired conclusion holds for each  $\delta$  and each equilibrium, and then show that  $\hat{T}$  can be fixed independent of the choice of  $\delta$  and the equilibrium.

**Lemma 9.** *For any  $\delta < 1$ , any strategy profile  $(\sigma_0^*, \sigma_1^*, \sigma_2^*)$  where  $(\sigma_0^*, \sigma_2^*) \in B_1(s_1^*)^H$  (and hence, any Nash equilibrium), and any  $\zeta > 0$ , there exists a set of infinite histories  $G(\zeta) \subset H^\infty$  satisfying  $\mathbb{Q}(G(\zeta)) > 1 - \zeta$  and a period  $\hat{T}$  such that, for any  $h \in G(\zeta)$  and any  $t \geq \hat{T}$ , we have  $\mu_t(\cdot|h) \in M_\zeta$ .*

*Proof.* Since  $\mathbb{Q}$  is absolutely continuous relative to  $\mathbb{P}$  and  $\mu_t(\cdot|h)$  is a martingale relative to  $\mathbb{P}$ ,  $\mu_t(\cdot|h)$  converges  $\mathbb{Q}$ -almost surely to some limit distribution  $\mu_\infty(\cdot|h)$  (e.g., Mailath and Samuelson (2006), Lemma 15.4.2).

We show that, for  $\mathbb{Q}$ -almost all histories  $h \in H^\infty$ ,  $\mu_\infty(\{\omega_R, \omega_{s_1^*}\}|h) = 1$ . Suppose that  $\mu_\infty(\Omega \setminus \{\omega_R\}|h) > 0$ , let  $\omega_{s_1} \in \Omega \setminus \{\omega_R\}$  satisfy  $\mu_\infty(\omega_{s_1}|h) > 0$ , and let  $c > 0$  and  $T$  satisfy  $\mu_t(\omega_{s_1}|h) > c$  for all  $t > T$ . Suppose also that the set of signals  $y_1$  that realize infinitely often in  $h$  is precisely  $Y_1^* = \text{supp}(\rho(\cdot|s_1^*(Y_0)))$ , the set of signals that arise with positive probability when player 1 plays  $s_1^*$ . This supposition is without loss as  $\rho(\cdot|a_0)$  has full support, so such histories occur  $\mathbb{Q}$ -almost surely. Next, for any  $\Omega' \subseteq \Omega \setminus \{\omega_R\}$ , let  $p_{Y_1}(\sigma_0^*, \tilde{s}_1|h^t, \Omega')$  denote the equilibrium distribution of  $y_1$  conditional on reaching history  $h^t$  and the event  $\omega \in \Omega'$ ; when  $\Omega'$  is a singleton,  $\Omega' = \{\hat{s}_1\}$ , we write this as  $p_{Y_1}(\sigma_0^*, \hat{s}_1|h^t)$ . Then, for any  $y_1 \in Y_1^*$ , we have

$$\begin{aligned} & \left| \mu_{t+1}(\omega_{s_1}|\Omega \setminus \{\omega_R\}, h^t, y_1) - \mu_t(\omega_{s_1}|\Omega \setminus \{\omega_R\}, h^t) \right| \\ &= \left| \frac{p_{Y_1}(\sigma_0^*, s_1|h^t)[y_1] \mu_t(\omega_{s_1}|\Omega \setminus \{\omega_R\}, h^t)}{p_{Y_1}(\sigma_0^*, \tilde{s}_1|h^t, \Omega \setminus \{\omega_R\})[y_1]} - \mu_t(\omega_{s_1}|\Omega \setminus \{\omega_R\}, h^t) \right| \\ &= \frac{\mu_t(\omega_{s_1}|\Omega \setminus \{\omega_R\}, h^t)}{p_{Y_1}(\sigma_0^*, \tilde{s}_1|h^t, \Omega \setminus \{\omega_R\})[y_1]} \left| p_{Y_1}(\sigma_0^*, s_1|h^t)[y_1] - p_{Y_1}(\sigma_0^*, \tilde{s}_1|h^t, \Omega \setminus \{\omega_R\})[y_1] \right| \\ &> c \left| p_{Y_1}(\sigma_0^*, s_1|h^t)[y_1] - p_{Y_1}(\sigma_0^*, \tilde{s}_1|h^t, \Omega \setminus \{\omega_R\})[y_1] \right|. \end{aligned}$$

Since  $\mu_t(\cdot|h)$  converges,  $y_1$  realizes infinitely often, and  $c > 0$ , this implies

$$\lim_{t \rightarrow \infty} \left| p_{Y_1}(\sigma_0^*, s_1|h^t)[y_1] - p_{Y_1}(\sigma_0^*, \tilde{s}_1|h^t, \Omega \setminus \{\omega_R\})[y_1] \right| = 0.$$

At the same time, applying the argument in Lemma 2 conditional on the event  $\omega \neq \omega_R$  implies that, for  $\mathbb{Q}$ -almost all histories  $h \in H^\infty$ ,

$$\lim_{t \rightarrow \infty} \left| p_{Y_1}(\sigma_0^*, \tilde{s}_1|h^t, \Omega \setminus \{\omega_R\}) - p_{Y_1}(\sigma_0^*, s_1^*|h^t) \right| = 0.$$

In particular, since  $p_{Y_1}(\sigma_0^*, s_1^* | h^t)[y_1] = 0$  for all  $y_1 \notin Y_1^*$ , this implies that, for all  $y_1 \notin Y_1^*$ ,

$$\lim_{t \rightarrow \infty} p_{Y_1}(\sigma_0^*, \tilde{s}_1 | h^t, \Omega \setminus \{\omega_R\})[y_1] = 0.$$

Since we have already shown that  $p_{Y_1}(\sigma_0^*, \tilde{s}_1 | h^t, \Omega \setminus \{\omega_R\})[y_1]$  and  $p_{Y_1}(\sigma_0^*, s_1 | h^t)[y_1]$  have the same limit for all  $y_1 \in Y_1^*$ , we have

$$\lim_{t \rightarrow \infty} \left\| p_{Y_1}(\sigma_0^*, s_1 | h^t) - p_{Y_1}(\sigma_0^*, \tilde{s}_1 | h^t, \Omega \setminus \{\omega_R\}) \right\| = 0.$$

Thus, by the triangle inequality,

$$\lim_{t \rightarrow \infty} \left\| p_{Y_1}(\sigma_0^*, s_1 | h^t) - p_{Y_1}(\sigma_0^*, s_1^* | h^t) \right\| = 0.$$

Finally, since  $\omega_{s_1} \in \Omega$  and  $s_1^*$  is not behaviorally confounded, this implies that  $s_1 = s_1^*$ , and hence  $\mu_\infty(\{\omega_R, \omega_{s_1^*}\} | h) = 1$ .

To complete the proof, recall that Egorov's theorem shows that if a sequence of functions  $f_n : H^\infty \rightarrow \mathbb{R}$  converges  $\mathbb{Q}$ -almost surely to  $f$ , then for all  $\zeta > 0$ , there exists  $G(\zeta) \subset H^\infty$  satisfying  $\mu(G(\zeta)) \geq 1 - \zeta$  such that  $f_n \rightarrow f$  uniformly on  $G(\zeta)$ . Thus, Lemma 9 follows from Egorov's theorem applied to the sequence of conditional beliefs  $\mu_t(\{\omega_{s_1^*}, \omega_R\} | h)$  and the definition of uniform convergence.  $\square$

We now show that  $\hat{T}$  can be chosen as a function only of  $\zeta$  and not of  $\delta$  or the equilibrium strategies. To this end, let  $\mathbb{Q}^{\sigma_0, \sigma_2}$  be the probability measure on  $H^\infty$  induced by strategies  $(\sigma_0, s_1^*, \sigma_2)$ , and let  $\mu_t^{\sigma_0, \sigma_2}(\omega_{s_1^*} | \Omega \setminus \{\omega_R\}, h^t)$  be the conditional belief that the long-run player is of type  $\omega_{s_1^*}$  conditional on being irrational. This is well-defined because, conditional on the event  $\Omega \setminus \{\omega_R\}$ , the rational long-run player's strategy does not affect  $\mu_t^{\sigma_0, \sigma_2}$  once  $(\sigma_0, \sigma_2)$  are given. Next, let  $L^{\sigma_0, \sigma_2} \subset H^\infty$  be the set of all histories  $h$  where  $\mu_t^{\sigma_0, \sigma_2}(\omega_{s_1^*} | \Omega \setminus \{\omega_R\}, h^t) \rightarrow 1$  as  $t \rightarrow \infty$ . By Lemma 9,  $\mathbb{Q}^{\sigma_0, \sigma_2}(L^{\sigma_0, \sigma_2}) = 1$  for all  $(\sigma_0, \sigma_2) \in B_1(s_1^*)^H$ .

We show that  $B_1(s_1^*)^H$  is compact in an appropriate topology. Endow  $B_1(s_1^*)^H$  with the topology induced by the metric

$$d((\sigma_0, \sigma_2), (\sigma'_0, \sigma'_2)) = \sup_{h \in H} \left\{ \sum_{t=0}^{\infty} \frac{1}{2^t} \|(\sigma_0(h^t), \sigma_2(h^t)) - (\sigma'_0(h^t), \sigma'_2(h^t))\| \right\}.$$

This is the sup-norm over undominated short-run player strategies  $(\sigma_0, \sigma_2) : H \rightarrow B_1(s_1^*)$ .

Endowing  $H^\infty$  with the product topology, we have that  $H$  is a dense subset of  $H^\infty$ <sup>25</sup>. We can thus continuously extend  $B_1(s_1^*)^H$  to the larger space  $B_1(s_1^*)^{H^\infty}$ , that is, the space  $B_1(s_1^*)^{H^\infty}$  is the set of all continuous functions from infinite sequences of signals  $H^\infty$  into sequences of strategies  $B_1(s_1^*)^\infty$  under the sup norm. Thus, since  $H^\infty$  has countable dense subset  $H$ , a standard diagonalization argument implies that the space  $(B_1(s_1^*)^{H^\infty}, d)$  is compact.

We are now ready to prove that  $\hat{T}(\zeta, \sigma_0^*, \sigma_2^*)$  can be chosen independent of the choice of  $(\sigma_0^*, \sigma_2^*)$  (and hence also independent of  $\delta$ ). Suppose for contradiction that there exists  $\zeta > 0$  such that for each  $T \in \mathbb{N}$ , there exist  $(\sigma_0^T, \sigma_2^T) \in B_1(s_1^*)^H$  and a set of histories  $E_T(\zeta) \subset H^\infty$  such that  $\mathbb{Q}^{\sigma_0^T, \sigma_2^T}(E_T(\zeta)) > \zeta$  but  $\mu_T(\cdot|h) \notin M_\zeta$  for all  $h \in E_T(\zeta)$ . Taking a subsequence if necessary and using compactness of  $B_1(s_1^*)^H$ , we have  $(\sigma_0^T, \sigma_2^T) \rightarrow (\sigma_0^\infty, \sigma_2^\infty) \in B_1(s_1^*)^H$ . Since  $\mu_T$  depends only on the history up to time  $T$ , this implies that  $\mathbb{Q}^{\sigma_0^T, \sigma_2^T}(E_T(\zeta)) > \zeta$  for all  $T' \geq T$ . Thus, since  $\mathbb{Q}_{T'}$  is continuous in strategies as a finite-dimensional measure, passing  $(\sigma_0^T, \sigma_2^T)$  to the limit (while fixing the time  $T'$  and the set of histories  $E_T(\zeta)$ ) gives  $\mathbb{Q}^{\sigma_0^\infty, \sigma_2^\infty}(E_T(\zeta)) > \zeta$  for all  $T'$  sufficiently large; and then taking  $T' \rightarrow \infty$  gives  $\mathbb{Q}^{\sigma_0^\infty, \sigma_2^\infty}(E_T(\zeta)) > \zeta$ . As this holds for all  $T$ , we have a sequence of events  $\{E_T(\zeta)\}_{T \in \mathbb{N}}$  such that  $\mathbb{Q}^{\sigma_0^\infty, \sigma_2^\infty}(E_T(\zeta)) > \zeta$  for all  $T$ . From here, we can conclude<sup>26</sup>

$$\mathbb{Q}^{\sigma_0^\infty, \sigma_2^\infty} \left( \limsup_{T \rightarrow \infty} E_T(\zeta) \right) \geq \limsup_{n \rightarrow \infty} \frac{\left( \sum_{k=1}^n \mathbb{Q}^{\sigma_0^\infty, \sigma_2^\infty}(E_k(\zeta)) \right)^2}{\sum_{1 \leq j, k \leq n} \mathbb{Q}^{\sigma_0^\infty, \sigma_2^\infty}(E_j(\zeta) \cap E_k(\zeta))} \geq \frac{n^2 \zeta^2}{n^2} = \zeta^2.$$

Thus, for any history  $h \in \limsup_{T \rightarrow \infty} E_T(\zeta)$ , there is a sequence of times  $\{T_n\}$  such that  $\mu_{T_n}(\cdot|h) \notin M_\zeta$  for all  $n$ . Since  $\zeta^2 < \zeta$ , this implies that  $\mu_{T_n}(\cdot|h) \notin M_{\zeta^2}$ . Thus, for any  $h \in E_\infty(\zeta)$ ,  $\mu_t(\cdot|h) \notin M_{\zeta^2}$  for infinitely many  $T$ ; but  $\mathbb{Q}^{\sigma_0^\infty, \sigma_2^\infty}(E_\infty(\zeta)) > \zeta^2$ . But this contradicts Lemma 9 for the strategies  $(\sigma_0^\infty, \sigma_2^\infty) \in B_1(s_1^*)^H$ , completing the proof.

### B.3 Proof of Lemma 4

Note that if  $h^t \in H_\eta^t$  and  $\mu_t(\cdot|h^t) \in M_0$  then

$$\left| \mu_t(\omega_R|h^t) \right| \left| p(\sigma_0^*(h^t), \sigma_1^*(h^t), \sigma_2^*(h^t)) - p(\sigma_0^*, \sigma_1^*, \sigma_2^*|h^t) \right| < \eta$$

and  $(\sigma_0^*(h^t), \sigma_2^*(h^t)) \in B_\eta(s_1^*)$ . Thus, since  $p(\sigma_0^*(h^t), \sigma_1^*, \sigma_2^*(h^t))$  is continuous in  $\mu_t(\cdot|h^t)$ , there exists a strictly positive function  $\zeta(\eta)$  satisfying  $\lim_{\eta \rightarrow 0} \zeta(\eta) = 0$  such that if  $h^t \in H_\eta^t$  and

<sup>25</sup>Formally,  $H = \bigcup_t H^t$  is isomorphic to the set of finite cylinders that generate  $H^\infty$ .

<sup>26</sup>This is a consequence of the Kochen-Stone Theorem; see, for example, Theorem 1.3 of Arthan and Oliva (2021). We thank Eric Gao for pointing us to this result.

$\mu_t(\cdot|h^t) \in M_{\zeta(\eta)}$  then

$$|\mu_t(\omega_R|h^t)| \left| \left| p(\sigma_0^*(h^t), \sigma_1^*(h^t, \omega_R), \sigma_2^*(h^t)) - p(\sigma_0^*(h^t), s_1^*, \sigma_2^*(h^t)) \right| \right| < 2\eta$$

and  $(\sigma_0^*(h^t), \sigma_2^*(h^t)) \in B_{2\eta}(s_1^*)$ .

Now fix any  $c > 0$ . If  $\mu_t(\omega_R|h^t) \geq c$  then

$$\left| \left| p(\sigma_0^*(h^t), \sigma_1^*(h^t, \omega_R), \sigma_2^*(h^t)) - p(\sigma_0^*(h^t), s_1^*, \sigma_2^*(h^t)) \right| \right| < \frac{2\eta}{c}.$$

Hence, as  $\eta \rightarrow 0$ , Lemma 1 implies that  $\|\sigma_1^*(h^t, \omega_R) - s_1^*\| \rightarrow 0$ , and hence  $(\sigma_0^*(h^t), \sigma_2^*(h^t)) \in \hat{B}_{\xi_1(\eta)}(s_1^*)$  for some strictly positive function  $\xi_1(\eta)$  satisfying  $\xi_1(\eta) \rightarrow 0$ . If instead  $\mu_t(\omega_R|h^t) < c$  then  $\|\sigma_1^*(h^t) - s_1^*\| \leq 1 - \zeta(\eta) - c$ , and hence  $(\sigma_0^*(h^t), \sigma_2^*(h^t)) \in \hat{B}_{\zeta(\eta)+c}(s_1^*)$ . Taking  $\xi(\eta) = \max\{\xi_1(\eta), \zeta(\eta) + c\}$  completes the proof.

## B.4 Proof of Proposition 5

Let  $\gamma$  be feasible in  $\text{OT}(\rho, \phi)$  and  $\text{supp}(\gamma)$  not strictly  $u$ -CM. Let  $\{(x_i, y_i)_{i=1}^N\} \subset \text{supp}(\gamma)$  be a collection of pairs witnessing a violation and set  $\varepsilon = \min_i \gamma((x_i, y_i))$ . Define  $\gamma' \in \Delta(X \times Y)$  by

$$\gamma'(x, y) = \begin{cases} \gamma(x, y) - \varepsilon & \text{if } (x, y) \in \{(x_i, y_i)_{i=1}^N\} \setminus \{(x_i, y_{i+1})_{i=1}^N\}, \\ \gamma(x, y) + \varepsilon & \text{if } (x, y) \in \{(x_i, y_{i+1})_{i=1}^N\} \setminus \{(x_i, y_i)_{i=1}^N\}, \\ \gamma(x, y) & \text{otherwise.} \end{cases}$$

Then  $\gamma' \neq \gamma$  is feasible in  $\text{OT}(\rho, \phi)$  and  $\int u(x, y) d\gamma \leq \int u(x, y) d\gamma'$  (since  $\{(x_i, y_i)_{i=1}^N\}$  witnesses a violation of strict  $u$ -cyclical monotonicity), so  $\gamma$  does not uniquely solve  $\text{OT}(\rho, \phi)$ .

Conversely, if  $\gamma$  is feasible in  $\text{OT}(\rho, \phi)$  and strictly  $u$ -cyclically monotone, consider any  $\gamma' \neq \gamma$  that is feasible in  $\text{OT}(\rho, \phi)$ . Since  $\gamma$  and  $\gamma'$  are both feasible in  $\text{OT}(\rho, \phi)$  and  $\gamma \neq \gamma'$ , there exists  $\{(x_i, y_i)_{i=1}^N\} \subset \text{supp}(\gamma)$  such that  $\{(x_i, y_{i+1})_{i=1}^N\} \subset \text{supp}(\gamma')$ . (To see this, let  $(x_1, y_1)$  be any pair such that  $\gamma(x_1, y_1) > \gamma'(x_1, y_1)$ . Since  $\gamma$  and  $\gamma'$  transport the same mass into  $y_1$ , there exists  $x_2$  such that  $\gamma(x_2, y_1) < \gamma'(x_2, y_1)$ . But now, since  $\gamma$  and  $\gamma'$  transport the same mass out of  $x_2$ , there exists  $y_2$  such that  $\gamma(x_2, y_2) > \gamma'(x_2, y_2)$ . Continuing in this

manner and using finiteness of  $X \times Y$  yields a cycle.) Let  $\varepsilon = \min_i \gamma'(x_i, y_{i+1})$ , and let

$$\gamma''(x, y) = \begin{cases} \gamma'(x, y) - \varepsilon & \text{if } (x, y) \in \{(x_i, y_{i+1})_{i=1}^N\} \setminus \{(x_i, y_i)_{i=1}^N\}, \\ \gamma'(x, y) + \varepsilon & \text{if } (x, y) \in \{(x_i, y_i)_{i=1}^N\} \setminus \{(x_i, y_{i+1})_{i=1}^N\}, \\ \gamma'(x, y) & \text{otherwise.} \end{cases}$$

Then  $\gamma''$  is feasible in  $\text{OT}(\rho, \phi)$  and  $\int u(x, y) d\gamma'' > \int u(x, y) d\gamma'$  (since  $\{(x_i, y_i)_{i=1}^N\}$  is contained in the strictly  $u$ -cyclically monotone set  $\text{supp}(\gamma)$ ), so  $\gamma'$  does not solve  $\text{OT}(\rho, \phi)$ . Thus, since no  $\gamma' \neq \gamma$  solves  $\text{OT}(\rho, \phi)$ , and  $\text{OT}(\rho, \phi)$  has a solution as a continuous maximization problem over a compact set,  $\gamma$  must uniquely solve  $\text{OT}(\rho, \phi)$ .

## B.5 Proof of Proposition 6

The result is obvious if  $\bar{v}_1^{\text{CM}} \geq \bar{u}_1$ , so suppose  $\bar{v}_1^{\text{CM}} < \bar{u}_1$ , and fix  $\varepsilon < 2(\bar{u}_1 - \bar{v}_1^{\text{CM}})$ , a prior  $\mu_0$  with  $\mu_0(\omega_R) > 0$ , and an equilibrium  $(\sigma_0^*, \sigma_1^*, \sigma_2^*)$ . We start with an additional lemma.

**Lemma 10.** *For all  $\xi > 0$ , there exists  $\varsigma > 0$  such that, for any  $u_1$ -cyclically monotone strategy  $s_1$ , any strategy  $s'_1$  satisfying  $\|s_1 - s'_1\| < \varsigma$ , and any  $(\alpha_0, \alpha_2) \in B(s_1)$ , we have  $u_1(\alpha_0, s_1, \alpha_2) < \bar{v}_1^{\text{CM}} + \xi$ .*

*Proof.* Suppose for contradiction that there exists  $\varepsilon > 0$  and a sequence of strategies  $\tilde{s}_1^n$ , each within  $1/n$  of a  $u_1$ -cyclically monotone strategy  $s_1^n$ , and  $(\alpha_0^n, \alpha_2^n) \in B(\tilde{s}_1^n)$  such that  $u_1(\alpha_0^n, s_1^n, \alpha_2^n) > \bar{v}_1^{\text{CM}} + \xi$ . Taking a subsequence if necessary and noting that the set of  $u_1$ -cyclically monotone strategies is closed (as they are characterized by their support),  $s_1^n$  converges to a  $u_1$ -cyclically monotone strategy  $s_1$ . Moreover,  $\tilde{s}_1^n$  also converges to  $s_1$ , by the triangle inequality. But this yields a contradiction, as

$$\bar{v}_1^{\text{CM}} + \xi \leq \limsup_{n \rightarrow \infty} u_1(\alpha_0^n, \tilde{s}_1^n, \alpha_2^n) \leq \sup_{(\alpha_0, \alpha_2) \in B(s_1)} u_1(\alpha_0, s_1, \alpha_2) \leq \bar{v}_1^{\text{CM}},$$

where the second inequality follows because  $B(\cdot)$  is upper hemi-continuous and  $u_1$  is continuous, and the third inequality follows because  $s_1$  is  $u_1$ -cyclically monotone.  $\square$

Now, note that at any history  $h^t$  where  $\mu_t(\omega_R|h^t) > 1 - \varsigma$ , we have  $\|\sigma_1^*(h^t) - \sigma_1^*(h^t, \omega_R)\| < \varsigma$ . Thus, by Lemmas 5 and 10, there exists  $\varsigma > 0$  such that at any history  $h^t$  where  $\mu_t(\omega_R|h^t) >$



$1 - \varsigma$ , we have  $u_1(\sigma_0^*, \sigma_1^*, \sigma_2^* | h^t, \omega_R) < \bar{v}_1^{\text{CM}} + \varepsilon/2$ . Since  $\mu_t(\omega_R | h^t, \omega_R)$  is a  $\mathbb{P}$ -submartingale,

$$(1 - \mathbb{P}(\mu_t(\omega_R | h^t, \omega_R) > 1 - \varsigma))(1 - \varsigma) + \mathbb{P}(\mu_t(\omega_R | h^t, \omega_R) > 1 - \varsigma)(1) \geq \mu_0(\omega_R) \iff \mathbb{P}(\mu_t(\omega_R | h^t, \omega_R) > 1 - \varsigma) \geq 1 - \frac{1 - \mu_0(\omega_R)}{\varsigma}.$$

Therefore, the long-run player's expected payoff in each period  $t$  is at most

$$\left(1 - \frac{1 - \mu_0(\omega_R)}{\varsigma}\right) \left(\bar{v}_1^{\text{CM}} + \frac{\varepsilon}{2}\right) + \frac{1 - \mu_0(\omega_R)}{\varsigma} \bar{u}_1.$$

This payoff is less than  $\bar{v}_1^{\text{CM}} + \varepsilon$  whenever

$$\mu_0(\omega_R) > 1 - \frac{\varepsilon \varsigma}{2(\bar{u}_1 - \bar{v}_1^{\text{CM}}) - \varepsilon},$$

completing the proof.

## B.6 Proof of Proposition 9

Suppose that  $s_1 : \Theta \rightarrow \Delta(R)$  is monotone with respect to  $(\succsim_\Theta, \succsim_R)$ .

First,  $G(s_1)$  cannot contain a cycle  $(\theta_1, r_1), (\theta_2, r_1), \dots, (\theta_K, r_K), (\theta_1, r_K)$ . To see this, suppose otherwise, and let  $\theta_1 \prec \dots \prec \theta_K$ , without loss. Since  $r_k \in \text{supp}(s_1(\theta_k)) \cap \text{supp}(s_1(\theta_{k+1}))$  for  $k \in \{1, \dots, K-1\}$  and  $r_K \in \text{supp}(s_1(\theta_K))$ , monotonicity requires  $r_1 \prec \dots \prec r_K$ . But this gives a contradiction, since  $r_K \in \text{supp}(s_1(\theta_1))$  and  $r_1 \in \text{supp}(s_1(\theta_2))$ .

Next,  $G(s_1)$  cannot contain a forbidden triple. We have already explained why it cannot contain a Type (1) forbidden triple. The argument for why it cannot contain a Type (2) forbidden triple is identical, with the roles of states and actions interchanged.

Conversely, suppose that  $G(s_1)$  is acyclic and does not contain a forbidden triple. It suffices to consider the case where  $G(s_1)$  is connected, as otherwise the orders on the states and actions in each connected component of  $G(s_1)$  can be appended to one another. So suppose that  $G(s_1)$  is connected, and let  $(\theta_1, r_1), (\theta_2, r_1), \dots, (\theta_K, r_K)$  be any maximum path in  $G(s_1)$  (i.e., any path of maximum length), supposing that such a path ends with an action. (The argument for the case where all maximum paths both start and end at actions or at states is almost identical.) Define the orders  $\prec_\Theta$  on  $\{\theta_1, \dots, \theta_K\}$  and  $\prec_R$  on  $\{r_1, \dots, r_K\}$  by  $\theta_1 \prec_\Theta \dots \prec_\Theta \theta_K$  and  $r_1 \prec_R \dots \prec_R r_K$ .

We claim that for any  $\theta \in \Theta \setminus \{\theta_1, \dots, \theta_K\}$ , there exists  $k \in \{1, \dots, K-1\}$  such that

$\text{supp}(s_1(\theta)) = \{r_k\}$ . Note such a state  $\theta$  is linked to at most one  $r_k \in \{r_1, \dots, r_{K-1}\}$ , as if  $\theta$  is linked to distinct  $r_k, r'_k$  then appending  $\theta$  to both ends of the path from  $r_k$  to  $r'_k$  forms a cycle. In addition,  $\theta$  cannot be linked to  $r_K$ , as then it could be appended to the maximum path. Finally,  $\theta$  cannot be linked to both some  $r_k \in \{r_1, \dots, r_{K-1}\}$  and some  $r \in R \setminus \{r_1, \dots, r_{K-1}\}$ . For, if  $k = 1$  then replacing  $(\theta_1, r_1)$  with  $(\theta, r), (\theta, r_1)$  at the beginning of the maximum path would lengthen it; and if  $k \geq 2$  then the set of states  $\{\theta_k, \theta, \theta_{k+1}\}$  together with the set of actions  $\{r, r_{k-1}, r_k, r_{k+1}\}$  would be a forbidden triple, as  $\{r_{k-1}, r_k\} \in \text{supp}(s_1(\theta_k))$ ,  $\{r, r_k\} \in \text{supp}(s_1(\theta))$ , and  $\{r_k, r_{k+1}\} \in \text{supp}(s_1(\theta_{k+1}))$  (see Figure 2).

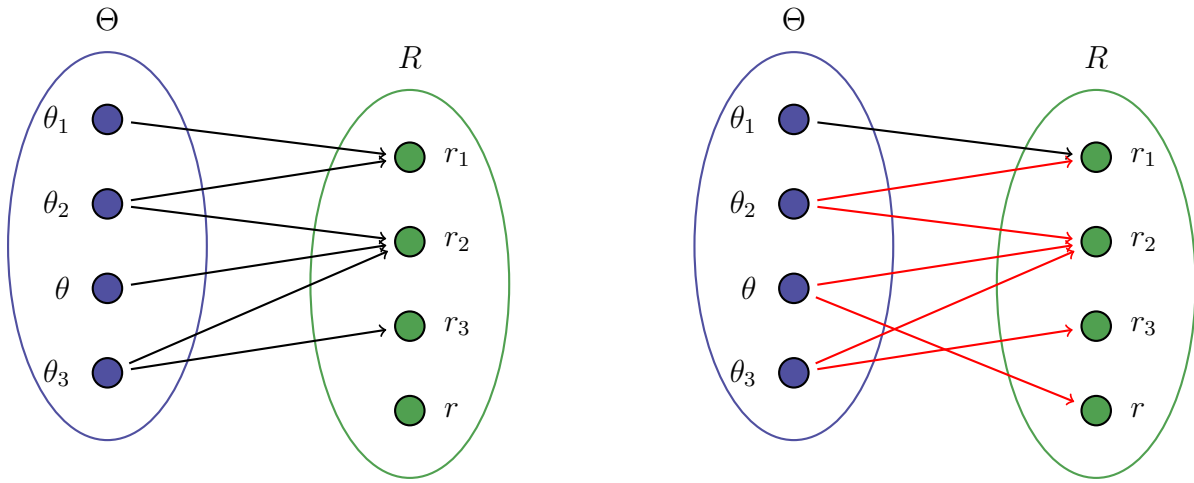


Figure 2: Each State  $\theta \notin \{\theta_1, \dots, \theta_K\}$  Has Only One Neighbor

Notes. If  $\theta \notin \{\theta_1, \dots, \theta_K\}$  is linked to  $r_k \in \{r_1, \dots, r_K\}$  and  $r \notin \{r_1, \dots, r_K\}$ , then  $\{\theta_k, \theta, \theta_{k+1}\}$  together with  $\{r, r_{k-1}, r_k, r_{k+1}\}$  is a forbidden triple.

Given this claim, we can extend  $\prec_\Theta$  to  $\Theta$  by ordering each  $\theta \in \Theta \setminus \{\theta_1, \dots, \theta_K\}$  such that  $\text{supp}(s_1(\theta)) = \{r_k\}$  in between  $\theta_k$  and  $\theta_{k+1}$  (and ordering multiple such states arbitrarily between  $\theta_k$  and  $\theta_{k+1}$ ).

Similarly, for any  $r \in R \setminus \{r_1, \dots, r_K\}$ , there exists  $k \in \{2, \dots, K\}$  such that  $r_k$  is linked only to  $\theta_k$  in  $G(s_1)$ . Extend  $\prec_R$  to  $R$  by ordering each such  $r$  in between  $r_{k-1}$  and  $r_k$ .

Note that for any  $k \geq 2$  and any  $r \in \text{supp}(s_1(\theta_k))$ , we have  $r_{k-1} \prec_R r \prec_R r_k$ . This follows because if  $r \notin \{r_1, \dots, r_K\}$  then  $r_{k-1} \prec_R r \prec_R r_k$  by construction, and if  $r = r_{\tilde{k}}$  for some  $\tilde{k} \notin \{k-1, k\}$ , then  $G(s_1)$  contains a cycle starting with  $(\theta_k, r_{\tilde{k}})$  and then following the maximum path back to  $\theta_k$ .

Finally, we claim that  $s_1$  is monotone with respect to  $(\prec_\Theta, \prec_R)$ . To see this, fix any  $\theta \succ_\Theta \theta'$ . Let  $\tilde{k} = \max\{k : \theta \prec_\Theta \theta_k\}$ . If  $\theta \succ_\Theta \theta_{\tilde{k}} \prec_\Theta \theta'$ , then  $\text{supp}(s_1(\theta)) = \{r_{\tilde{k}}\}$  and  $r_{\tilde{k}} \prec_R r$

for all  $r \in \text{supp}(s_1(\theta'))$ . If  $\theta = \theta_{\bar{k}} \succ_{\Theta} \theta'$ , then  $r_{\theta_{\bar{k}-1}}$  is the lowest action in  $\text{supp}(s_1(\theta))$ , and no action in  $\text{supp}(s_1(\theta'))$  is above  $r_{\theta_{\bar{k}-1}}$ . Lastly, if  $\theta \succ_{\Theta} \theta' \succsim_{\Theta} \theta_{\bar{k}}$ , then  $\text{supp}(s_1(\theta)) = \text{supp}(s_1(\theta')) = \{r_{\bar{k}}\}$ . Thus, in all cases, monotonicity holds.

## References

- Acemoglu, Daron, and Alexander Wolitzky.** 2024. “Mistrust, Misperception, and Misunderstanding: Imperfect Information and Conflict Dynamics.” *Handbook of the Economics of Conflict*.
- Arieli, Itai, Yakov Babichenko, Rann Smorodinsky, and Takuro Yamashita.** 2023. “Optimal Persuasion via Bi-Pooling.” *Theoretical Economics* 18 (1): 15–36.
- Arthan, Rob, and Paulo Oliva.** 2021. “On the Borel-Cantelli Lemmas, the Erdős–Rényi Theorem, and the Kochen-Stone Theorem.” *Journal of Logic and Analysis* 13 (6): .
- Avenhaus, Rudolf, Bernhard von Stengel, and Shmuel Zamir.** 2002. “Inspection Games.” In *Handbook of Game Theory with Economic Applications*, Volume 3. Chap. 51 1947–1987, Elsevier.
- Ball, Ian, and Deniz Kattwinkel.** 2024. “Quota Mechanisms: Finite-Sample Optimality and Robustness.” *Working Paper*.
- Battigalli, Pierpaolo, and Joel Watson.** 1997. “On “reputation” refinements with heterogeneous beliefs.” *Econometrica* 369–374.
- Best, James, and Daniel Quigley.** 2024. “Persuasion for the Long Run.” *Journal of Political Economy* 132 (5): 1305–1337.
- Chakraborty, Archishman, and Rick Harbaugh.** 2010. “Persuasion by Cheap Talk.” *American Economic Review* 100 (5): 2361–2382.
- Chen, Ying, Navin Kartik, and Joel Sobel.** 2008. “Selecting cheap-talk equilibria.” *Econometrica* 76 (1): 117–136.
- Clark, Daniel, Drew Fudenberg, and Alexander Wolitzky.** 2021. “Record-Keeping and Cooperation in Large Societies.” *The Review of Economic Studies* 88 (5): 2179–2209.

- Dworczak, Piotr, and Giorgio Martini.** 2019. “The simple economics of optimal persuasion.” *Journal of Political Economy* 127 (5): 1993–2048.
- Ely, Jeffrey, Drew Fudenberg, and David K Levine.** 2008. “When is reputation bad?” *Games and Economic Behavior* 63 (2): 498–526.
- Ely, Jeffrey, and Juuso Valimaki.** 2003. “Bad Reputation.” *Quarterly Journal of Economics* 118 (3): 785–814.
- Escobar, Juan F., and Juuso Toikka.** 2013. “Efficiency in Games with Markovian Private Information.” *Econometrica* 81 (5): 1737–1767.
- Frankel, Alexander.** 2014. “Aligned Delegation.” *American Economic Review* 104 (1): 66–83.
- Fudenberg, Drew, Ying Gao, and Harry Pei.** 2022. “A Reputation for Honesty.” *Journal of Economic Theory* 204 105508.
- Fudenberg, Drew, and David Levine.** 1989. “Reputation and Equilibrium Selection in Games with a Patient Player.” *Econometrica* 57 (4): 759–778.
- Fudenberg, Drew, and David Levine.** 1992. “Maintaining a Reputation when Strategies are Imperfectly Observed.” *Review of Economic Studies* 59 (3): 561–579.
- Gossner, Olivier.** 2011. “Simple Bounds on the Value of a Reputation.” *Econometrica* 79 (5): 1627–1641.
- Heller, Yuval, and Erik Mohlin.** 2018. “Observations on Cooperation.” *The Review of Economic Studies* 85 (4): 2253–2282.
- Jackson, Matthew O., and Hugo F. Sonnenschein.** 2007. “Overcoming Incentive Constraints by Linking Decisions.” *Econometrica* 75 (1): 241–257.
- Kamenica, Emir, and Matthew Gentzkow.** 2011. “Bayesian Persuasion.” *American Economic Review* 101 (6): 2590–2615.
- Kartik, Navin.** 2009. “Strategic communication with lying costs.” *The Review of Economic Studies* 76 (4): 1359–1395.
- Kleiner, Andreas, Benny Moldovanu, and Philipp Strack.** 2021. “Extreme Points and Majorization: Economic Applications.” *Econometrica* 89 (5): 1671–1700.

- Kolotilin, Anton.** 2018. “Optimal information disclosure: A linear programming approach.” *Theoretical Economics* 13 (2): 607–635.
- Kuvalekar, Aditya, Elliot Lipnowski, and Joao Ramos.** 2022. “Goodwill in communication.” *Journal of Economic Theory* 203 105467.
- Lin, Xiao, and Ce Liu.** 2024. “Credible Persuasion.” *Journal of Political Economy* 132 (7): .
- Liu, Qingmin.** 2011. “Information acquisition and reputation dynamics.” *The Review of Economic Studies* 78 (4): 1400–1425.
- Liu, Qingmin, and Andrzej Skrzypacz.** 2014. “Limited Records and Reputation Bubbles.” *Journal of Economic Theory* 151 2–29.
- Mailath, George J., and Larry Samuelson.** 2006. *Repeated Games and Reputations: Long-Run Relationships*. Oxford University Press.
- Margaria, Chiara, and Alex Smolin.** 2018. “Dynamic Communication with Biased Senders.” *Games and Economic Behavior* 110 330–339.
- Mathevet, Laurent, David Pearce, and Ennio Stacchetti.** 2024. “Reputation and Information Design.” *Working Paper*.
- Matsushima, Hitoshi, Koichi Miyazaki, and Nobuyuki Yagi.** 2010. “Role of Linking Mechanisms in Multitask Agency with Hidden Information.” *Journal of Economic Theory* 145 (6): 2241–2259.
- Meng, Delong.** 2021. “On the value of repetition for communication games.” *Games and Economic Behavior* 127 227–246.
- Morris, Stephen.** 2001. “Political correctness.” *Journal of Political Economy* 109 (2): 231–265.
- Myerson, Roger B.** 1982. “Optimal coordination mechanisms in generalized principal–agent problems.” *Journal of mathematical economics* 10 (1): 67–81.
- Pei, Harry.** 2020. “Reputation Effects Under Interdependent Values.” *Econometrica* 88 (5): 1671–1700.

- Pei, Harry.** 2023. “Repeated Communication with Private Lying Costs.” *Journal of Economic Theory* 210 105668.
- Pei, Harry.** 2024. “Reputation Effects under Short Memories.” *Journal of Political Economy* 132 (10): .
- Rahman, David M.** 2024. “Detecting Profitable Deviations.” *Journal of Mathematical Economics* 111 102946.
- Rayo, Luis, and Ilya Segal.** 2010. “Optimal Information Disclosure.” *Journal of Political Economy* 118 (5): 949–987.
- Renault, Jérôme, Eilon Solan, and Nicolas Vieille.** 2013. “Dynamic Sender–Receiver Games.” *Journal of Economic Theory* 148 (2): 502–534.
- Rochet, Jean-Charles.** 1987. “A Necessary and Sufficient Condition for Rationalizability in a Quasi-Linear Context.” *Journal of Mathematical Economics* 16 (2): 191–200.
- Santambrogio, Filippo.** 2015. “Optimal transport for applied mathematicians.” *Birkhäuser, NY* 55 (58-63): 94.
- Schelling, Thomas C.** 1966. *Arms and Influence*. Chap. 1 74, The Henry L. Stimson Lectures Series, Yale University Press.
- Sorin, Sylvain.** 1999. “Merging, Reputation, and Repeated Games with Incomplete Information.” *Games and Economic Behavior* 29 274–308.
- Spence, Michael.** 1973. “Job Market Signaling.” *The Quarterly Journal of Economics* 87 (3): 355–374.
- Takahashi, Satoru.** 2010. “Community enforcement when players observe partners’ past play.” *Journal of Economic Theory* 145 (1): 42–62.
- Watson, Joel.** 1993. “A “reputation” refinement without equilibrium.” *Econometrica* 199–205.