Econometrica, Vol. 93, No. 6 (November, 2025), 2007–2042

MARGINAL REPUTATION

DANIEL LUO
Department of Economics, MIT

ALEXANDER WOLITZKY
Department of Economics, MIT

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. Applications include deterrence, delegation, signaling, and persuasion. 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, deterrence, delegation, signaling, Bayesian persuasion.

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, delegation, 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

Daniel Luo: daniel57@mit.edu
Alexander Wolitzky: wolitzky@mit.edu

For helpful comments, we thank seminar participants at ESWC, Northwestern, SITE, UT Austin, and VSET; and Sandeep Baliga, Ian Ball, V. Bhaskar, Drew Fudenberg, Eric Gao, Navin Kartik, Anton Kolotilin, Stephen Morris, David Pearce, Harry Pei, and Tom Wiseman. Funding Statement: Open Access funding enabled and organized by Massachusetts Institute of Technology WOA Institution: Massachusetts Institute of Technology Consortia Name: MIT Hybrid 2025

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 came in 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 nontrivial 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—meaning that they are both best responses to some long-run player strategy that induces the same signal distribution as the Stackelberg strategy—which implies that the payoff lower bound of Fudenberg and Levine (1992) 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. (Our theorem also requires that the Stackelberg strategy is not confounded by a different commitment type of the long-run player—we discuss this second condition below.)

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 strictly supermodular in one-dimensional signals and actions, a strategy is confound-defeating if and only if every selection from the graph of 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. Combining these results gives a unique payoff prediction in many games when the long-run player is patient and is rational with high probability.

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 strictly supermodular in her own (one-dimensional) type and action, a patient sender can secure her best commitment payoff from any monotone signaling strategy.[1] This follows because these conditions imply that the sender's payoff is strictly supermodular in her own type and action for any receiver strategy, so our results for one-dimensional supermodular games apply.

We then consider repeated cheap talk games with state-independent sender preferences. Our results do not apply to cheap talk games, because when the sender's "action" is a payoff-irrelevant message, her payoff cannot be strictly supermodular in this action and her type. However, our results imply that perturbing the sender's payoff by adding a small, strictly submodular "lying cost" can provide 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 can thus provide a reputational foundation for a large class of communication mechanisms, even when the history of realized states is unobserved.

Our main result 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 that our result goes through when the Stackelberg type is *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 much more 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."

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 secure her commitment payoff against the least favorable of

[1] 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.

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 for a patient long-run player's payoff, which Ely and Valimaki (2003) show are much too loose in a class of repeated delegation games. 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.[2]

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.[3] Other papers in the reputation literature where supermodularity 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, 2007, 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 via cheap talk when the marginal distribution over messages is observed. In their setting, a joint distribution over states and receiver 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. payoff-relevant sender actions) and come from different strategic considerations (static cheap talk subject to a "credibility" constraint vs. long-run reputation formation).[4]

We are also not the first to discuss reputation foundations for the Bayesian persuasion commitment assumption. This assumption has been controversial ever since its introduction by Kamenica and Gentzkow (2011) and Rayo and Segal (2010), who suggested reputation as a possible foundation. Mathevet, Pearce, and Stacchetti (2024) observe that reputation effects yield the commitment payoff when a long-run sender faces a sequence of

[2]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, Fudenberg, and Levine (2008) it is not identified when the short-run player exits; see Section 6.2.

[3]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.

[4]Renault, Solan, and Vieille (2013) characterize the set of equilibrium payoffs in repeated cheap talk games with patient players via an optimality condition over joint distributions of states and receiver actions similar to that in Lin and Liu (2024).

short-run receivers who observe the history of messages *and states*, and 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 two examples, including the deterrence game discussed above. 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 and not behaviorally confounded. Section 5 characterizes confound-defeatingness via cyclical monotonicity and gives an upper bound for the long-run player's payoff. Section 6 applies our results to one-dimensional supermodular games. Section 7 considers communication games. Section 8 extends our results to allow indistinguishable commitment types by introducing our salience notion. Section 9 summarizes the paper and discusses some possible extensions.

2. MOTIVATING EXAMPLES

2.1. *Deterrence With Private Signals*

We begin with the example of a simple deterrence game.[5] A long-run player with discount factor δ 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, drawn with conditional probability $\Pr(c|C) = \Pr(d|D) = p \in (1/2, 1)$, 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:[6]

	C	D
A	1	$1+g$
F	$-l$	0

Short-run player payoff

	c	d
A	1	y
F	x	0

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

[5]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.

[6]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 mixes 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.

the short-run player strictly prefers to take C if the long-run player plays the "deterrence" strategy, *(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.[7] Finally, assume that the long-run player is committed to (A, F) with probability μ_0 and is rational otherwise.[8]

What can be said about the equilibria of this game when the discount factor δ 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). 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 $\delta \to 1$, 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.[9]

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* $\liminf_{\delta \to 1} \underline{U}_1(\delta) \geq p$ *for all* $\mu_0 > 0$.

(2) *If* $x + y > 1$, *then* $\lim_{\mu_0 \to 0} \overline{U}_1(\delta) = 1 - p + py$ *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.

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 $A \succ F$, $c \succ d$) 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.[10]

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

[7]These two conditions holds iff $p > \max\{\frac{1+g+l}{2+g+l}, \frac{1}{2-y}\}$.

[8]The role of this assumption in the context of the current example in discussed in Sections 4.2 and 8.

[9]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.

[10]This also relies on the assuming that the Stackelberg strategy is not confounded by another commitment type. For example, if there is a second commitment type that fights with probability p after each signal, the conclusion of Proposition 1.1 does not hold.

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.[11]

2.2. *A Repeated Trust Game*

For another example, 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(ood), B(ad)}, drawn i.i.d. across periods with equal probability on each state. In each period, the long-run player observes θ before taking an action $a_1 \in$ {H(igh Effort), L(ow Effort)}. Simultaneously (and having observed the history of past actions, but not states), the short-run player takes an action $a_2 \in$ {T(rust), N(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
H	$1, 2$	$-1, 0$
L	$2, -1$	$0, 0$

Payoffs in state $\theta = G$

	T	N
H	$1 - w, -1$	$-1 - z, 0$
L	$2, -1$	$0, 0$

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, θ is the quality of the day's catch, and player 2 is a customer who wants to eat only fish that is 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 μ_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* $\liminf_{\delta \to 1} \underline{U}_1(\delta) \geq 3/2$ *for all* $\mu_0 > 0$.
(2) *If* $\max\{w, z\} < 0$, *then* $\lim_{\mu_0 \to 0} \overline{U}_1(\delta) = 0$ *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 (θ, 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 (θ, a_1) for any a_2, which we show limits her to her minmax payoff when the prior commitment probability is small.[12]

[11]The submodular case of Proposition 1 follows from Corollary 3 in Section 6.

[12]If one of $\{w, z\}$ is positive and the other negative, the long-run player's payoff is supermodular in (θ, 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 θ 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).

3. MODEL

We study 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, delegation, and communication games, the first short-run player is Nature. In addition, in communication games, 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. There are also two finite signal sets, Y_0 and Y_1, with generic elements y_0 and y_1.

The stage game timing is as follows:

(1) Player 0 takes an action a_0. This generates a signal y_0, drawn from a distribution $\rho_0(\cdot|a_0)$, which is observed by player 1 only.

(2) Players 1 and 2 simultaneously take actions a_1 and a_2. This generates a signal y_1, drawn from a distribution $\rho_1(\cdot|a_1, a_2)$, which is 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 $s_1 : Y_0 \to \Delta(A_1)$. 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 α_2) over player 1's private signal y_0 and action a_1 according to

$$\gamma(\alpha_0, s_1)[y_0, a_1] = \sum_{a_0 \in A_0} \alpha_0(a_0)\rho_0(y_0|a_0)s_1(y_0)[a_1],$$

and induces a joint distribution $p(\alpha_0, s_1, \alpha_2) \in \Delta(Y_1)$ over the public signal y_1 according to

$$p(\alpha_0, s_1, \alpha_2)[y_1] = \sum_{a_0 \in A_0} \sum_{y_0 \in Y_0} \sum_{a_1 \in A_1} \sum_{a_2 \in A_2} \alpha_0(a_0)\rho_0(y_0|a_0)s_1(y_0)[a_1]\alpha_2(a_2)\rho_1(y_1|a_1, a_2).$$

We maintain the following assumption on signals.

ASSUMPTION 1:

(1) *The distribution of the signal y_0 has full support*: $\rho_0(y_0|a_0) > 0$ *for all* $y_0 \in Y_0$, $a_0 \in A_0$.

(2) *The support of the distribution of the signal y_1 is independent of a_2*: $\rho_1(y_1|a_1, a_2) > 0 \implies \rho_1(y_1|a_1, a_2') > 0$ *for all* $y_1 \in Y_1$, $a_1 \in A_1$, $a_2 \neq a_2' \in A_2$.[13]

(3) *The signal y_1 statistically identifies the long-run player's action*: *for any* $a_2 \in A_2$, *the* $|A_1|$ *vectors* $\rho_1(\cdot|a_1, a_2)_{a_1 \in A_1}$ *are linearly independent in* $\mathbb{R}^{|Y_1|}$.

[13]This assumption rules out perfect monitoring of player 2's action. However, this is unnecessary: the analysis is unaffected by the introduction of an arbitrary additional public signal y_2, independent of (y_0, y_1) conditional on (a_0, a_1, a_2), whose distribution $\rho_2(y_2|a_2)$ depends only on a_2.

However, we emphasize that, while the public signal y_1 identifies player 1's action a_1, it does 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. We also denote the marginal of $\gamma(\alpha_0, s_1)$ over Y_0 (which depends only on α_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 $u_0 : A_0 \times A_1 \to \mathbb{R}$ for player 0 and $u_i : Y_0 \times A_1 \times A_2 \to \mathbb{R}$ 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.[14] Finally, in a slight abuse of notation, we also 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 and delegation games fit by making A_0 a singleton (i.e., making player 0 Nature). Communication games fit by making A_0 a singleton; viewing $\rho_0(y_0)$ as the prior distribution of a payoff-relevant state y_0; letting $\rho(y_1|a_1, a_2) = \mathbf{1}(\{y_1 = a_1\})$ (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 (α_0, α_2) satisfying

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

so that player 0 best responds to s_1, and player 2 best responds to α_0 and s_1. With this 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

$$V(s_1) = \inf_{(\alpha_0, \alpha_2) \in B(s_1)} u_1(\alpha_0, s_1, \alpha_2).^{15}$$

Finally, we employ the following definition.[16]

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 *η-confirmed best response* to s_1 if there exists s_1' such that:

[14] However, none of our results requires this assumption, with the exception of Theorem 2 in Section 8.

[15] Recall that the *lower* (resp., *upper*) commitment payoff results from the *least favorable* (resp., *most favorable*) short-run best response.

[16] Here and throughout the paper, $||\cdot||$ denotes the sup norm.

(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 *η-confirms* (α_0, α_2) *against* s_1.

Let $B_\eta(s_1)$ be the set of η-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. Since $B_1(s_1)$ does not depend on s_1, we abbreviate it to B_1. In addition, $B_\eta(s_1) \downarrow B_0(s_1)$ as $\eta \to 0$, 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 $t = 0, 1, 2, \ldots$. 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 replies after observing only the public history of signals. A period-t public history is denoted $h^t = (y_{1,t'})_{t'=0}^{t-1} \in Y_1^t$. Let H^t be the set of period-t (public) histories, $H = \bigcup_t H^t$ the set of all finite histories, and $H^\infty = Y_1^\infty$ the set of infinite histories. A repeated game strategy σ_i for player i maps public histories to stage game strategies: formally, σ_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 $i = 1$.[17]

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 ω_{s_1} plays s_1 in every period.[18] The type ω 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 (rational) 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, σ_1^* denotes the equilibrium strategy for the rational long-run player, and we let $\bar{\sigma}_1^*$ denote the corresponding unconditional long-run player strategy, averaging over Ω by updating the prior μ_0 by Bayes' rule. Note that $\bar{\sigma}_1^*(h^t)$ is defined only for histories h^t that arise on path for some long-run player strategy. Finally, given a strategy profile $(\sigma_0^*, \sigma_1^*, \sigma_2^*)$, we let $\mathbb{P} \in \Delta(H^\infty)$ denote the induced measure over infinite histories conditional on $\omega = \omega_R$, we let $\overline{\mathbb{P}}$ denote the corresponding unconditional measure (averaging over Ω), and we let $\mathbb{P}_t$ and $\overline{\mathbb{P}}_t$ denote the corresponding projections on H^t. Note that the set of on-path period-t histories under rational play is $\text{supp}(\mathbb{P}_t)$, and that the set of period-t histories where $\bar{\sigma}_1^*(h^t)$ is well-defined is the larger set $\text{supp}(\overline{\mathbb{P}}_t) \supset \text{supp}(\mathbb{P}_t)$.

[17]In principle, the long-run player could condition on her past private signals 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.

[18]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 9.

The key prior result in this context is the following.[19]

THEOREM 0—Fudenberg and Levine (1992): *For any strategy s_1^*, if $\omega_{s_1^*} \in \Omega$ then*

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

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.

4. THE COMMITMENT PAYOFF THEOREM

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

4.1. *The Confound-Defeating Property*

We give two equivalent definitions of the confound-defeating property. 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 5–7). 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 the optimal transport problem

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

DEFINITION 2: 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 $\mathrm{OT}(\rho(\alpha_0), \phi(\alpha_0, s_1^*); \alpha_2)$.

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 η-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.

PROPOSITION 3: *The two definitions of confound-defeatingness are equivalent.*

[19] Technically, our model is not a special case of Fudenberg and Levine's because they assume a single short-run opponent in each period, but the same argument applies. Other than this difference, our model specializes theirs to the case where the long-run player's strategy is a function $s_1 : Y_0 \to \Delta(A_1)$ and Assumption 1 holds.

It is immediate that if $\gamma(\alpha_0, s_1^*)$ is not the unique solution to $\mathrm{OT}(\rho(\alpha_0), \phi(\alpha_0, s_1^*); \alpha_2)$ then the first definition of confound-defeatingness cannot hold. The converse relies on Assumption 1(3) (player 1's action is identified). Without identification, the first definition of confound-defeatingness still gives a commitment payoff theorem, but it is more difficult to check and does 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 on-path history under rational play 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^*. 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 $h^t \in \mathrm{supp}(\mathbb{P}_t)$ where:*
(1) $\|p(\sigma_0^*, \bar{\sigma}_1^*, \sigma_2^*|h^t) - p(\sigma_0^*, s_1^*, \sigma_2^*|h^t)\| < \eta$, *and*
(2) $\|p(\sigma_0^*, \sigma_1^*, \sigma_2^*|h^t) - p(\sigma_0^*, s_1^*, \sigma_2^*|h^t)\| < \eta$,
we have $\|\sigma_1^*(h^t) - s_1^*\| \leq \varepsilon$.

PROOF: Suppose not, so there exists a history $h^t \in \mathrm{supp}(\mathbb{P}_t)$ where conditions (1) and (2) hold, but $\|\sigma_1^*(h^t) - s_1^*\| > \varepsilon$. 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 η-confirmed best reply to s_1^*, as $\bar{\sigma}_1^*(h^t)$ η-confirms it against s_1^*. Hence, condition (2) along with $\|\sigma_1^*(h^t) - s_1^*\| > \varepsilon$ and confound-defeatingness imply that there exists some strategy $\tilde{s}_1$ such that $p(\sigma_0^*, \tilde{s}_1, \sigma_2^*|h^t) = p(\sigma_0^*, \sigma_1^*, \sigma_2^*|h^t)$ and $u_1(\sigma_0^*, \tilde{s}_1, \sigma_2^*|h^t) > u_1(\sigma_0^*, \sigma_1^*, \sigma_2^*|h^t)$. But this implies that if the long-run player deviates from $\sigma_1^*(h^t)$ to $\tilde{s}_1$ at h^t, her continuation payoff is unchanged while her stage game payoff increases. So, since $h^t \in \mathrm{supp}(\mathbb{P}_t)$, this deviation is strictly profitable, contradicting the assumption that $(\sigma_0^*, \sigma_1^*, \sigma_2^*)$ is an equilibrium. *Q.E.D.*

4.2. *Behavioral Confounding*

Lemma 1 implies that if a strategy s_1^* is confound-defeating, it will not be confounded by the equilibrium play of the rational type of player 1. However, there remains the possibility of confounding by behavioral types. To address this, we make use of the following definition.

DEFINITION 3: Strategy s_1^* is *not behaviorally confounded* if, for any $\omega_{s_1'} \in \Omega$ such that $s_1' \neq s_1^*$ and any $(\alpha_0, \alpha_2) \in B_1$, we have $p(\alpha_0, s_1^*, \alpha_2) \neq p(\alpha_0, s_1', \alpha_2)$.

A strategy is not behaviorally confounded if the public signal distinguishes it from any other commitment type, whenever the short-run players take actions that best respond to some long-run player strategy. The definition allows the possibility that s_1 is indistinguishable from a mixture of two commitment types $\omega_{s_1'}, \omega_{s_1''} \in \Omega$.[20] Note also that if there is only one commitment type ω_{s_1} then s_1 is not behaviorally confounded.

[20]Ruling out this possibility would simplify the proof of Lemma 3 and would also let us replace B_1 with $B_0(s_1)$ in Definition 3. In particular, 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 $\bigcup_{s_1' \neq s_1^* : \omega_{s_1'} \in \Omega} p(\alpha_0, s_1', \alpha_2)$.

Our main result assumes that s_1^* is not behaviorally confounded. In games where player 0 is Nature (like trust and communication games), this is fairly innocuous, as α_0 is exogenous, so the identification condition $p(\alpha_0, s_1^*, \alpha_2) \neq p(\alpha_0, s_1', \alpha_2)$ for all $\alpha_2 \in B_1$ holds for generic $s_1^* \neq s_1'$. In games with a "real" player 0, it is much more restrictive, because α_0 is endogenous, so the identification condition need not hold generically. For example, in the deterrence game in Section 2, the pure Stackelberg strategy (A, F) is not behaviorally confounded if and only if 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 Section 8 we show that our results extend even if s_1^* is behaviorally confounded, so long as it has sufficiently high prior weight relative to any behavioral confound that induces an η-confirmed best response that is not also a best response to s_1^*.

4.3. *Payoff Lower Bound*

We are now prepared to state our main result.

THEOREM 1: *For any strategy s_1^*, if $\omega_{s_1^*} \in \Omega$ and s_1^* is confound-defeating and not behaviorally confounded, then*

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

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 $p(\sigma_0^*, s_1^*, \sigma_2^* | h^t)$ 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 $\omega_{s_1^*}$.[21] In total, the short-run players come to believe that player 1 is either the commitment type $\omega_{s_1^*}$ or is rational and playing 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 $\mathbb{Q} \in \Delta(H^\infty)$ denote the probability measure over infinite histories H^∞ induced by this deviation; that is, the measure induced by strategy profile $(\sigma_0^*, s_1^*, \sigma_2^*)$. We proceed in four steps.

Step 1: Teaching the Marginal Over Signals. For any $\eta > 0$, define the set of period-t histories where the equilibrium signal distribution is within η of that under the deviation by

$$H_\eta^t = \{h^t \in \operatorname{supp}(\bar{\mathbb{P}}_t) : \| p(\sigma_0^*, s_1^*, \sigma_2^* | h^t) - p(\sigma_0^*, \bar{\sigma}_1^*, \sigma_2^* | h^t) \| \leq \eta\}.$$

We 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 A.2.

[21] This step is not trivial, for example, because we allow s_1^* to be indistinguishable from a mixture of commitment types.

LEMMA 2: *We have*

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

Step 2: Ruling Out "Bad" Commitment Types. For any $\zeta > 0$, denote the set of beliefs with at most ζ 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 δ. The proof, which relies on the martingale convergence theorem, Assumption 1(1) and 1(2), and the assumption that s_1^* is not behaviorally confounded, is deferred to Appendix A.3. In what follows, given a history h, $\mu_t(\cdot|h) \in \Delta(\Omega)$ denotes the posterior belief over Ω conditional on the period t truncation h^t of h. (We also write $\mu_t(\cdot|h^t)$ for the posterior belief at period-t history h^t.)

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 δ and the choice of equilibrium) such that, for any $h \in G(\zeta)$ and any $t \geq \hat{T}(\zeta)$, we have $\mu_t(\cdot|h) \in M_\zeta$.*

Step 3: Inducing Short-Run Best Responses. For any $\xi > 0$, we say that a short-run player strategy (α_0, α_2) is a *ξ-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 hemicontinuous and $\hat{B}_0(s_1^*) = B(s_1^*)$, we have

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

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 the rational type, they will take a ξ-close best response to s_1^*. Its proof (deferred to Appendix A.4) relies on Lemma 1.

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

Step 4: Completing the Proof. 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 η,

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

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

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

As $\delta \to 1$, 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 remains valid for all large enough $\delta < 1$. Finally, taking η small enough so $\zeta(\eta)(\bar{V}(s_1^*) - \underline{u}_1) < \varepsilon/3$ gives the desired $V(s_1^*) - \varepsilon$ bound. *Q.E.D.*

5. CYCLICAL MONOTONICITY AND PAYOFF UPPER BOUND

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.

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.[22]

DEFINITION 4: Fix finite sets X, Y, and a function $u : X \times Y \to \mathbb{R}$. 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 4: *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*

$$\mathrm{OT}(\rho, \phi) : \max_{\gamma' \in \Delta(X\times Y)} \int u(x, y)\, d\gamma' \quad \textit{such that } \pi_X(\gamma') = \rho \textit{ and } \pi_Y(\gamma') = \phi$$

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

We apply Proposition 4 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

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

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 α_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* α_2, it is strictly $u_1(\cdot, \alpha_2)$-cyclically monotone for *all* α_2. In this case, the strict u_1-cyclical monotonicity of a set

[22]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 4.

$S \subset Y_0 \times A_1$ is well-defined independent of α_2.[23] 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 we denote the long-run player's greatest lower commitment payoff from a strictly u_1-cyclically monotone strategy by

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

COROLLARY 2: *Suppose u_1 is strictly cyclically separable. For any strategy s_1^*, if $\omega_{s_1^*} \in \Omega$, and s_1^* is strictly u_1-cyclically monotone and not behaviorally confounded, then*

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

In particular,

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

Thus, when the long-run player's payoff is strictly cyclically separable, the confound-defeating property in Theorem 1 can be replaced by strict u_1-cyclical monotonicity.

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-cyclical monotonicity for any α_2 (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, 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 α_2, it is $u_1(\cdot, \alpha_2)$-cyclically monotone for all α_2.

[23]Note that this case always applies in games without a player 2, such as deterrence games.

PROPOSITION 5: *Suppose u_1 is cyclically separable. Then for all $\varepsilon > 0$, there exists $\kappa > 0$ such that, for any prior μ_0 satisfying $\mu_0(\omega_R) > 1 - \kappa$ and any $\delta < 1$,*

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

A key step in the proof of Proposition 5 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 $h^t \in \operatorname{supp}(\mathbb{P}_t)$, $\sigma_1^*(h^t)$ is u_1-cyclically monotone.*

PROOF: Note that $\sigma_1^*(h^t)$ must solve

$$\mathrm{OT}\big(\sigma_0^*(h^t), \phi\big(\sigma_0^*(h^t), \sigma_1^*(h^t)\big); \sigma_2^*(h^t)\big).$$

This holds because, otherwise, there exists a strategy s_1 that gives player 1 a strictly higher payoff at history h^t than $\sigma_1^*(h^t)$ 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 $\operatorname{supp}(\gamma(\sigma_0^*(h^t), \sigma_1^*(h^t)))$ is $u_1(\cdot, \sigma_2^*(h^t))$-cyclically monotone. Hence, by cyclical separability, $\sigma_1^*(h^t)$ is u_1-cyclically monotone. *Q.E.D.*

Given Lemma 5, at any history h^t where the long-run player is thought to be rational with high probability, the short-run players must expect the long-run player strategy $\bar{\sigma}_1^*(h^t)$ to be u_1-cyclically monotone with high probability. When the short-run players best-respond to such a strategy, the long-run player's payoff cannot greatly exceed $\bar{v}_1^{\mathrm{CM}}$. Finally, if the prior probability of the rational type is close to 1, then for each period t, the probability that $\mu_t(\omega_R|h^t, \omega_R)$ is high is also close to 1 (since $\mu_t(\omega_R|h^t, \omega_R)$ is a $\mathbb{P}$-submartingale), so the long-run player's overall expected payoff also cannot greatly exceed $\bar{v}_1^{\mathrm{CM}}$.

Combining Corollary 2 and Proposition 5 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.

REMARK 1: A similar argument shows that, in both our general model (without cyclical separability) and in Fudenberg and Levine (1992)'s general model, the upper Stackelberg payoff $\sup_{s_1} \sup_{(\alpha_0,\alpha_2)\in B(s_1)} u_1(\alpha_0, s_1, \alpha_2)$ is an upper bound for $\limsup_{\mu_0:\mu_0(\omega_r)\to 1} \bar{U}_1(\delta)$ for any $\delta < 1$. The logic is that the long-run player's payoff cannot exceed the upper Stackelberg payoff at any history h^t where she is thought to be rational with high probability, and such histories occur with high probability on path when $\mu_0(\omega_r)$ is close to 1 and the long-run player is indeed rational. This is a very different upper bound than that given by Fudenberg and Levine, who show that the upper Stackelberg payoff allowing 0-confirmed best responses, $\sup_{s_1} \sup_{(\alpha_0,\alpha_2)\in B_0(s_1)} u_1(\alpha_0, s_1, s_2)$, is an upper bound for $\limsup_{\delta\to 1} \bar{U}_1(\delta)$ for any prior μ_0.

6. ONE-DIMENSIONAL SUPERMODULAR GAMES

We now show how our results apply in games where the long-run player's payoff is supermodular in a one-dimensional signal and action. This simple setting includes a wide range of applications, including games of deterrence, trust, delegation, and signaling.

6.1. *Supermodularity and Monotonicity*

The relevant supermodularity notion is strict supermodularity in (y_0, a_1) for all α_2.

DEFINITION 5: The long-run player's payoff u_1 is *strictly supermodular* if there exist total orders $(\succsim_{Y_0}, \succsim_{A_1})$ 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 $s_1 : Y_0 \to \Delta(A_1)$ is monotone if any selection from the graph of its support is monotone. Formally, we have the following.

DEFINITION 6: A 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 $a_1 \succsim a_1'$.

The following is the key result of this section.

PROPOSITION 6: *Let u_1 be strictly supermodular. For any 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 (α_0, α_2), s_1^* is monotone if and only if $\gamma(\alpha_0, s_1^*)$ is the unique solution to* $\text{OT}(\rho(\alpha_0), \phi(\alpha_0, s_1^*); \alpha_2)$.

PROOF: By Lemma 2.8 in Santambrogio (2015), if s_1^* is monotone then $\gamma(\alpha_0, s_1^*)$ is the unique *comonotone 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 γ, y_0 and a_1 are comonotone random variables. Conversely, since $\rho(\alpha_0)$ has full support, if s_1^* is not monotone then $\gamma(\alpha_0, s_1^*)$ is not comonotone. Finally, by Theorem 2.9 and Exercise 10 in Santambrogio (2015), when u_1 is strictly supermodular, the comonotone transport plan is the unique solution to $\text{OT}(\rho(\alpha_0), \phi(\alpha_0, s_1^*); \alpha_2)$. *Q.E.D.*

By Lemma 6, if s_1^* is confound-defeating then $\gamma(\alpha_0, s_1^*)$ is the unique solution to $\text{OT}(\rho(\alpha_0), \phi(\alpha_0, s_1^*); \alpha_2)$ 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 $\text{OT}(\rho(\alpha_0), \phi(\alpha_0, s_1^*); \alpha_2)$ for any (α_0, α_2), and hence is confound-defeating. Moreover, s_1^* is monotone if and only if $\gamma(\alpha_0, s_1^*)$ is comonotone, as shown in the proof of Lemma 6, and $\gamma(\alpha_0, s_1^*)$ is comonotone 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. *Q.E.D.*

We now 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

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

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

Combining Theorem 1 and Propositions 5 and 6, we obtain the following corollary, which specializes Corollary 2 and Proposition 5 to one-dimensional supermodular games.

COROLLARY 3: *Suppose u_1 is strictly supermodular. Then*

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

Conversely, for all $\varepsilon > 0$, there exists $\kappa > 0$ such that for any prior μ_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: in a one-dimensional supermodular game, a patient long-run player can secure 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 our results for delegation and communication games in the subsequent sections.

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 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 $\underline{v}_1^{\text{mon}} = \bar{v}_1^{\text{mon}}$ (e.g., if short-run players have a unique best response to any monotone strategy s_1 and the best monotone commitment strategy s_1^* satisfies $\omega_{s_1^*} \in \Omega$ and is not behaviorally confounded), Corollary 3 gives a unique payoff prediction as $\mu_0(\omega_R)$ and δ both approach 1. This holds, for example, in the deterrence and trust games.

6.2. *Delegation Games*

We apply Corollary 3 to repeated delegation games, where in each period a short-run player 2 chooses whether or not to delegate a decision to the long-run player 1, who then takes an action $a \in A \subset \mathbb{R}$ after observing a state $\theta \in \Theta \subset \mathbb{R}$, which is drawn i.i.d. across periods from a full-support distribution $\rho_0 \in \Delta(\Theta)$.[24] We assume A and Θ are finite.

[24] Lipnowski and Ramos (2020) study repeated delegation with two long-run players and no commitment types. We thank Navin Kartik for suggesting we discuss repeated delegation games.

Formally, player 2 can choose either a safe action $a_2 = S(afe)$, in which case the decision is delegated with probability $\varepsilon > 0$ (perhaps reflecting either a tremble or the occasional necessity of delegation), or action $a_2 = D(elegate)$, in which case delegation occurs with probability 1. Assume that player 2's action is observed, and player 1's action is observed if and only if the decision is delegated. (The state is never observed.) The assumption that delegation always occurs with positive probability ensures that Assumption 1 holds. In particular, player 1's action is identified even if player 2 always takes the safe action.

Normalize the players' no-delegation payoffs to 0, and denote a player i's payoff when the decision is delegated and action a is taken in state θ by $u_i(a, \theta)$. Assume that $u_1(a, \theta)$ is strictly supermodular. Then, for any long-run player strategy $s_1(\theta)$, payoffs are given by

$$u_i(s_1, S) = \varepsilon\mathbb{E}_\theta[u_i(s_1(\theta), \theta)] \quad \text{and} \quad u_i(s_1, D) = \mathbb{E}_\theta[u_i(s_1(\theta), \theta)].$$

Note that $u_i(s_1, D) > u_i(s_1, S)$ if and only if $u_i(s_1, D) > 0$.

The following result is an immediate application of Corollary 3.

PROPOSITION 7: *In repeated delegation, for any monotone strategy* s_1^*, *if* $\omega_{s_1^*} \in \Omega$, s_1^* *is not behaviorally confounded, and* $u_2(s_1^*, D) > 0$, *then* $\liminf_{\delta\to 1} \underline{U}_1(\delta) \geq u_1(s_1^*, D)$.

Proposition 7 contrasts with the results of Ely and Valimaki (2003) and Ely, Fudenberg, and Levine (2008), who provide conditions under which a patient long-run player's payoff approaches her minmax payoff in all equilibria of a repeated delegation game with commitment types. A key difference is our assumption that delegation always occurs with positive probability, so the long-run player's action is always identified. With this assumption, our Assumption 1 would also hold in Ely and Valimaki's model, and the commitment payoff theorem would apply whenever the Stackelberg type is present with positive probability and is not behaviorally confounded.

7. COMMUNICATION GAMES

We now consider implications of our results for repeated communication games. Recall that the model covers communication games by viewing $\rho_0(y_0)$ as the prior distribution of a payoff-relevant state y_0; letting $\rho_1(y_1|a_1) = \mathbf{1}(\{y_1 = a_1\})$; 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 θ.

7.1. *Signaling*

In a repeated 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 θ before taking an action $a_1 \in A_1 \subset \mathbb{R}$. The short-run player observes the current action a_1 (but not θ) and the history of past actions (but not past states) and then takes an action $r \in R$ in response. Assume Θ, A_1, and R are finite, the long-run player's payoff is given by $v(a_1, r) - w(a_1, \theta)$ for some functions v and w, 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 θ. Note that submodularity of w is a standard assumption in signaling theory, for example, in Spence (1973), $v(a_1, r) = r$ and $w(a_1, \theta) = a_1/\theta$.

The following result is another implication of Corollary 3.

PROPOSITION 8: *In repeated signaling, for any monotone strategy* s_1^*, *if* $\omega_{s_1^*} \in \Omega$ *and* s_1^* *is not behaviorally confounded, then* $\liminf_{\delta \to 1} \underline{U}_1(\delta) \geq V(s_1^*)$.

PROOF: We have $u_1(\theta, a_1, a_2) = (1 - \lambda)v(a_1, a_2(a_1)) - \lambda w(a_1, \theta)$, which is strictly supermodular in (θ, a_1) because w is strictly submodular. The result thus follows from Corollary 3. *Q.E.D.*

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, even if the history of states is never observed.

7.2. *Perturbed Cheap Talk*

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. The sender's utility $u_1(\theta, a_1, r)$ is then independent of a_1, and hence cannot be strictly supermodular in (θ, r). 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 θ. We ask what commitment payoffs $V(s_1)$ can be secured by leveraging such a grain of commitment.

By the revelation principle, for any set of sender actions A_1 and any strategy $\hat{s}_1 : \Theta \to \Delta(A_1)$, there exists a direct *communication mechanism* $s_1 : \Theta \to \Delta(R)$ such that $V(s_1) = V(\hat{s}_1)$. We restrict attention to direct communication mechanisms in this section, and in particular assume 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 (which is an immediate consequence of Corollary 3) provides a general sufficient condition for the commitment payoff $V(s_1)$ from a communication mechanism $s_1 : \Theta \to \Delta(R)$ to be approximately securable.

PROPOSITION 9: *If a communication mechanism* $s_1 : \Theta \to \Delta(R)$ *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* $w : R \times \Theta \to [0, 1]$,

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

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*

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

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 θ. In particular, if $R = \Theta$ and the receiver's optimal action in state θ 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 θ 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 9 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.

For a concrete example, consider Kamenica and Gentzkow's prosecutor-judge example, where a prosecutor discloses information about a defendant's guilt to maximize the probability of conviction. The prosecutor has state-independent utility $v(r) = 1(\{r = \text{Convict}\})$. If the prosecutor's payoff is augmented with an ε cost of recommending that an innocent defendant be convicted, then $w(\tilde{r}, \theta)$ is strictly submodular in the order Convict $\succ$ Acquit, Guilty $\succ$ Innocent, and our results imply that a patient prosecutor can secure her commitment payoff.[25]

To fully operationalize Proposition 9, it remains to characterize what mechanisms $s_1 : \Theta \to \Delta(R)$ are monotone with respect to some order $(\succsim_\Theta, \succsim_R)$. To do so, let $G(s_1)$ be the bipartite graph with vertices Θ and R, where a state θ 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 $s_1 : \Theta \to \Delta(R)$ 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 $k \in \{1, 2, 3\}$.

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 is nonmonotone 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 nonmonotone 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 $s_1 : \Theta \to \Delta(R)$ is monotone with respect to some order.

PROPOSITION 10: *A communication mechanism $s_1 : \Theta \to \Delta(R)$ 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 10 implies that the set of mechanisms that are monotone with respect to some order includes, for example, all partitions (i.e., deterministic mechanisms $s_1 : \Theta \to R$) and all linear partitions with randomization at the endpoints. This includes the set of

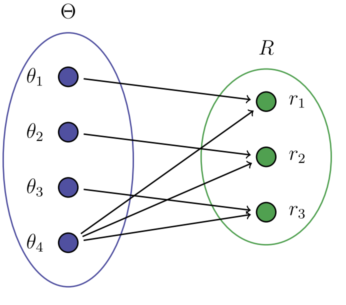

FIGURE 1.—A Type (1) Forbidden Triple. *Notes.* Monotonicity is violated for any placement of θ_4 in the order $\succsim_\Theta$.

[25]Another implication of our results is that 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 selection of the babbling equilibrium.

all monotone partitions, which are shown to be optimal in certain persuasion problems by Kolotilin (2018) and Dworczak 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 $(0, 1/3)$, $(1, 1/3)$, $(1, 2/3)$, $(0, 2/3)$.[26]

Proposition 10 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 $f : X \to Y$ is consistent with a joint distribution over $X \times Y$ that is comonotone 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.

8. A PAYOFF BOUND WITH BEHAVIORAL CONFOUNDING

Our last result generalizes Theorem 1 to the case where s_1^* is behaviorally confounded. This extension is particularly important for games where α_0 is endogenous, like the deterrence game in Section 2, as in these games s_1^* is often behaviorally confounded when there are multiple commitment types (as discussed in Section 4.2).

The details of the generalization are somewhat intricate, but the main idea is simple. 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 β with which the long-run weight on s_1^* exceeds this level under the deviation measure $\mathbb{Q}$ (where the player 1 always plays s_1^*). We call β the "salience" of s_1^*, and we establish a lower bound for the patient long-run player's payoff as a function of β. 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 μ over Ω' by $\mu(\cdot|\Omega')$. In addition, given a belief $\mu \in \Delta(\Omega \setminus \{\omega_R\})$, we slightly abuse notation by also denoting by $\mu(\cdot|\Omega \setminus \{\omega_R\})$ 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: Fix any strategy s_1^* and any $\eta, \varsigma > 0$, a number $c \in [0, 1]$ is a (s_1^*, η, ς)-*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

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

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

[26]This example is a discrete version of a "bi-pooling" policy. With a continuous state, Kleiner, Moldovanu, and Strack (2021) and Arieli, Babichenko, Smorodinsky, and Yamashita (2023) show that the bi-pooling policies are those that are uniquely optimal in some persuasion problem.

Let $c_{\eta,\varsigma}(s_1^*)$ denote the supremum over (s_1^*, η, ς)-confounding weights, with the convention that if no such weight exists, then $c_{\eta,\varsigma}(s_1^*) = -\infty$. Finally, let $c_0(s_1^*) = \lim_{\varsigma\to 0} \times \lim_{\eta\to 0} c_{\varsigma,\eta}(s_1^*)$.

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

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

with the convention that if $c_0(s_1^*) = -\infty$, then $\beta = 1$.

The logic of Definitions 8 and 9 is that conditional on the long-run player being irrational, $c_0(s_1^*)$ is the minimum weight on s_1^* that ensures short-run players best respond to s_1^*, and (by Bayes' rule) β is the minimum probability the long-run weight on s_1^* strictly exceeds $c_0(s_1^*)$.

The general version of our main result is as follows.

THEOREM 2: *For any strategy s_1^*, if $\omega_{s_1^*} \in \Omega$ and s_1^* is confound-defeating and has salience β, then*

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

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

Moreover, as $c_0(s_1^*) < 1$, $\beta \to 1$ whenever $\mu_0(\omega_{s_1^*}|\Omega_0(s_1^*) \setminus \{\omega_R\}) \to 1$, in which case Theorem 2 recovers the conclusion of Theorem 1 even if s_1^* is behaviorally confounded. But 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 $\mu_0(\omega_{s_1^*}|\Omega_0(s_1^*) \setminus \{\omega_R\})$.[27]

For example, consider the deterrence game from Section 2 with two commitment types: the pure Stackelberg type (A, F) and the type s_1 that takes A with probability p for

[27]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."

each signal. In this example, $c_0(A, F)$ 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 $c_0(A, F)$ and plays s_1 with probability $1 - c_0$, which is given by $c_0(A, F) = \frac{pg+(1-p)l}{(2p-1)(1+g)}$. The salience of type (A, F) is then

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

and Theorem 2 implies that as $\delta \to 1$ 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 $c_0(A, F) < 1$, so that as the prior weight on (A, F) relative to s_1 increases, β converges to 1 and the long-run player is assured her pure Stackelberg payoff p.

9. 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 is confound-defeating if and only if it is monotone. In delegation games with supermodular agent utility or in signaling games with state-independent sender preferences and strictly submodular signaling costs, a patient agent or sender can secure her commitment payoff from any monotone strategy. Finally, we characterized communication strategies that are monotone with respect to some order and 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 9 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.

APPENDIX A: APPENDIX: OMITTED PROOFS

A.1. *Proof of Proposition 3*

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 $\mathrm{OT}(\rho(\alpha_0), \phi(\alpha_0, s_1^*); \alpha_2)$. 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^η 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^η 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 $p(\alpha_0^\eta, s_1^\eta, \alpha_2^\eta) = p(\alpha_0^\eta, \tilde{s}, \alpha_2^\eta)$. Since a_1 is identified, whenever $\phi(\alpha_0^\eta, s_1^\eta) = \phi(\alpha_0^\eta, \tilde{s}_1)$ we know $p(\alpha_0^\eta, s^\eta, \alpha_2^\eta) = p(\alpha_0^\eta, \tilde{s}, \alpha_2^\eta)$, and hence s_1^η solves $\mathrm{OT}(\rho(\alpha_0^\eta), \phi(\alpha_0^\eta, s_1^\eta); \alpha_2^\eta)$. 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 $\mathrm{OT}(\rho(\alpha_0), \phi(\alpha_0, s_1); \alpha_2)$ is jointly upper hemicontinuous 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 $\mathrm{OT}(\rho(\alpha_0), \phi(\alpha_0, s_1^*); \alpha_2)$. Thus, the second definition fails.

A.2. *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,

$$\sum_t \mathbb{E}^{\mathbb{Q}}\left[d\left(p\left(\sigma_0^*, s_1^*, \sigma_2^* \middle| h^t\right) \middle\| p\left(\sigma_0^*, \bar{\sigma}_1^*, \sigma_2^* \middle| h^t\right)\right)\right] \leq -\log \mu_0(\omega_{s_1^*}).$$

Hence, by Markov's inequality,

$$\mathbb{E}^{\mathbb{Q}}\left[\#\left\{t : d\left(p\left(\sigma_0^*, s_1^*, \sigma_2^* \middle| h^t\right) \middle\| p\left(\sigma_0^*, \bar{\sigma}_1^*, \sigma_2^* \middle| h^t\right)\right) > \frac{\eta^2}{2}\right\}\right] < -\frac{2\log \mu_0(\omega_{s_1^*})}{\eta^2}.$$

On the other hand, by Pinsker's inequality,

$$d\left(p\left(\sigma_0^*, s_1^*, \sigma_2^* \middle| h^t\right) \middle\| p\left(\sigma_0^*, \bar{\sigma}_1^*, \sigma_2^* \middle| h^t\right)\right) \leq \frac{\eta^2}{2} \quad \Longrightarrow \quad h^t \in H_\eta^t.$$

This gives the desired bound.

A.3. *Proof of Lemma 3*

We first show that the desired conclusion holds for each strategy profile $(\sigma_0^*, \sigma_1^*, \sigma_2^*)$ with $(\sigma_0^*, \sigma_2^*) \in B_1^H$ (and hence for any equilibrium), and then show that $\hat{T}$ can be fixed independent of the choice of δ and the equilibrium.

LEMMA 7: *For any strategy profile $(\sigma_0^*, \sigma_1^*, \sigma_2^*)$ where $(\sigma_0^*, \sigma_2^*) \in B_1^H$, 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 $\overline{\mathbb{P}}$ and $\mu_t(\cdot|h)$ is a martingale relative to $\overline{\mathbb{P}}$, $\mu_t(\cdot|h)$ 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 $Y_1^* = \text{supp}(\rho_1(\cdot|s_1^*(Y_0), a_2))$ (which is independent of the choice of a_2 by Assumption 1(2)), the set of signals that arise with positive probability when player 1 plays s_1^*. This supposition is without loss as $\rho_0(\cdot|a_0)$ has full support (by Assumption 1(1)), 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 distribution of y_1 conditional on reaching history h^t and the event $\omega \in \Omega'$; when Ω' 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

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

where the first equality is by Bayes' rule, the second pulls out a common term, and the inequality is by the definition of c. Since $\mu_t(\cdot|h)$ converges, y_1 realizes infinitely often, and $c > 0$, this implies

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

At the same time, applying the argument in Lemma 2 conditional on the event $\omega \neq \omega_R$ and the fact s_1^* is not behaviorally confounded implies that, for $\mathbb{Q}$-almost all histories $h \in H^\infty$,

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

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 $y_1 \notin Y_1^*$,

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

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

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

Thus, by the triangle inequality,

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

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 $f_n : H^\infty \to \mathbb{R}$ converges $\mathbb{Q}$-almost surely to f, then for all $\zeta > 0$, there exists $G(\zeta) \subset H^\infty$ satisfying $\mathbb{Q}(G(\zeta)) \geq 1 - \zeta$ such that $f_n \to f$ uniformly on $G(\zeta)$. Thus, Lemma 7 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. *Q.E.D.*

We now show that $\hat{T}$ can be chosen as a function only of ζ and not of δ or the equilibrium strategies. To this end, let $\mathbb{Q}^{\sigma_0,\sigma_2}$ be the probability measure on H^∞ 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 $\mathbb{Q}^{\sigma_0,\sigma_2}$-almost surely 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 (σ_0, σ_2) are given. Next, let $L^{\sigma_0,\sigma_2} \subset H^\infty$ be the set of all histories h where $\mu_t^{\sigma_0,\sigma_2}(\omega_{s_1^*}|\Omega \setminus \{\omega_R\}, h^t) \to 1$ as $t \to \infty$. By Lemma 7, $\mathbb{Q}^{\sigma_0,\sigma_2}(L^{\sigma_0,\sigma_2}) = 1$ for all $(\sigma_0, \sigma_2) \in B_1^H$.

Following Mertens, Sorin, and Zamir (2015, p. 173), we identify B_1^H with the space $(B_1^\infty)^{H^\infty}$ of maps from infinite histories h to sequences of stage game strategies $(\alpha_0, \alpha_2)^\infty$. We show that $(B_1^\infty)^{H^\infty}$ is compact in the topology of uniform convergence. Endow $H^\infty = Y_1^\infty$ with the product topology, where each coordinate is endowed with the discrete topology; by Tychonoff's theorem, this makes H^∞ compact. Endow B_1^∞ with the product topology, which is generated by the metric

$$\tilde{d}\big((\alpha_0, \alpha_2)^\infty, (\alpha_0', \alpha_2')^\infty\big) = \sum_{t=0}^{\infty} \frac{1}{2^t} \big\|(\alpha_0, \alpha_2)_t - (\alpha_0', \alpha_2')_t\big\|.$$

This is again compact by Tychonoff's theorem because $\Delta(A_0) \times \Delta(A_2)$ is compact. Note that each $(\sigma_0, \sigma_2) : H^\infty \to B_1^\infty$ is continuous with these topologies, because if two sequences of histories agree for the first T periods then the $\tilde{d}$-distance between their images under (σ_0, σ_2) is at most $1/2^{T-1}$. Moreover, because the modulus of continuity is independent of both the choice of histories and the choice of (σ_0, σ_2), the family of functions $(B_1^\infty)^{H^\infty}$ is equicontinuous. Next, endow $(B_1^\infty)^{H^\infty}$ with the topology of uniform convergence, which is generated by the sup norm

$$d\big((\sigma_0, \sigma_2), (\sigma_0', \sigma_2')\big) = \sup_{h \in H} \tilde{d}\big((\sigma_0(h), \sigma_2(h)), (\sigma_0'(h), \sigma_2'(h)\big).$$

Thus, $(B_1^\infty)^{H^\infty}$ is equicontinuous and closed.

Moreover, as B_1^∞ is compact, for any $h \in H^\infty$ the closure of the set $\{(\sigma_0(h), \sigma_2(h))\}_{(\sigma_0,\sigma_2)\in(B_1^\infty)^{H^\infty}}$ is a compact subset of B_1^∞. We thus satisfy the conditions of Theorem 47.1 in Munkres (2000) (generalized Arzela–Ascoli), implying that $(B_1^\infty)^{H^\infty}$ is compact in the topology of uniform convergence.[28]

We can now prove that $\hat{T}(\zeta, \sigma_0^*, \sigma_2^*)$ can be chosen independent of the choice of (σ_0^*, σ_2^*) (and hence also independent of δ). Suppose by contradiction that the statement is false. Then there exists $\zeta > 0$ such that for each time $T \in \mathbb{N}$, there exist strategies $(\sigma_0^T, \sigma_2^T) \in B_1^H$ 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 $(B_1^\infty)^{H^\infty}$, we have $(\sigma_0^T, \sigma_2^T) \to (\sigma_0^\infty, \sigma_2^\infty) \in B_1^H$ in the topology of uniform convergence. Since μ_T depends only on the history up to time T, this implies that $\mathbb{Q}_{T'}^{\sigma_0^T,\sigma_2^T}(E_T(\zeta)) > \zeta$ for all $T' \geq T$, where $\mathbb{Q}_{T'}$ is the measure $\mathbb{Q}$ conditioned on the time-T' σ-algebra. Thus,

[28]Theorem 47.1 of Munkres proves compactness of $(B_1^\infty)^{H^\infty}$ in the topology of compact convergence, which Theorem 46.7 of Munkres shows coincides with the topology of uniform convergence when H^∞ is compact.

since $\mathbb{Q}_{T'}$ is continuous in strategies as a finite-dimensional measure, passing (σ_0^T, σ_2^T) to the limit (while fixing the time T' and the set of histories $E_T(\zeta)$) gives $\mathbb{Q}_{T'}^{\sigma_0^\infty, \sigma_2^\infty}(E_T(\zeta)) > \zeta$ for all T' sufficiently large; and then taking $T' \to \infty$ 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[29]

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

Thus, for any history $h \in \limsup_{T\to\infty} E_T(\zeta) = E_\infty(\zeta)$, 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 7 for the strategies $(\sigma_0^\infty, \sigma_2^\infty) \in B_1^H$, completing the proof.

A.4. *Proof of Lemma 4*

Fix any $h^t \in H_\eta^t$ and $\mu_t(\cdot|h^t) \in M_0$. Then

$$\begin{aligned}&\big\|p\big(\sigma_0^*, \bar{\sigma}_1^*, \sigma_2^*\big|h^t\big) - p\big(\sigma_0^*, s_1^*, \sigma_2^*\big|h^t\big)\big\| \\ &\quad = \big|\mu_t\big(\omega_R|h^t\big)\big|\big\|p\big(\sigma_0^*, \sigma_1^*, \sigma_2^*\big|h^t\big) - p\big(\sigma_0^*, s_1^*, \sigma_2^*\big|h^t\big)\big\| < \eta.\end{aligned}$$

Since $p(\sigma_0^*, s_1^*, \sigma_2^*|h^t)$ is continuous in $\mu_t(\cdot|h^t)$, there exists a strictly positive function $\zeta(\eta)$ satisfying $\lim_{\eta\to 0}\zeta(\eta) = 0$ such that if $h^t \in H_\eta^t$ and $\mu_t(\cdot|h^t) \in M_{\zeta(\eta)}$ then

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

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

Now fix $c : (0, 1] \to (0, 1]$ satisfying $c(\eta) \to 0$, $\frac{\eta}{c(\eta)} \to 0$ as $\eta \to 0$. If $\mu_t(\omega_R|h^t) \geq c(\eta)$ then

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

and in addition $h^t \in \operatorname{supp}(\mathbb{P})$. Hence, as $\eta \to 0$, Lemma 1 implies that $||\sigma_1^*(h^t) - s_1^*|| \to 0$, and so $(\sigma_0^*(h^t), \sigma_2^*(h^t)) \in \hat{B}_{\xi_1(\eta)}(s_1^*)$ for some strictly positive function $\xi_1(\eta)$ satisfying $\xi_1(\eta) \to 0$. If instead $\mu_t(\omega_R|h^t) < c(\eta)$, then $||\bar{\sigma}_1^*(h^t) - s_1^*|| \leq 1 - \zeta(\eta) - c(\eta)$, and hence $(\sigma_0^*(h^t), \sigma_2^*(h^t)) \in \hat{B}_{\zeta(\eta)+c(\eta)}(s_1^*)$. Taking $\xi(\eta) = \max\{\xi_1(\eta), \zeta(\eta) + c(\eta)\}$ completes the proof.

A.5. *Proof of Proposition 4*

The forward direction. Let γ be feasible in $\text{OT}(\rho, \phi)$ with $\operatorname{supp}(\gamma)$ not strictly u-cyclically monotone. Let $\{(x_i, y_i)_{i=1}^N\} \subset \operatorname{supp}(\gamma)$ be a collection of pairs witnessing a vi-

[29]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.

olation 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 $\text{OT}(\rho, \phi)$ (since γ and γ' transport the same mass into and out of each point) 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 γ does not uniquely solve $\text{OT}(\rho, \phi)$.

Conversely, if γ is feasible in $\text{OT}(\rho, \phi)$ and $\text{supp}(\gamma)$ is strictly u-cyclically monotone, consider any feasible $\gamma' \neq \gamma$ in $\text{OT}(\rho, \phi)$. Since γ and γ' are both feasible 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 γ and γ' 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 γ and γ' 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 γ'' 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 γ' 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, γ must uniquely solve $\text{OT}(\rho, \phi)$.

A.6. *Proof of Proposition 5*

The result is obvious if $\bar{v}_1^{\text{CM}} = \bar{u}_1$, so suppose $\bar{v}_1^{\text{CM}} < \bar{u}_1$, and fix $\varepsilon < 2(\bar{u}_1 - \bar{v}_1^{\text{CM}})$, a prior μ_0 with $\mu_0(\omega_R) > 0$, and an equilibrium $(\sigma_0^*, \sigma_1^*, \sigma_2^*)$. We start with an additional lemma.

LEMMA 8: *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 $\xi > 0$ 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 $u_1(\alpha_0^n, s_1^n, \alpha_2^n) \geq \bar{v}_1^{\text{CM}} + \xi$. 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), and hence compact, 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

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

The first inequality is by definition of $\tilde{s}_1^n$, the second inequality follows because $B(\cdot)$ is upper hemicontinuous and u_1 is continuous, and the third inequality follows because s_1 is u_1-cyclically monotone. *Q.E.D.*

Now, note that at any history h^t where $\mu_t(\omega_R|h^t) > 1-\varsigma$, we have $\|\bar{\sigma}_1^*(h^t) - \sigma_1^*(h^t)\| < \varsigma$. Thus, by Lemmas 5 and 8, there exists $\varsigma > 0$ such that at any history h^t where $\mu_t(\omega_R|h^t) > 1-\varsigma$, we have $u_1(\sigma_0^*, \sigma_1^*, \sigma_2^*|h^t) < \bar{v}_1^{\mathrm{CM}} + \varepsilon/2$. Since $\mu_t(\omega_R|h^t, \omega_R)$ is a $\mathbb{P}$-submartingale,

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

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

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

This completes this proof, as this payoff is less than $\bar{v}_1^{\mathrm{CM}} + \varepsilon$ whenever

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

A.7. *Proof of Proposition 10*

Suppose that $s_1 : \Theta \to \Delta(R)$ 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 $r_k \in \operatorname{supp}(s_1(\theta_k)) \cap \operatorname{supp}(s_1(\theta_{k+1}))$ for $k \in \{1, \dots, K-1\}$ and $r_K \in \operatorname{supp}(s_1(\theta_K))$, monotonicity requires $r_1 \prec \dots \prec r_K$. But this gives a contradiction, since $r_K \in \operatorname{supp}(s_1(\theta_1))$ and $r_1 \in \operatorname{supp}(s_1(\theta_2))$.

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 $\operatorname{supp}(s_1(\theta)) = \{r_k\}$. Note such a state θ is linked to at most one $r_k \in \{r_1, \dots, r_{K-1}\}$, as if θ is linked to distinct r_k, r_k' then appending θ to both ends of the path from r_k to r_k' forms a cycle. In addition, θ cannot be linked to r_K, as then it could be appended to the maximum path. Finally, θ 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 (θ_1, r_1) with (θ, r), (θ, 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 $\{r_{k-1}, r_k\} \in \operatorname{supp}(s_1(\theta_k))$, $\{r, r_k\} \in \operatorname{supp}(s_1(\theta))$, and $\{r_k, r_{k+1}\} \in \operatorname{supp}(s_1(\theta_{k+1}))$ (see Figure 2).

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

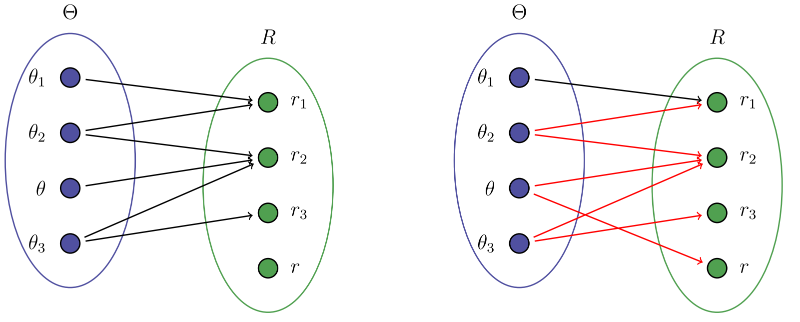

FIGURE 2.—Each State $\theta \notin \{\theta_1, \ldots, \theta_K\}$ Has Only One Neighbor. *Notes.* If $\theta \notin \{\theta_1, \ldots, \theta_K\}$ is linked to $r_k \in \{r_1, \ldots, r_K\}$ and $r \notin \{r_1, \ldots, r_K\}$, then $\{\theta_k, \theta, \theta_{k+1}\}$ together with $\{r, r_{k-1}, r_k, r_{k+1}\}$ is a forbidden triple.

Similarly, for any $r \in R \setminus \{r_1, \ldots, r_K\}$, there exists $k \in \{2, \ldots, K\}$ such that r_k is linked only to θ_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 $r_{k-1} \precsim_R r \precsim_R r_k$. This follows because if $r \notin \{r_1, \ldots, r_K\}$ then $r_{k-1} \precsim_R r \precsim_R r_k$ 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 θ_k.

Finally, we claim that s_1 is monotone with respect to $(\succsim_\Theta, \succsim_R)$. To see this, fix any $\theta \succ_\Theta \theta'$. Let $\tilde{k} = \max\{k : \theta \succsim_\Theta \theta_k\}$. If $\theta \succ_\Theta \theta_{\tilde{k}} \succsim_\Theta \theta'$, then $\text{supp}(s_1(\theta)) = \{r_{\tilde{k}}\}$ and $r_{\tilde{k}} \succsim_R r$ for all $r \in \text{supp}(s_1(\theta'))$. If $\theta = \theta_{\tilde{k}} \succ_\Theta \theta'$, then $r_{\tilde{k}-1}$ is the lowest action in $\text{supp}(s_1(\theta))$, and no action in $\text{supp}(s_1(\theta'))$ is above $r_{\tilde{k}-1}$. Lastly, if $\theta \succ_\Theta \theta' \succsim_\Theta \theta_{\tilde{k}}$, then $\text{supp}(s_1(\theta)) = \text{supp}(s_1(\theta')) = \{r_{\tilde{k}}\}$. Thus, in all cases, monotonicity holds.

A.8. *Proof of Theorem 2*

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

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

The proof of Lemma 9 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 ϑ can be chosen uniformly over a certain set of beliefs μ.

LEMMA 10: *For any $s_1 \in \Delta(A_1)^{Y_0}$ and $\varsigma > 0$, let*

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

Then there exists $\vartheta(\varsigma, s_1) > 0$, vanishing as $\varsigma \to 0$, 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 ς, 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 $(s_1^{n_k}, \tilde{s}_1^{n_k}) \to (s_1', \tilde{s}_1)$ 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. *Q.E.D.*

PROOF OF LEMMA 9: We show there exists strictly positive functions $\zeta(\eta)$ and $\xi(\eta)$, vanishing as $\eta \to 0$, and $\bar{T}(\eta)$ such that, for all $t > \bar{T}(\eta)$,

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

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

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 η, 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 $1 - \beta_{\varsigma,\eta}$. 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 η. 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 $\eta \to 0$. Thus, as η 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 10. Since $\mu_t(\cdot|h^t, \Omega \setminus \{\omega_R\}) \in C_\varsigma(s_1^*)$, Lemma 10 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 (η, ς)-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

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

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

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

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

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

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

$$q_{\eta,s}\bigg(\frac{c_{\eta,s}}{1-\eta}\bigg) + (1-q_{\eta,s})(1) \ge \mu_0\big(\omega_{s_1^*}|\Omega_\eta\big(s_1^*\big) \setminus \{\omega_R\}\big) \quad \Longleftrightarrow$$

$$q_{\eta,s} \le \min\left\{\frac{1-\mu_0\big(\omega_{s_1^*}|\Omega_\eta\big(s_1^*\big) \setminus \{\omega_R\}\big)}{1-\dfrac{c_{\eta,s}}{1-\eta}}, 1\right\} = 1-\beta_{\eta,s},$$

completing the proof. *Q.E.D.*

REFERENCES

Acemoglu, Daron and Alexander Wolitzky (2024), "Mistrust, Misperception, and Misunderstanding: Imperfect Information and Conflict Dynamics." In *Handbook of the Economics of Conflict*. [2011]

Arieli, Itai, Yakov Babichenko, Rann Smorodinsky, and Takuro Yamashita (2023), "Optimal Persuasion via bi-Pooling." *Theoretical Economics*, 18 (1), 15–36. [2029]

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). [2035]

Avenhaus, Rudolf, Bernhard von Stengel, and Shmuel Zamir (2002), "Inspection Games." In *Handbook of Game Theory With Economic Applications*, Vol. 3, 1947–1987, Elsevier. Chapter 51. [2011]

Ball, Ian and Deniz Kattwinkel (2024), "Quota Mechanisms: Finite-Sample Optimality and Robustness." Working Paper. [2010,2021]

Battigalli, Pierpaolo and Joel Watson (1997), "On "Reputation" Refinements With Heterogeneous Beliefs." *Econometrica*, 369–374. [2031]

Best, James and Daniel Quigley (2024), "Persuasion for the Long Run." *Journal of Political Economy*, 132 (5), 1305–1337. [2011]

Chakraborty, Archishman and Rick Harbaugh (2007), "Comparative Cheap Talk." *Journal of Economic Theory*, 132 (1), 70–94. [2010]

Chen, Ying, Navin Kartik, and Joel Sobel (2008), "Selecting Cheap-Talk Equilibria." *Econometrica*, 76 (1), 117–136. [2027]

Clark, Daniel, Drew Fudenberg, and Alexander Wolitzky (2021), "Record-Keeping and Cooperation in Large Societies." *The Review of Economic Studies*, 88 (5), 2179–2209. [2010]

Dworczak, Piotr and Giorgio Martini (2019), "The Simple Economics of Optimal Persuasion." *Journal of Political Economy*, 127 (5), 1993–2048. [2029]

Ely, Jeffrey and Juuso Valimaki (2003), "Bad Reputation." *Quarterly Journal of Economics*, 118 (3), 785–814. [2010,2026,2030]

Ely, Jeffrey, Drew Fudenberg, and David K. Levine (2008), "When Is Reputation Bad?" *Games and Economic Behavior*, 63 (2), 498–526. [2010,2026,2030]

Escobar, Juan F. and Juuso Toikka (2013), "Efficiency in Games With Markovian Private Information." *Econometrica*, 81 (5), 1737–1767. [2010]

Frankel, Alex (2014), "Aligned Delegation." *American Economic Review*, 104 (1), 66–83. [2010]

Fudenberg, Drew and David Levine (1989), "Reputation and Equilibrium Selection in Games With a Patient Player." *Econometrica*, 57 (4), 759–778. [2008,2031]

Fudenberg, Drew and David Levine (1992), "Maintaining a Reputation When Strategies Are Imperfectly Observed." *Review of Economic Studies*, 59 (3), 561–579. [2008-2010,2012,2015-2017,2019,2023,2031]

Fudenberg, Drew, Ying Gao, and Harry Pei (2022), "A Reputation for Honesty." *Journal of Economic Theory*, 204, 105508. [2011]

Gossner, Olivier (2011), "Simple Bounds on the Value of a Reputation." *Econometrica*, 79 (5), 1627–1641. [2008,2010,2012,2019,2032]

Heller, Yuval and Erik Mohlin (2018), "Observations on Cooperation." *The Review of Economic Studies*, 85 (4), 2253–2282. [2010]

Jackson, Matthew O. and Hugo F. Sonnenschein (2007), "Overcoming Incentive Constraints by Linking Decisions." *Econometrica*, 75 (1), 241–257. [2010]

Kamenica, Emir and Matthew Gentzkow (2011), "Bayesian Persuasion." *American Economic Review*, 101 (6), 2590–2615. [2010,2028]

Kartik, Navin (2009), "Strategic Communication With Lying Costs." *The Review of Economic Studies*, 76 (4), 1359–1395. [2027]

Kleiner, Andreas, Benny Moldovanu, and Philipp Strack (2021), "Extreme Points and Majorization: Economic Applications." *Econometrica*, 89 (5), 1671–1700. [2029]

Kolotilin, Anton (2018), "Optimal Information Disclosure: A Linear Programming Approach." *Theoretical Economics*, 13 (2), 607–635. [2029]

Kuvalekar, Aditya, Elliot Lipnowski, and Joao Ramos (2022), "Goodwill in Communication." *Journal of Economic Theory*, 203, 105467. [2011]

Lin, Xiao and Ce Liu (2024), "Credible Persuasion." *Journal of Political Economy*, 132 (7). [2010,2024]

Lipnowski, Elliot and Joao Ramos (2020), "Repeated Delegation." *Journal of Economic Theory*, 188, 105040. [2025]

Liu, Qingmin (2011), "Information Acquisition and Reputation Dynamics." *The Review of Economic Studies*, 78 (4), 1400–1425. [2010]

Liu, Qingmin and Andrzej Skrzypacz (2014), "Limited Records and Reputation Bubbles." *Journal of Economic Theory*, 151, 2–29. [2010]

Mailath, George J. and Larry Samuelson (2006), *Repeated Games and Reputations: Long-Run Relationships*. Oxford University Press. [2032]

Margaria, Chiara and Alex Smolin (2018), "Dynamic Communication With Biased Senders." *Games and Economic Behavior*, 110, 330–339. [2010]

Mathevet, Laurent, David Pearce, and Ennio Stacchetti (2024), "Reputation and Information Design." Working Paper. [2010]

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. [2010]

Meng, Delong (2021), "On the Value of Repetition for Communication Games." *Games and Economic Behavior*, 127, 227–246. [2010]

Mertens, Jean-François, Sylvain Sorin, and Shmuel Zamir (2015), *Repeated Games*, Vol. 55. Cambridge University Press. [2034]

Morris, Stephen (2001), "Political Correctness." *Journal of Political Economy*, 109 (2), 231–265. [2028]

Munkres, James R. (2000), *Topology*, second edition. Prentice Hall. [2034]

Myerson, Roger B. (1982), "Optimal Coordination Mechanisms in Generalized Principal–Agent Problems." *Journal of mathematical economics*, 10 (1), 67–81. [2031]

Pei, Harry (2020), "Reputation Effects Under Interdependent Values." *Econometrica*, 88 (5), 1671–1700. [2010, 2013]

Pei, Harry (2023), "Repeated Communication With Private Lying Costs." *Journal of Economic Theory*, 210, 105668. [2011]

Pei, Harry (2024), "Reputation Effects Under Short Memories." *Journal of Political Economy*, 132 (10). [2010]

Rahman, David M. (2024), "Detecting Profitable Deviations." *Journal of Mathematical Economics*, 111, 102946. [2010]

Rayo, Luis and Ilya Segal (2010), "Optimal Information Disclosure." *Journal of Political Economy*, 118 (5), 949–987. [2010]

Renault, Jérôme, Eilon Solan, and Nicolas Vieille (2013), "Dynamic Sender–Receiver Games." *Journal of Economic Theory*, 148 (2), 502–534. [2010]

Rochet, Jean-Charles (1987), "A Necessary and Sufficient Condition for Rationalizability in a Quasi-Linear Context." *Journal of Mathematical Economics*, 16 (2), 191–200. [2009,2010,2021,2024]

Santambrogio, Filippo (2015), *Optimal Transport for Applied Mathematicians*, Vol. 55 (58–63), 94, Birkäuser, NY. [2023,2024]

Schelling, Thomas C. (1966), *Arms and Influence*. The Henry L. Stimson Lectures Series. Yale University Press. Chapter 1, 74. [2007]

Sorin, Sylvain (1999), "Merging, Reputation, and Repeated Games With Incomplete Information." *Games and Economic Behavior*, 29, 274–308. [2019]

Spence, Michael (1973), "Job Market Signaling." *The Quarterly Journal of Economics*, 87 (3), 355–374. [2026]

Takahashi, Satoru (2010), "Community Enforcement When Players Observe Partners' Past Play." *Journal of Economic Theory*, 145 (1), 42–62. [2010]

Watson, Joel (1993), "A "Reputation" Refinement Without Equilibrium." *Econometrica*, 199–205. [2031]

Co-editor Marina Halac handled this manuscript.

Manuscript received 17 December, 2024; final version accepted 20 July, 2025; available online 29 July, 2025.