Justified Communication Equilibrium

Daniel Clark and Drew Fudenberg*

August 8, 2020

Abstract

Justified communication equilibrium (JCE) is an equilibrium refinement for signaling games with cheap-talk communication. A strategy profile must be a JCE to be a stable outcome of non-equilibrium learning when the receivers are *initially trusting* and senders play many more times than receivers. In this model, the counterfactual "speeches" that have been informally used to motivate past refinements are messages that are actually sent. Stable profiles need not be perfect Bayesian equilibria, so JCE sometimes preserves equilibria that existing refinements eliminate. Despite this, it has much of the flavor of NWBR, and coincides with it in *co-monotonic* signaling games.

*Department of Economics, MIT. We thank Kevin He, David K. Levine, and Alex Wolitzky for helpful comments, and NSF grant SES 1643517 for financial support.

1 Introduction

Although it is not usually included in formal analyses of signaling games, cheap-talk communication is available in many of the settings signaling games are intended to model. Thus it is desirable for game-theoretic analysis of signaling games to explicitly include cheap-talk communication in the extensive form, because incorporating cheap talk changes the action space of the sender, which can lead to different results. The prevalence of cheap-talk communication also makes it important to extend the learning-theoretic foundations for equilibrium concepts to settings where players can costlessly communicate.

Signaling games with or without cheap talk can have a great many equilibria. This paper provides a learning foundation for *justified communication equilibrium* (JCE), which is a new equilibrium refinement for signaling games with cheap-talk communication. For a given signal and strategy profile, a sender type is *justified* if some conceivable response makes the type indifferent between playing the signal or conforming to the strategy profile, and makes all other types weakly prefer to conform. A *justified response* is a convex combination of best responses to beliefs that assign probability 1 to the justified types for that signal. JCE requires that for every signal, there is at least one message that induces the receiver to play a justified response.

We study the limits of steady states in an overlapping generations learning environment where agents are patient have long expected lifetimes, and the senders on average live much longer and play many more repetitions of the game than the receivers do. This fits settings where the senders are institutions and the receivers are individuals (or families, clans, etc.), since institutions will typically be involved in many more interactions than individuals. We say that the strategy profiles corresponding to these steady states are *stable*.

We analyze the stable profiles under the assumptions that the message space is sufficiently rich and that receivers are *initially trusting*, which means roughly that the receivers' prior leads them to trust messages they have not previously observed to be lies. We view initial trust as a mild and plausible assumption on how receivers respond to messages. Section 3.1 discusses how it relates to past work on the interpretation of communication.

JCE emerges as a necessary condition for stability in our learning model because when senders are long-lived most of them play a best response to the aggregate play of the receivers. Consequently, receivers are very unlikely to encounter a signal being played by a non-justified type. Initial trust then means that most receivers will trust a message claiming to be a justified type, and so play a justified response.

Because we formally include cheap talk in the extensive form of the games we model, we can give the first learning foundation for "speeches" of the sort Cho and Kreps [1987] used to motivate the Intuitive Criterion, while avoiding some of the conceptual difficulties of an informal approach, since these messages are not counterfactual but are actually sent. Thus we are able to avoid the "Stiglitz critique," which argues that the arguments Cho and Kreps [1987] used to motivate how out-of-equilibrium speeches "should" be interpreted are internally contradictory. Instead, our model is in the spirit of what Fudenberg and Kreps [1988] called a "complete theory," as all of the relevant messages -are- sent by young senders who want to learn how the receivers respond.

Our results can be seen as both a validation of and a correction to previous signaling game refinements, which are roughly but not exactly in line with the implications of non-equilibrium learning. Specifically, the restrictions imposed by JCE on off-path play have some of the flavor of traditional signaling game refinements, such as the Intuitive Criterion (Cho and Kreps [1987]), D1 (Banks and Sobel [1987]), and NWBR (Cho and Kreps [1987])[1] and the intuitions behind these concepts have a connection to non-equilibrium learning. However, this connection is inexact, as none of the traditional equilibrium refinements is a necessary condition for stability in our learning model.[2] Indeed, as we show in Example 6, the stable outcomes of our learning model

[1] When we refer to NWBR in this paper we mean "Never a weak best response" in the sense of Cho and Kreps [1987] and Cho and Sobel [1990].

[2] Moreover, as far as we know they have not been shown to be necessary in -any- learning model.

need not be perfect Bayesian equilibria (Fudenberg and Tirole [1991a]), since the response to an off-path signal can be a mixture over pure best responses corresponding to different beliefs that need not itself be a best response. For this reason, JCE is not a refinement of perfect Bayesian equilibrium, but instead is a refinement of perfect Bayesian equilibrium with heterogeneous off-path beliefs (PBE-H, Fudenberg and He [2018]).

We explore the relationships of JCE with previous equilibrium refinements later in the paper, but we give two general results here as a preview: every JCE is a PBE-H that passes the "Intuitive Criterion Test," and every PBE-H that satisfies NWBR is path-equivalent to a JCE. We also show that JCE and NWBR are essentially equivalent in the special but important class of *co-monotonic* signaling games. In many co-monotonic games, JCE selects only equilibria that approximate the least-cost separating equilibria. While existing refinements like D1 and NWBR also make this selection in these games, JCE provides the first learning justification for this selection. Moreover, JCE is a strict subset of the RCE of Fudenberg and He [2020a], which is the strongest existing equilibrium refinement for signaling games that has a learning foundation. We summarize work on learning foundations in Section 5.

2 Model

2.1 Signaling Games

We study signaling games between a sender (player 1) and a receiver (player 2). The sender has a finite type space Θ, a finite signal space S, and a finite message space M. The receiver has a finite action space A. The utility function of the sender is $u_1 : \Theta \times S \times A \to \mathbb{R}$ and the utility function of the receiver is $u_2 : \Theta \times S \times A \to \mathbb{R}$. Both players' utilities depend on the signal and type of the sender and the action of the receiver. Neither player's utility depends directly on the message of the sender. The full-support prior distribution over the sender's type is $\lambda \in \Delta(\Theta)$.

The sender first observes their type and then chooses a signal $s \in S$ and message $m \in M$. The receiver observes the sender's choice of (s, m), but not the sender's type, then selects their action $a \in A$, after which payoffs are realized. Throughout, we denote the set of pure strategies for a fixed sender type by $X = S \times M$ and the set of pure receiver strategies by $Y = A^{S \times M}$. We also denote the set of sender behavior strategies by $\Pi_1 = \Delta(S \times M)^{\Theta}$, and the set of receiver behavior strategies by $\Pi_2 = \Delta(A)^{S \times M}$. Finally, we write $u_1(\theta, \pi)$ and $u_2(\pi)$ for the expected payoffs from strategy profile π, and we let

$$BR(p, s) = \arg\max_{a \in A} u_2(p, s, a) \text{ and } BR(\widetilde{\Theta}, s) = \cup_{p \in \Delta(\widetilde{\Theta})} BR(p, s)$$

denote the best responses for the receiver to signal s under belief $p \in \Delta(\Theta)$ or all p with support $\widetilde{\Theta}$ respectively.

2.2 Learning Environment

We consider an overlapping generations learning environment where time is discrete and doubly infinite, $t \in \{..., -2, -1, 0, 1, 2, ...\}$, and where, for each $\theta \in \Theta$, there is a continuum of mass $\lambda(\theta)$ of agents in the role of a type θ sender, and there is a continuum of agents of mass 1 in the receiver role. The agents have geometric lifespans: agents in sender roles have continuation probability $\gamma_1 \in [0, 1)$, while agents in the receiver role have continuation probability $\gamma_2 \in [0, 1)$. Each period newborn agents replace the departing agents so the sizes of the various populations are constant.

Every period agents are anonymously matched into sender-receiver pairs: Each sender agent is equally likely to be paired with any of the current receiver agents and vice-versa. Within each match, the sender plays a signal $s \in S$ and a message $m \in M$. The receiver observes the sender's choice of (s, m) and then responds with some action $a \in A$.

At the end of each period, both players in a given match observe its outcome, which consists of the type of the sender, the signal and message chosen by the sender, and the action chosen by the receiver. At the beginning of their lives, sender agents have a non-doctrinaire prior $g_1 \in \Delta(\Pi_2)$ over the receiver's behavior strategies, while receiver agents have a non-doctrinaire prior $g_2 \in \Delta(\Delta(\Theta \times S \times M))$ over the distribution of sender type, signal, and message.[3] Upon observing the outcome of a match, agents update their beliefs in accordance with Bayes' rule, which is always applicable because the priors assign positive probability to any finite sequence of observations. Define $\mathcal{H}_{1,t} = (S \times M \times A)^t$ to be the set of histories that a sender of age t could have observed, and let $\mathcal{H}_1 = \cup_{t \in \mathbb{N}} \mathcal{H}_{1,t}$ be the collection of all such sequences. Likewise, the relevant pieces of information for the receiver are the type, signal choice, and message choice of the sender. Let $\mathcal{H}_{2,t} = (\Theta \times S \times M)^t$ denote the set of possible sequences of such triples that a receiver agent with age t could have observed, and let $\mathcal{H}_2 = \cup_{t \in \mathbb{N}} \mathcal{H}_{2,t}$ be the collection of all such sequences.

All agents are rational Bayesians who maximize their expected total payoff. Agents in the role of a type θ sender use some *policy* $\mathbf{x}_\theta : \mathcal{H}_1 \to X$ that is optimal given the prior g_1, continuation probability γ_1, and discount factor $\delta \in [0, 1)$. Agents in the receiver role use a *policy* $\mathbf{y} : \mathcal{H}_2 \to Y$ that is optimal given the prior g_2. Because the receivers observe the type of the sender at the end of each match, their optimal policy is simply to choose an action that maximizes their expected payoff in the current match. This is not the case for the senders: Because the senders' observations depend on their play, the senders have an incentive to "experiment" with various signal-message pairs that have the potential to lead to higher-than-equilibrium payoffs. The force behind our equilibrium refinement is that these sets of potentially improving deviations are different for different sender types.

[3] Here "non-doctrinaire" means "described by a continuous density function that is strictly positive on the interior of the probability simplex." Because we allow for correlated beliefs, it would be equivalent to view the receivers' beliefs as being over the distribution of sender types and sender behavior strategies.

2.3 Steady States and Aggregate Play

At every period t, the *state* of the system, denoted $\mu_t = (\mu_{1,t}, \mu_{2,t}) \in \Delta(\mathcal{H}_1)^{\Theta} \times \Delta(\mathcal{H}_2)$, is the shares of agents in a given player role with the various possible histories. Given μ_t the profile $\mathbf{x} = \{\mathbf{x}_\theta\}_{\theta\in\Theta}$ of sender policies induces a sender behavior strategy $\sigma_1(\mu_{1,t}) \in \Pi_1$ that we call the *aggregate sender play.* Similarly, the receiver policy $\mathbf{y}$ induces a receiver behavior strategy $\sigma_2(\mu_{2,t}) \in \Pi_2$ that we call the *aggregate receiver play.* We call $\sigma(\mu_t) = (\sigma_1(\mu_{1,t}), \sigma_2(\mu_{2,t})) \in \Pi_1 \times \Pi_2$ the *aggregate strategy profile.* (Online Appendix Section OA.1 gives formal definitions of the mappings $\sigma_1 : \Delta(\mathcal{H}_1)^{\Theta} \to \Pi_1$ and $\sigma_2 : \Delta(\mathcal{H}_2) \to \Pi_2$, as well as other objects introduced in this subsection.)

A policy profile generates an *update rule* $\mathbf{f} : \Delta(\mathcal{H}_1)^{\Theta} \times \Delta(\mathcal{H}_2) \to \Delta(\mathcal{H}_1)^{\Theta} \times \Delta(\mathcal{H}_2)$, taking the state in period t to the state in period $t+1$, a mapping $\mathscr{R}_1^{\delta,\gamma_1,\gamma_2} : \Pi_2 \to \Pi_1$ that describes the aggregate play of the senders when the aggregate play of the receivers is fixed at π_2, and a mapping $\mathscr{R}_2^{\delta,\gamma_1,\gamma_2} : \Pi_1 \to \Pi_2$ from the aggregate sender strategy to the aggregate receiver strategy. We refer to the mapping $\mathscr{R}^{\delta,\gamma_1,\gamma_2}(\pi_1, \pi_2) \equiv (\mathscr{R}_1^{\delta,\gamma_1,\gamma_2}(\pi_2), \mathscr{R}_2^{\delta,\gamma_1,\gamma_2}(\pi_1))$ as the *aggregate response mapping.* Straightforward arguments show that these mappings are continuous.

We study this system's steady states, those μ satisfying $\mathbf{f}(\mu) = \mu$, and denote the corresponding *steady state profiles* for given priors and continuation probabilities by $\Pi^*(g, \delta, \gamma_1, \gamma_2) \subseteq \Pi_1 \times \Pi_2$.

Proposition 1. $\Pi^*(g, \delta, \gamma_1, \gamma_2)$ *consists of the strategy profiles that are fixed points of the aggregate response mapping, and it is non-empty for all* $g = (g_1, g_2)$*,* δ*, and* γ_1, γ_2*.*

The proof follows standard lines and is omitted.

We will study the limit of steady state play when both senders and receivers live a long time, so they can acquire enough observations to outweigh their prior.[4] In addition, we assume that the receivers play much less often than the senders. This fits settings where the senders are institutions who, over time, interact with a large number

[4] We need to let the agents acquire arbitrarily large data sets for the steady state profiles to be Nash equilibria, because we do not restrict the initial confidence of their priors, so it may take a great many observations before their posteriors reflect their data.

of individuals in the role of the receivers, as with firms who signal their knowledge about their productivity, future growth, etc. using offers of incentive pay to potential employees. While employees may interact with a large number of firms over their lifetime, or observe family members and other relations do so, it is unlikely that any given individual will be involved in (or have access to information concerning) as many interactions as the typical large firm.

Formally, we consider the limit of steady state aggregate play in the iterated limit $\lim_{\gamma_2 \to 1} \lim_{\delta \to 1} \lim_{\gamma_1 \to 1} \Pi^*(g, \delta, \gamma_1, \gamma_2)$.[5] We will call these the *stable profiles.*

Definition 1. $\pi = (\pi_1, \pi_2) \in \Pi_1 \times \Pi_2$ *is* **stable** *if there is a sequence* $\{\gamma_{2,j}\}_{j \in \mathbb{N}} \to 1$*, sequences* $\{\delta_{j,k}\}_{j,k \in \mathbb{N}}$ *with* $\lim_{j \to \infty} \delta_{j,k} = 1$ *for all* j*, and sequences* $\{\gamma_{1,j,k,l}\}_{j,k,l \in \mathbb{N}}$ *with* $\lim_{l \to \infty} \gamma_{1,j,k,l} = 1$ *for all* j, k*, such that* $\pi = \lim_{j \to \infty} \lim_{k \to \infty} \lim_{l \to \infty} \pi_{j,k,l}$ *for some sequence* $\pi_{j,k,l} \in \Pi^*(g, \delta_{1,j,k}, \gamma_{1,j,k,l}, \gamma_{2,j})$*.*

A corollary of Proposition 1 is that there are stable strategy profiles.

Corollary 1. *Stable strategy profiles exist.*

Fudenberg and Levine [1993, 2006] and Fudenberg and He [2018, 2020a,b] studied a different notion of stability, where all agents have the same expected lifetimes.[6] In our case of senders who are much longer lived than receivers, most receivers never encounter young senders. Young senders are most likely to "experiment" and play signal-message pairs that are off-path according to the limit profile. This means that most receivers have little experience with off-path play by the senders, which facilitates the analysis of the stable profiles. When receivers have the same or longer expected lifetimes as the senders they are likely to encounter some young senders, which makes the corresponding limits more difficult to analyze.

[5]We conjecture that our results would also hold with $\delta \equiv 1$ but have not been able to prove this.

[6]These papers set $\gamma_1 = \gamma_2 = \gamma$ and studied the iterated limit in which first agents become arbitrarily long-lived ($\gamma \to 1$) and then arbitrarily patient ($\delta \to 1$), so the receiver's continuation parameter converges to 1 faster than the sender's patience level. When γ_1 converges to 1 faster than γ_2, this is no longer needed. We focus on the iterated limit where δ converges to 1 before γ_2 does because it affords slightly cleaner results and greater analytical simplicity. We discuss modifications to this in Section 6.

2.4 Key Assumptions

Our results about stable profiles use two additional assumptions. First, we assume that the sender message space is rich enough that every subset of sender types has a corresponding message.

Assumption 1. (Richness) $|M| \geq 2^{|\Theta|}$.

Assumption 1 implies that for each signal $s \in S$ and subset of sender types $\widetilde{\Theta} \subseteq \Theta$ there is a message $m_{s,\widetilde{\Theta}} \in M$ that can be interpreted as saying "my type is in the set $\widetilde{\Theta}$."[7] Our next assumption is that when the sender plays s and sends the message $m_{s,\widetilde{\Theta}}$, the receiver "trusts" the message provided that they have not previously encountered any sender with any other type $\theta \notin \widetilde{\Theta}$ playing signal s and sending message $m_{s,\widetilde{\Theta}}$.

Assumption 2. (Initially Trusting) *For every $s \in S$ and $\widetilde{\Theta} \subseteq \Theta$, there is some $m_{s,\widetilde{\Theta}} \in M$ such that $\mathbf{y}(h_2)[s, m_{s,\widetilde{\Theta}}] \in BR(\widetilde{\Theta}, s)$ for every $h_2 \in \mathcal{H}_2$ in which, for all $\theta' \notin \widetilde{\Theta}$, $(\theta', s, m_{s,\widetilde{\Theta}})$ has not been observed.*

Without any assumptions on the receiver's prior, stability offers little predictive power beyond PBE-H, as we show by example in D.1. The assumption of initial trust amounts to an implicit restriction on the receiver prior which can be overturned with little data; whether it is in fact overturned depends on the data that the agents observe. It seems plausible that receivers may give the sender the "benefit of the doubt" and act in accordance with claims they have not previously seen proved false. This does not require that the receivers are certain that these messages are true, only that they give them a sufficiently high probability of being true. Of course, the receiver may quickly learn to distrust such messages if they do prove to be false, which is why Assumption 2 is only applied to claims for which no contradictory evidence exists.[8] We maintain

[7] "My type is in this set" need not be the literal content of the message. For instance, $m_{s,\widetilde{\Theta}}$ could represent an argument like e.g. "I played signal s so you should believe my type is in $\widetilde{\Theta}$ because..."

[8] This assumption is similar to the "believe-unless-refuted" condition of Lipman and Seppi [1995], and is also related to notions of credibility in Rabin [1990], Farrell [1993], and Clark [2020]. We discuss these connections in more detail in Section 5.

Assumptions 1 and 2 throughout the main text; in Section 6 we discuss an alternative assumption to initial trust which gives a similar refinement to JCE.

3 Justified Communication Equilibrium

This section defines justified communication equilibrium, and states and proves our main result: Every stable profile is a justified communication equilibrium. Informally, justified communication equilibrium amounts to adding two additional restrictions to Nash equilibrium. First, at off-path signal-message pairs, every action the receiver plays with positive probability must be a best response to some belief about the sender's type. Second, for each off-path signal there must be a "justified receiver response" that deters all types from playing it. We need some additional notation to state these conditions formally.

3.1 Notation and Definitions

The set of actions that are a best response to some belief about θ is $BR(\Theta, s)$; these are the *undominated responses* to s; the other responses are *conditionally dominated* in the sense of Fudenberg and Tirole [1991b]. Thus $\Delta(BR(\Theta, s))$ is the set of receiver strategies that assign probability 1 to undominated responses. The set of mixed best responses to s for beliefs supported on $\widetilde{\Theta}$ is

$$MBR(\widetilde{\Theta}, s) = \{\alpha \in \Delta(A) : \exists p \in \Delta(\widetilde{\Theta}) \text{ s.t. } u_2(p, s, \alpha) \geq u_2(p, s, a) \ \forall a \in A\}.$$

Note that $MBR(\widetilde{\Theta}, s)$ is in general different from $\Delta(BR(\widetilde{\Theta}, s))$, the set of mixed actions with support contained in $BR(\widetilde{\Theta}, s)$. The difference between these two sets is important, as we will see.

Definition 2 (Fudenberg and He [2018])**.** *Strategy profile* $\pi = (\pi_1, \pi_2)$ *is a* **perfect Bayesian equilibrium with heterogeneous off-path beliefs (PBE-H)** *if*

1. *For each* $\theta \in \Theta$, $u_1(\theta, \pi) = \max_{(s,m)\in S\times M} u_1(\theta, s, \pi_2(\cdot|s, m))$.
2. *For each on-path signal-message pair* (s, m) *and each* $a \in A$ *with* $\pi_2(a|s, m) > 0$, $a \in BR(p(\cdot|s, m), s)$, *where* $p(\cdot|s, m)$ *is the posterior belief given* (s, m) *obtained through Bayes' rule.*
3. *For each off-path signal-message pair* (s, m) *and each* $a \in A$ *with* $\pi_2^*(a|s, m) > 0$, *there exists a belief* $p \in \Delta(\Theta)$ *such that* $a \in BR(p, s)$.

Conditions 1 and 2 of Definition 2 are the conditions for a strategy profile to be a Nash equilibrium. Condition 3 lets the receiver's response to an off-path signal-message pair (s, m) be a mixture over several actions, each of which is a response to a possibly different belief about the sender's type. For this reason, Conditions 1–3 together are slightly weaker than perfect Bayesian equilibria (PBE, Fudenberg and Tirole [1991a]), which replaces 3 with the condition that the receiver response to each s is in $MBR(\Theta, s)$.[9]

Justified communication equilibrium adds the "justified-response" condition to PBE-H. To define this condition, for each (s, π) let

$$\widetilde{D}_\theta(s, \pi) = \{\alpha \in \Delta(BR(\Theta, s)) : u_1(\theta, s, \alpha) > u_1(\theta, \pi)\}.$$

This is the set of mixtures over receiver responses to s that are optimal for some belief over Θ and would make type θ strictly prefer s to their outcome under π.[10] Let

$$\widetilde{D}^0_\theta(s, \pi) = \{\alpha \in \Delta(BR(\Theta, s)) : u_1(\theta, s, \alpha) = u_1(\theta, \pi)\}$$

be the corresponding set for which type θ would be indifferent with their outcome under π. For every $s \in S$ and $\pi \in \Pi_1 \times \Pi_2$, let

$$\Theta^\dagger(s, \pi) = \{\theta \in \Theta : \widetilde{D}^0_\theta(s, \pi) \not\subseteq \cup_{\theta' \neq \theta} \widetilde{D}_{\theta'}(s, \pi)\}$$

[9] Recall that PBE and sequential equilibria are equivalent in signaling games. Fudenberg and He [2018] showed that PBE-H is a necessary condition for learning outcomes in their setting.

[10] This set is very similar to the set D_θ used by Banks and Sobel [1987] to define divine equilibrium; we discuss the differences in Section 3.3.

be the set of types θ where there is some mixed receiver action $\alpha \in \Delta(BR(\Theta, s))$ that makes θ indifferent between (s, α) to their outcome under π and no other type θ' strictly prefers (s, α) to their outcome under π.

Definition 3. *The set of* **justified types** *for signal s given profile π is*

$$\overline{\Theta}(s, \pi) = \begin{cases} \Theta^{\dagger}(s, \pi) & \textit{if } \Theta^{\dagger}(s, \pi) \neq \emptyset \\ \Theta & \textit{if } \Theta^{\dagger}(s, \pi) = \emptyset \end{cases}.$$

The **justified responses** $\alpha \in \Delta(BR(\overline{\Theta}(s, \pi), s))$ *assign positive probability only to actions that are best responses to beliefs with support in* $\overline{\Theta}(s, \pi)$.

Definition 4. *The strategy profile $\pi = (\pi_1, \pi_2)$ is a* **justified communication equilibrium (JCE)** *if*

1. *It is a PBE-H.*
2. *For each $s \in S$, there is some $m \in M$ such that $\pi_2(\cdot|s, m) \in \Delta(BR(\overline{\Theta}(s, \pi), s))$.*

This says that in a JCE, the receiver's response to each signal is justified for at least one message. Since every JCE is a PBE-H, there is a justified receiver response to every signal that gives each sender type a weakly lower payoff than they obtain under the JCE. Note that each type that plays an on-path signal is a justified type for that signal, so the condition of Definition 5 holds automatically for these signals. The substance of the definition comes from its requirements for off-path signals. We now show that these restrictions are satisfied by stable strategy profiles given Assumption 2.

3.2 Only Justified Communication Equilibria are Stable

Theorem 1. *If $\pi = (\pi_1, \pi_2) \in \Pi_1 \times \Pi_2$ is stable, then it is a justified communication equilibrium.*

Condition 3 of Definition 2 follows from the fact that the receivers in our model myopically optimize because their observations do not depend on their play. We establish the two other conditions of Definition 2, as well as the additional requirement in Definition 4, using three supporting lemmas, the proofs of which are in A.1.

To establish that every stable profile is a perfect Bayesian equilibrium with heterogeneous off-path beliefs, we first show that the aggregate sender play is optimal given the aggregate receiver play.

Lemma 1. *Suppose that* $\pi = (\pi_1, \pi_2) \in \Pi_1 \times \Pi_2$ *is stable. Then Condition 1 of Definition 2 holds: For each* $\theta \in \Theta$*,* $\pi_1(\cdot|\theta)$ *puts support only on those sender signal-message pairs that are optimal for type* θ *under the receiver behavior strategy* π_2*.*

The proof of Lemma 1 first shows that for fixed $\gamma_2 \in [0, 1)$, aggregate sender play is optimal given the aggregate receiver play in the iterated limit where first $\gamma_1 \to 1$ then $\delta \to 1$. As in Fudenberg and Levine [1993], this holds because each sender type will experiment enough to drive the option value of experimentation to 0, so that aggregate sender play is optimal in the limit. The conclusion of Lemma 1 then follows from combining this optimality condition with the fact that the sender best response has a closed graph.

The next lemma shows that in a stable profile, aggregate receiver play is a best response to (on-path) aggregate play by the senders.

Lemma 2. *Suppose that* $\pi = (\pi_1, \pi_2) \in \Pi_1 \times \Pi_2$ *is stable. Then Condition 2 of Definition 2 holds: For any sender signal-message pair* $(s, m) \in S \times M$ *that occurs with positive probability under* π*,* $\pi_2(\cdot|s, m)$ *puts support only on receiver actions that are best-responses to* s *and the posterior belief induced by* λ *and* $\{\pi_1(s, m|\theta)\}_{\theta \in \Theta}$ *under Bayes' rule.*

The proof of Lemma 2 shows that receivers will get enough observations of on-path play for their data to swamp their priors. By the law of large numbers their sample converges to the population distribution, and since receivers myopically optimize the lemma follows.

Neither Lemma 1 nor Lemma 2 requires Assumptions 1 and 2. Lemma 1 implies Condition 1 of the definition of JCE, and Lemma 2 implies Condition 2. Since we have already noted that Condition 3 holds, we conclude that a stable profile must be a PBE-H.

However, the next lemma does require both assumptions. The lemma shows that, for fixed $s \in S$ and $\widetilde{\Theta} \subseteq \Theta$, if every type $\theta \notin \widetilde{\Theta}$ strictly prefers their payoff under π to their payoff from playing $(s, m_{s,\widetilde{\Theta}})$ (and having the receiver respond with $\pi_2(\cdot|s, m_{s,\widetilde{\Theta}})$), then the aggregate receiver response to $(s, m_{s,\widetilde{\Theta}})$ must be supported on $BR(\widetilde{\Theta}, s)$.

The proof of the lemma, and thus of Theorem 1, fails without Assumption 2, and *a fortiori* without any cheap-talk messages at all. Moreover, Example D.1 shows that, without initial trust, there can be stable profiles that are not JCE. This is why the learning foundation of JCE requires cheap-talk communication and initial trust.

Lemma 3. *Suppose that $\pi = (\pi_1, \pi_2) \in \Pi_1 \times \Pi_2$ is stable. Fix $s \in S$ and $\widetilde{\Theta} \subseteq \Theta$. If $u_1(\theta, s, \pi_2(\cdot|s, m_{s,\widetilde{\Theta}})) < u_1(\theta, \pi)$ for all $\theta \notin \widetilde{\Theta}$, then $\pi_2(BR(\widetilde{\Theta}, s)|s, m_{s,\widetilde{\Theta}}) = 1$.*

To get some intuition for this result, note that, when $u_1(\theta, s, \pi_2(\cdot|s, m_{s,\widetilde{\Theta}})) < u_1(\theta, \pi)$ for all $\theta \notin \widetilde{\Theta}$, Lemma 2 implies that the aggregate probability that a type outside of $\widetilde{\Theta}$ plays $(s, m_{s,\widetilde{\Theta}})$ is small if the prevailing aggregate strategy profile is close to π, δ is close to 1, and, given δ, γ_1 is close to 1. For a fixed receiver continuation probability $\gamma_2 < 1$, the share of receivers in the population who have witnessed a sender with type outside of $\widetilde{\Theta}$ play the signal-message pair $(s, m_{s,\widetilde{\Theta}})$ becomes arbitrarily small as the aggregate probability of such play by types outside of $\widetilde{\Theta}$ approaches 0. Recall that receivers who have never observed a type outside of $\widetilde{\Theta}$ play $(s, m_{s,\widetilde{\Theta}})$ would respond to $(s, m_{s,\widetilde{\Theta}})$ with some action in $BR(\widetilde{\Theta}, s)$. Combining these facts, it follows that the share of receivers who play some action in $BR(\widetilde{\Theta}, s)$ in response to $(s, m_{s,\widetilde{\Theta}})$ becomes arbitrarily close to 1 in the iterated limit where $\gamma_1 \to 1$ then $\delta \to 1$ then $\gamma_2 \to 1$.

Proof of Theorem 1. We have already established that π is a PBE-H. We now show that the condition in Definition 4 holds.

Let $\{\pi_{j,k,l} \in \Pi^*(g, \delta_{j,k}, \gamma_{1,j,k,l}, \gamma_{2,j})\}_{j,k,l \in \mathbb{N}}$ be a sequence of steady state profiles such that $\lim_{j\to\infty} \lim_{k\to\infty} \lim_{l\to\infty} \pi_{j,k,l} = \pi$, where $\lim_{j\to\infty} \gamma_{2,j} = 1$, $\lim_{k\to\infty} \delta_{j,k} = 1$ for all j, and $\lim_{l\to\infty} \gamma_{1,j,k,l} = 1$ for all j, k. Fix $s \in S$. Since $u_1(\theta, s, \pi_2(\cdot|s, m_{s,\overline{\Theta}(s,\pi)})) \leq u_1(\theta, \pi)$

holds for all $\theta \in \overline{\Theta}(s,\pi)$, it must be that $u_1(\theta, s, \pi_2(\cdot|s, m_{s,\overline{\Theta}(s,\pi)})) < u_1(\theta, \pi)$ for all $\theta \notin \overline{\Theta}(s,\pi)$. Lemma 3 then implies that $\pi_2(\cdot|s, m_{s,\overline{\Theta}(s,\pi)}) \in \Delta(BR(\overline{\Theta}(s,\pi), s))$. ■

Theorem 1 shows that only JCE can be stable. This means that we can safely eliminate strategy profiles that are not JCE, and, as we will see in the next subsection, JCE rule out strategy profiles that commonly used refinements like the Intuitive Criterion and D1 preserve. The following example shows that stable profiles need not be PBE. This means that sequential equilibrium and its refinements eliminate too many profiles, as they are not necessary conditions for stability.[11]

Example 1. The sender's type space is $\Theta = \{\theta_1, \theta_2\}$, and the receiver's prior is that both types are equally likely. The sender's signal space is $S = \{In, Out\}$, their message space contains $\{m_{\theta_1}, m_{\theta_2}, m_\Theta\}$, and the receiver's action space is $A = \{a_1, a_2, a_3\}$. The payoffs to the sender and receiver are given in Table 1 below.

θ_1	a_1	a_2	a_3
In	$-2, 2$	$1, -2$	$1, .3$
Out	$0, 0$	$0, 0$	$0, 0$

θ_2	a_1	a_2	a_3
In	$1, -2$	$-2, 2$	$1, .3$
Out	$0, 0$	$0, 0$	$0, 0$

Table 1: The payoffs for Example 1.

This game does not have a perfect Bayesian equilibrium in which both types play Out, because $u_1(\theta_1, In, \alpha) \leq 0$ only if $\alpha[a_1] > 0$ and $u_1(\theta_2, In, \alpha) \leq 0$ only if for $\alpha[a_2] > 0$, yet there is no mixed best response to In where the receiver assigns positive probability to both a_1 and a_2. However, the profile π in which both sender types play Out and the receiver always responds to In with $(1/2)a_1 + (1/2)a_2$ is a justified communication equilibrium, because $\overline{\Theta}(In, \pi) = \Theta$, and $a_1, a_2 \in BR(\Theta, s)$. Moreover, C.1 shows that both types playing Out can be a stable outcome. □

Even though we give direct proofs of stability in most of our examples, as in Example 1 above, it is difficult to give very general sufficient conditions for stability. It is

[11] The equilibrium refinements in Fudenberg and He [2018] and Fudenberg and He [2020a] also relax PBE to PBE-H, but those papers do not show that this relaxation is needed.

nevertheless possible to give sufficient conditions for stability in some classes of games. Specifically, Appendix B shows that, in "strictly monotonic" games, all JCE that satisfy a "uniform justification" requirement are outcome-equivalent to stable profiles.[12]

3.3 Relation to Other Equilibrium Refinements

By definition, JCE need only be PBE-H and not perfect Bayesian equilibria. The Intuitive Criterion (Cho and Kreps [1987]), D1 (Banks and Sobel [1987]), and NWBR (Kohlberg and Mertens [1986], Cho and Kreps [1987]) were all formulated as refinements of sequential equilibrium. However, the procedures they use to restrict out-of-equilibrium beliefs and equilibrium outcomes can be adapted to develop tests for any PBE-H. When this is done, we can more naturally compare the predictions of the modified versions of these refinements with JCE.

Figure 1 presents a visual summary of the general relationships between JCE and appropriately modified versions of the previously mentioned refinements, as well as rationality-compatible equilibrium (RCE, Fudenberg and He [2020a]), the strongest existing signaling game refinement with a learning foundation. In particular, every JCE is a PBE-H that satisfies RCE and a modified version of the Intuitive Criterion we call the *Intuitive Criterion Test (IC Test).* JCE and D1 are not nested: there are some JCE which do not satisfy D1 and there are some PBE-H satisfying D1 that are not JCE. We provide detailed definitions of these refinements and formal proofs of their relationships with JCE in Appendix C. For the remainder of this subsection, we focus on analyzing the relationship of JCE and NWBR, because as we will see these two refinements are very similar. Moreover, as Section 4 shows, there is a broad class of games in which JCE and NWBR are equivalent and make particularly strong predictions.

To compare NWBR with JCE, we first develop some notation. For every $\theta \in \Theta$,

[12] Moreover, this is true for all non-doctrinaire priors, even those that do not satisfy initial trust.

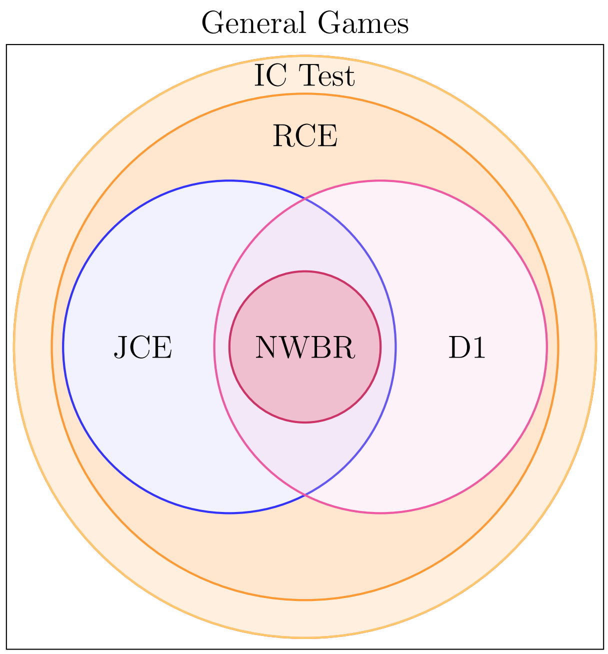

Figure 1: JCE and other restrictions of PBE-H.

$s \in S$, and $\pi \in \Pi_1 \times \Pi_2$, let

$$D_\theta(s, \pi) = \{\alpha \in MBR(\Theta, s) : u_1(\theta, s, \alpha) > u_1(\theta, \pi)\}$$

be the set of mixed receiver best responses to s that are optimal for some belief over Θ and would make type θ strictly prefer s to their equilibrium outcome. Let

$$D^0_\theta(s, \pi) = \{\alpha \in MBR(\Theta, s) : u_1(\theta, s, \alpha) = u_1(\theta, \pi)\}$$

be the corresponding set for which type θ would be indifferent with their equilibrium outcome. For every $s \in S$ and $\pi \in \Pi_1 \times \Pi_2$, let $\Theta^\ddagger(s, \pi) = \{\theta \in \Theta : D^0_\theta(s, \pi) \not\subseteq \cup_{\theta' \neq \theta} D_{\theta'}(s, \pi)\}$ be the set of types θ where there is some mixed receiver best response $\alpha \in MBR(\Theta, s)$ that makes θ indifferent between (s, α) and their equilibrium outcome, and no other type θ' strictly prefers (s, α) to their equilibrium outcome.

Let $\widehat{\Theta}(s,\pi) \subseteq \Theta$ be the set of types given by

$$\widehat{\Theta}(s,\pi) = \begin{cases} \Theta^{\ddagger}(s,\pi) & \text{if } \Theta^{\ddagger}(s,\pi) \neq \emptyset \\ \Theta & \text{if } \Theta^{\ddagger}(s,\pi) = \emptyset. \end{cases}$$

Definition 5 (Kohlberg and Mertens [1986], Cho and Kreps [1987])**.** *Strategy profile π passes the* **never a weak best response (NWBR)** *criterion if, for every $s \in S$, there is some $\alpha \in MBR(\widehat{\Theta}(s,\pi), s)$ such that $u_1(\theta,\alpha) \leq u_1(\theta,\pi)$ for all $\theta \in \Theta$.*

Compared to $\overline{\Theta}$, which is defined using $\widetilde{D}_\theta$ and $\widetilde{D}^0_\theta$, $\widehat{\Theta}$ differs in that the set of mixed best responses $MBR(\Theta, s)$ is replaced by the convex hull of the best responses $\Delta(BR(\Theta, s))$. This difference, along with the difference between sequential equilibrium and PBE-H, account for how JCE differs from NWBR. In fact, NWBR is a stronger requirement than JCE, since $\widehat{\Theta} \subseteq \overline{\Theta}$. Thus every profile that satisfies NWBR is path-equivalent to a justified communication equilibrium.[13]

Proposition 2. *If $\pi = (\pi_1, \pi_2) \in \Pi_1 \times \Pi_2$ is a PBE-H that satisfies NWBR, then π is path-equivalent to a JCE.*

A.2 proves Proposition 2, and Example OA 3 in Online Appendix Section OA.6.3 shows that there can be JCE that are sequential equilibria but do not satisfy NWBR.[14]

An immediate consequence of Proposition 2 is that the requirements of NWBR and those of JCE are equivalent whenever $MBR(\Theta, s) = \Delta(BR(\Theta, s))$ for all s. One setting that satisfies this condition is when there are at most two undominated receiver responses to each signal.

[13] A PBE-H strategy profile that satisfies NWBR is not necessarily a PBE, since the receiver's response to off-path play need not be optimal given any single belief over the sender's type. However, every such profile is path-equivalent to a PBE since the receiver's response to off-path play can always be replaced by the mixed best response satisfying NWBR that deters all sender types.

[14] Fudenberg and Kreps [1988] and Sobel, Stole, and Zapater [1990] recognized that the convex hull of the best responses is more natural in a learning setting, and discussed some of the differences this could make, but neither paper explicitly showed that NWBR rules out a profile that is stable in a learning model.

Proposition 3. *In any game in which* $MBR(\Theta, s) = \Delta(BR(\Theta, s))$ *for all* $s \in S$*, a strategy profile* $\pi = (\pi_1, \pi_2) \in \Pi_1 \times \Pi_2$ *is a justified communication equilibrium if and only if it is path-equivalent to a PBE-H that satisfies NWBR.*

Section 4 shows that the equivalence between JCE and NWBR holds, albeit for quite different reasons, in a broad class of "co-monotonic" games.

4 Co-Monotonic Signaling Games

Here we restrict attention to *co-monotonic* signaling games, which are games in which all sender types share the same preference over mixed receiver actions with support in $BR(\Theta, s)$.

Definition 6. *A signaling game is* **co-monotonic** *if, for all* $\theta, \theta' \in \Theta$*,* $s \in S$*, and* $\alpha, \alpha' \in \Delta(BR(\Theta, s))$*,* $u_1(\theta, s, \alpha) \geq u_1(\theta, s, \alpha')$ *if and only if* $u_1(\theta', s, \alpha) \geq u_1(\theta', s, \alpha')$*.*

This is a subset of the *monotonic* signaling games studied in Cho and Sobel [1990], where the sender types are required to share the same preference only over the receiver mixed best responses $MBR(\Theta, s)$ rather than the convex hull of those responses.[15]

A sufficient condition on sender preferences for a signaling game to be co-monotonic is that there be functions $v : S \times A \to \mathbb{R}$, $\omega : \Theta \times S \to \mathbb{R}_{++}$, and $\nu : \Theta \times S \to \mathbb{R}$ such that $u_1(\theta, s, a) = \omega(\theta, s)v(s, a) + \nu(\theta, s)$ for all $\theta \in \Theta$, $s \in S$, $a \in A$. Many games, including the following simple economic example, satisfy this condition.

Example 2. The sender's type space is $\Theta = \{1, 2, 3\}$, and the receiver's prior is that all types are equally likely. The sender's signal space is $S = \{0, 10, 20, ..., 100\}$, and the receiver's action space is $A = \{0, 10, 20, ..., 60\}$. The payoffs to the sender and receiver are $u_1(\theta, s, a) = \theta a - s$ and $u_2(\theta, s, a) = \theta a + s - a^2/40$, respectively.

One interpretation of this game is as follows. The sender is a firm, who is better informed about the productivity of effort by their employee, who is the receiver. The

[15] Recall that the convex hull of the mixed best responses, i.e. $co(MBR(\Theta, s))$, equals $\Delta(BR(\Theta, s))$. Also, note that this kind of co-monotonicity is not related to the comonotonicity used in probability theory.

firm's choice of s reflects a payment to the employee that can provide a costly signal of the firm's knowledge, captured by θ. The action a represents the employee's choice of effort level, which is costly to the agent. Both the firm and employee value greater productivity.

OA.5 in the Online Appendix shows that JCE selects only equilibria that approximate the least-cost separating equilibrium of this game. $\square$

We now explore JCE's relationship with other refinements in co-monotonic games. Co-monotonicity implies that, for all s, any mixture over receiver best responses $\alpha \in \Delta(BR(\Theta, s))$ has a corresponding receiver mixed receiver best response $\alpha' \in MBR(\Theta, s)$ such that $u_1(\theta, s, \alpha) = u_1(\theta, s, \alpha')$ for all θ. This ensures that $\overline{\Theta} = \widehat{\Theta}$.

Lemma 4. *In a co-monotonic signaling game,* $\overline{\Theta}(s, \pi) = \widehat{\Theta}(s, \pi)$ *for all* $s \in S$, $\pi = (\pi_1, \pi_2) \in \Pi_1 \times \Pi_2$.

The proof of Lemma 4 is in A.3.

Additionally, in co-monotonic games, all types agree about which receiver best responses are least desirable. Combining this with Lemma 4 shows that JCE and NWBR (Definition 5) select the same profiles up to path-equivalence. The learning foundation for JCE thus provides a foundation for the predictions of NWBR in the class of co-monotonic games.

Proposition 4. *In a co-monotonic signaling game,* π *is a justified communication equilibrium if and only if it is path-equivalent to a PBE-H that satisfies NWBR.*

Proof of Proposition 4. Suppose that π is a PBE-H that satisfies NWBR. Then, for every $s \in S$, there is some $\alpha_s \in MBR(\widehat{\Theta}(s, \pi), s)$ such that $u_1(\theta, s, \alpha_s) \leq u_1(\theta, \pi)$ for all $\theta \in \Theta$. Since the game is co-monotonic, Lemma 4 implies that $\alpha_s \in \Delta(BR(\overline{\Theta}(s, \pi), s))$, so π is a justified communication equilibrium.

If π is a justified communication equilibrium, it is a PBE-H. Moreover, for every $s \in S$, there is some $\alpha_s \in \Delta(BR(\overline{\Theta}(s, \pi), s))$ such that $u_1(\theta, s, \alpha_s) \leq u_1(\theta, \pi)$ for all $\theta \in \Theta$. Because the game is co-monotonic, there exists some $a_s \in BR(\overline{\Theta}(s, \pi), s)$ such that $a_s \in \arg\min_{a \in BR(\overline{\Theta}(s,\pi),s)} u_1(\theta, s, a)$ for all $\theta \in \Theta$. Thus, $u_1(\theta, s, a_s) \leq u_1(\theta, \pi)$ for all $\theta \in \Theta$. Since the game is co-monotonic, Lemma 4 implies that $a_s \in BR(\widehat{\Theta}(s,\pi), s)$, so π is a PBE-H that satisfies NWBR. ■

In co-monotonic signaling games, any outcome emerging from a PBE-H could also be sustained in a PBE by having the receiver respond to off-path play with an action that the various sender types unanimously agree is the worst. Consequently, every PBE-H is path-equivalent to a PBE. This results in the following corollary to Proposition 4.

Corollary 2. *In a co-monotonic signaling game, every justified communication equilibrium is path-equivalent to a PBE that satisfies NWBR.*

Cho and Sobel [1990] showed that NWBR selects the Riley outcome (Riley [1979]) in a class of monotonic games with a continuum of actions. The definition of JCE can be applied as is to signaling games with infinite actions, and the equivalence of JCE and NWBR in Proposition 4 continues to hold in all co-monotonic signaling games. Thus, JCE selects the Riley outcome in all co-monotonic games that satisfy the additional assumptions of Cho and Sobel [1990] and, by a closed graph argument, also only selects equilibria that are close to the Riley outcome when the action space is a sufficiently fine finite grid.

5 Related Work

Fudenberg and Kreps [1988] introduced the analysis of non-equilibrium learning in extensive-form games, and announced a program of deriving equilibrium refinements from learning foundations, but did not provide details. Our steady state formulation is more in the spirit of Fudenberg and Levine [1993]. Fudenberg and Levine [1993] and Fudenberg and Kreps [1994] provided conditions for rational players to do enough experimentation to rule out non-Nash outcomes.[16] Fudenberg and Levine [2006] used

[16]Kalai and Lehrer [1993], Lehrer and Solan [2007], Esponda [2013], Battigalli, Francetich, Lanzani,

a steady-state learning model to study equilibrium refinements in a class of games of perfect information, and showed that all “subgame confirmed” equilibria are stable.

In signaling games without cheap talk, Fudenberg and He [2018] analyzed the steady states of a model where senders and receivers have identically-distributed geometric lifetimes.[17] It assumed that the senders' prior beliefs over receiver strategies are independent, so that the senders' optimal policy is given by the Gittins index (Gittins [1979]). This characterization leads to restrictions on the relative probabilities with which various sender types experiment with different signals, and thus to restrictions on the receivers' off-path beliefs. Fudenberg and He [2020a] extended Fudenberg and He [2018] by supposing that the senders assign probability 0 to receivers playing conditionally dominated actions, and gave a learning foundation for rationality-compatible equilibrium (RCE). RCE is weaker than JCE, as shown in Proposition 9. In particular, in co-monotonic signaling games RCE, unlike JCE, permits equilibria that NWBR rules out. This paper obtains a stronger conclusion without assuming independent priors by explicitly modeling cheap-talk messages and combining the assumptions of initially trusting receivers and relatively long-lived senders.

We view initial trust as a plausible and appealingly simple assumption concerning the interpretation of communication. Initial trust has a similar form to the “believe-unless-refuted” condition of Lipman and Seppi [1995], which is an equilibrium refinement for signaling games with multiple receivers and partial provability. There, each receiver can learn from the refutations provided by other receivers. Initial trust is also related to the restrictions imposed directly by Rabin [1990], Farrell [1993], and Clark [2020] on how receivers respond to certain “credible” messages in signaling game equi-

and Marinacci [2019] study rational learning without assuming that agents are patient. Other papers such as Battigalli [1987], Rubinstein and Wolinsky [1994], Dekel, Fudenberg, and Levine [1999], Esponda [2013], Battigalli, Cerreia-Vioglio, Maccheroni, and Marinacci [2015], Fudenberg and Kamada [2015], and Fudenberg and Kamada [2018] studied equilibrium concepts motivated by rational learning without providing an explicit foundation in a learning model. A separate body of work, e.g. Binmore and Samuelson [1999], Nöldeke and Samuelson [1993], Hart [2002], Jehiel and Samet [2005], studied evolutionary or boundedly rational learning dynamics in extensive form games.

[17] Fudenberg and Levine [1993] assumed agents had fixed finite lifetimes. We do not think the difference is important for our results.

libria.[18] In these papers, common knowledge of the equilibrium to be played figures heavily in determining the credibility of messages. Such restrictions are not applicable to the learning setting of this paper. Moreover, deriving restrictions on equilibria from learning models and assumptions on the receiver's prior yields more insight than directly imposing restrictions on equilibria.

6 Conclusion

We have shown how adding cheap-talk communication to signaling games lets us provide a learning-theoretic foundation for the concept of justified communication equilibrium. We recovered some of the intuitions that underlie traditional equilibrium refinements, whose predictions were by and large sensible in the games where they were used. At the same time, we confirmed that some of the worries in the literature about the details of these refinements were well founded, and pointed out how those refinements need to be modified to accord with the implications of non-equilibrium learning.

Of course, there are multiple ways that one can formulate models of non-equilibrium learning, just as there are many definitions of forward induction, and several variants of the Kohlberg and Mertens [1986] axioms. In our opinion, it is easier to judge the plausibility of assumptions on learning models than of axiomatic conditions on equilibrium concepts, especially axioms that are imposed without any reference to how equilibrium play might arise. For this reason, our work makes a valuable contribution even in settings such as co-monotonic signaling games, where the predictions of JCE coincide with those of past work. Outside of those cases, not only does JCE have the benefit of a learning foundation, it is also easier to compute, because it is easier to determine the convex hull of the pure best responses than to determine the smaller

[18] Rabin [1990] and Farrell [1993] only analyzed cheap-talk games, but their refinements can be extended to games where the sender also has costly signals. Matthews, Okuno-Fujiwara, and Postlewaite [1991], Blume and Sobel [1995], Zapater [1997], Olszewski [2006], and Chen, Kartik, and Sobel [2008] also studied refinements in cheap-talk games.

and non-convex set of mixed best responses.

Replacing initial trust with alternative assumptions can lead to similar solution concepts. For example, if receivers know the payoff functions of the senders, as in Fudenberg and He [2020a], then receivers who are long-lived may feel that they have acquired a good sense of each sender type's equilibrium payoff. In Online Appendix Section OA.7.1, we discuss a weakened version of initial trust which only requires receivers to trust previously unencountered claims if they are consistent with the receiver's evaluation of sender incentives. Any stable profile under this assumption must satisfy a refinement that is similar to, but weaker than JCE.

We close with some comments on various extensions of our results. We note first that all of the profiles that we prove are stable in our examples would also be stable under a more general version of the iterated limit where first $\gamma_1 \to 1$ and then $(\gamma_2, \delta) \to (1, 1)$, without putting conditions on the relative speed with which γ_2 and δ converge to 1. Moreover, Theorem 1 continues to apply to all stable profiles that are on-path strict for the receiver, and also to all profiles where each sender type's payoff would not be changed if the receiver deviated, as when all of the types choose an "exit" option.

Also, some papers, such as Kartik, Ottaviani, and Squintani [2006] and Kartik [2009], study signaling games where lying is costly so communication is not "cheap talk." As discussed in OA.7.2, our analysis can be extended to such settings with the complication that the set of justified types for a given signal may not be the same across messages.

Finally, JCE has no cutting power in games where the sender's only actions are cheap-talk messages. We think that developing similar learning foundations for refinements in these games is a promising area for future research. If that is done, it would probably be straightforward to extend the analysis to settings with cheap talk and multiple audiences, as in Goltsman and Pavlov [2011].

A Omitted Proofs

A.1 Proofs of Supporting Results for Theorem 1

We use the following lemma in the proofs of several results, as well as the analysis of several examples.

Lemma 5. *Given $\gamma_2 \in [0,1)$, suppose that $\pi_{\gamma_2} = (\pi_{1,\gamma_2}, \pi_{2,\gamma_2}) = \lim_{k\to\infty} \lim_{l\to\infty} \pi_{\gamma_2,k,l}$ for some sequence of steady state profiles $\pi_{\gamma_2,k,l} \in \Pi^*(g, \delta_k, \gamma_{1,k,l}, \gamma_2)$, where the sequence $\{\delta_k\}_{k\in\mathbb{N}}$ is such that $\lim_{k\to\infty} \delta_k = 1$ and the collection of sequences $\{\gamma_{1,k,l}\}_{k,l\in\mathbb{N}}$ is such that $\lim_{l\to\infty} \gamma_{1,k,l} = 1$ for all k. Then, for each $\theta \in \Theta$, $\pi_{1,\gamma_2}(\cdot|\theta)$ puts support only on those sender signal-message pairs that are optimal for type θ under the receiver behavior strategy π_{2,γ_2}.*

Lemma 5 follows from the proof of Proposition 5 in Fudenberg and He [2018], which establishes that, in a learning setting with the same continuation probability γ across sender and receiver agents, the sender aggregate play is optimal given the receiver aggregate play in any profile which emerges in the iterated limit where first $\gamma \to 1$ then $\delta \to 1$.[19] Precisely the same arguments can be used to establish that, for fixed γ_2, any profile which emerges in the limit where first $\gamma_1 \to 1$ then $\delta \to 1$ must feature optimal sender aggregate play given the receiver aggregate play.

A.1.1 Proof of Lemma 1

Proof. Let $\{\pi_{j,k,l} \in \Pi^*(g, \delta_{j,k}, \gamma_{1,j,k,l}, \gamma_{2,j})\}_{j,k,l\in\mathbb{N}}$ be a sequence of steady state profiles such that $\lim_{j\to\infty} \lim_{k\to\infty} \lim_{l\to\infty} \pi_{j,k,l} = \pi$, where $\lim_{j\to\infty} \gamma_{2,j} = 1$, $\lim_{k\to\infty} \delta_{j,k} = 1$ for all j, and $\lim_{l\to\infty} \gamma_{1,j,k,l} = 1$ for all j, k. By Lemma 5, for every $\theta \in \Theta$, $\pi_{1,\gamma_{2,j}}(\cdot|\theta) = \lim_{k\to\infty} \lim_{l\to\infty} \pi_{1,j,k,l}(\cdot|\theta)$ puts support only on those sender signal-message pairs that

[19] The proof of Proposition 5 in Fudenberg and He [2018] adapts various arguments from Fudenberg and Levine [1993] to a setting with geometrically distributed agent lifetimes, and does not use the assumption that the sender's prior over receiver behavior strategies is independent across the responses to different signals.

are optimal under $\pi_{2,\gamma_{2,j}} = \lim_{k\to\infty}\lim_{l\to\infty}\pi_{2,j,k,l}$. Combining this with the upper hemicontinuity of optimal play implies that, for every $\theta \in \Theta$, $\pi_1 = \lim_{j\to\infty}\lim_{k\to\infty}\lim_{l\to\infty}\pi_{1,j,k,l}$ puts support only on those sender signal-message pairs that are optimal under $\pi_2 = \lim_{j\to\infty}\lim_{k\to\infty}\lim_{l\to\infty}\pi_{2,j,k,l}$. ■

A.1.2 Proof of Lemma 2

Proof. Let $q(\theta, s, m) = \lambda(\theta)\pi_1(s, m|\theta)$ be the distribution over sender type, signal, and message induced by λ and π_1. Also, let X^{on} be the set of sender signal-message pairs that occur with positive probability under π, and let $p(\theta|s, m)$ denote the conditional probability of θ given $(s, m) \in X^{on}$.

For $\varepsilon > 0$, let $Q_\varepsilon = \{q' \in \Delta(\Theta \times S \times M) : |q' - q| \le \varepsilon\}$. By upper hemicontinuity, there exists some $\varepsilon > 0$ such that every receiver whose belief $\widetilde{g}_2 \in \Delta(\Delta(\Theta \times S \times M))$ puts probability at least $1 - \varepsilon$ on Q_ε will respond to every $(s, m) \in X^{on}$ with some action belonging to $BR(p(\cdot|s, m), s)$.

Given the non-doctrinaire prior g_2, Theorem 4.2 of Diaconis and Freedman [1990] implies that there is some $T > 0$ such that a receiver who has lived more than T periods assigns posterior probability of at least $1 - \varepsilon$ to probability distributions q' within $\varepsilon/2$ distance of the empirical distribution they have observed. Fix $\eta > 0$. By the law of large numbers, we can take this T to be such that, with probability at least $1 - \eta$, a receiver who has lived more than T periods assigns probability of at least $1 - \varepsilon$ to probability distributions within $\varepsilon/2$ of the true distribution, regardless of the signal-message pair the receiver observes in the current period.

Let $\{\delta_n\}_{n\in\mathbb{N}}$ be a sequence of sender discount factors, $\{\gamma_{1,n}\}_{n\in\mathbb{N}}$ be a sequence of sender continuation probabilities, $\{\gamma_{2,n}\}_{n\in\mathbb{N}}$ be a sequence of receiver continuation probabilities, and $\pi_n = (\pi_{1,n}, \pi_{2,n}) \in \Pi^*(g, \delta_n, \gamma_{1,n}, \gamma_{2,n})$ a sequence of steady state profiles such that $\lim_{n\to\infty}\gamma_{2,n} = 1$ and $\lim_{n\to\infty}\pi_{1,n} = \pi_1$. The share of receivers in the population who have lived more than T periods, $\gamma_{2,n}^T$, converges to 1 as $n \to \infty$. Moreover, $q_n(\theta, s, m) = \lambda(\theta)\pi_{1,n}(s, m|\theta) \to q$ as $n \to \infty$. Thus, for every $(s, m) \in X^{on}$

and $\eta > 0$, there exists some $N \in \mathbb{N}$ such that $\pi_{2,n}(BR(p(\cdot|s,m),s)|s,m) \geq 1-\eta$ for all $n > N$. ■

A.1.3 Proof of Lemma 3

Proof. Let $\{\pi_{j,k,l} \in \Pi^*(g, \delta_{j,k}, \gamma_{1,j,k,l}, \gamma_{2,j})\}_{j,k,l\in\mathbb{N}}$ be a sequence of steady state profiles such that $\lim_{j\to\infty} \lim_{k\to\infty} \lim_{l\to\infty} \pi_{j,k,l} = \pi$, where $\lim_{j\to\infty} \gamma_{2,j} = 1$, $\lim_{k\to\infty} \delta_{j,k} = 1$ for all j, and $\lim_{l\to\infty} \gamma_{1,j,k,l} = 1$ for all j,k. Since $u_1(\theta, s, \pi_2(\cdot|s, m_{s,\widetilde{\Theta}})) < u_1(\theta,\pi)$ for all $\theta \notin \widetilde{\Theta}$, Lemma 5 implies that there is some $J \in \mathbb{N}$ such that $\lim_{k\to\infty} \lim_{l\to\infty} \pi_{j,k,l}(s, m_{s,\widetilde{\Theta}}|\theta') = 0$ holds for all $\theta' \notin \widetilde{\Theta}$, $j > J$. Receivers who have never observed the signal-message pair $(s, \widetilde{\Theta})$ played by a type outside of $\widetilde{\Theta}$ would respond to this pair with an action belonging to $BR(\widetilde{\Theta}, s)$. Thus if $\lim_{k\to\infty} \lim_{l\to\infty} \pi_{j,k,l}(s, m_{s,\widetilde{\Theta}}|\theta') = 0$ for all $\theta' \notin \widetilde{\Theta}$ then $\lim_{k\to\infty} \lim_{l\to\infty} \pi_{2,j,k,l}(BR(\widetilde{\Theta}, s)|s, m_{s,\widetilde{\Theta}}) = 1$. Since this holds for all $j > J$, we have that $\pi_2(BR(\widetilde{\Theta}, s)|s, m_{s,\widetilde{\Theta}}) = \lim_{j\to\infty} \lim_{k\to\infty} \lim_{l\to\infty} \pi_{2,j,k,l}(BR(\widetilde{\Theta}, s)|s, m_{s,\widetilde{\Theta}}) = 1$. ■

A.2 Proof of Proposition 2

The reason this proposition only shows that a PBE-H that satisfies NWBR is guaranteed to be path-equivalent to a JCE, rather than guaranteed to be a JCE, is that to be a JCE, for every $s \in S$, there must be some $m \in M$ such that $\pi_2(\cdot|s,m) \in \Delta(BR(\overline{\Theta}(s,\pi), s))$. NWBR imposes requirements which must be satisfied by some receiver response that would deter all types from playing a given off-path signal, but does not impose requirements about the receiver's response to off-path play as dictated by the candidate profile.

The proof of Proposition 2 uses the following two lemmas.

Lemma 6. $\Theta^{\ddagger}(s,\pi) \subseteq \Theta^{\dagger}(s,\pi)$ *for all* $s \in S$, $\pi = (\pi_1,\pi_2) \in \Pi_1 \times \Pi_2$.

Proof. If $\theta \notin \Theta^{\dagger}(s,\pi)$, then by definition $\widetilde{D}^0_\theta(s,\pi) \subseteq \cup_{\theta'\neq\theta}\widetilde{D}_{\theta'}(s,\pi)$. And if $\alpha \in D^0_\theta(s,\pi)$, then $\alpha \in MBR(\Theta, s) \subseteq \Delta(BR(\Theta, s))$ and $u_1(\theta,s,\alpha) = u_1(\theta,\pi)$, so $\alpha \in \widetilde{D}^0_\theta(s,\pi)$. Since $\widetilde{D}^0_\theta(s,\pi) \subseteq \cup_{\theta'\neq\theta}\widetilde{D}_{\theta'}(s,\pi)$, there is some $\theta' \neq \theta$ such that $u_1(\theta', s, \alpha) >$

$u_1(\theta', \pi)$, or equivalently $\alpha \in D_{\theta'}(s, \pi)$. As α is an arbitrary element of $D^0_\theta(s, \pi)$, we conclude that $D^0_\theta(s, \pi) \subseteq \cup_{\theta' \neq \theta} D_{\theta'}(s, \pi)$, so $\theta \notin \Theta^\ddagger(s, \pi)$. ■

Lemma 7. *If $\pi = (\pi_1, \pi_2) \in \Pi_1 \times \Pi_2$ is a PBE-H that satisfies NWBR, then, for every $s \in S$, either*

1. *$\Theta^\ddagger(s, \pi) \neq \emptyset$, or*
2. *$u_1(\theta, s, a) < u_1(\theta, \pi)$ for all $\theta \in \Theta$, $a \in BR(\Theta, s)$.*

Proof. Let $\pi = (\pi_1, \pi_2) \in \Pi_1 \times \Pi_2$ be a PBE-H that satisfies NWBR. Fix $s \in S$ and suppose that $\Theta^\ddagger(s, \pi) = \emptyset$. Let $\mathcal{A}_- = \{\alpha \in MBR(\Theta, s) : u_1(\theta, s, \alpha) < u_1(\theta, \pi)\ \forall \theta \in \Theta\}$ be the set of receiver mixed best responses that make playing s strictly worse for every type than their outcome under π. Similarly, let $\mathcal{A}_+ = \{\alpha \in MBR(\Theta, s) : \exists \theta \in \Theta \text{ s.t. } u_1(\theta, s, \alpha) > u_1(\theta, \pi)\}$ be the set of receiver mixed best responses that make some type strictly better off by playing s than receiving their outcome under π. $\mathcal{A}_-$ and $\mathcal{A}_+$ are disjoint open subsets of $MBR(\Theta, s)$, and $\mathcal{A}_- \cup \mathcal{A}_+ = MBR(\Theta, s)$ since $\Theta^\ddagger(s, \pi) = \emptyset$. As $MBR(\Theta, s)$ is connected, either $\mathcal{A}_- = MBR(\Theta, s)$ or $\mathcal{A}_+ = MBR(\Theta, s)$. $\mathcal{A}_+ = MBR(\Theta, s)$ is not possible when π is a PBE-H that satisfies NWBR since then, for every $\alpha \in MBR(\widehat{\Theta}(s, \pi), s)$, there is some θ such that $u_1(\theta, s, \alpha) > u_1(\theta, \pi)$. Therefore, $\mathcal{A}_- = MBR(\Theta, s)$, which gives $u_1(\theta, s, a) < u_1(\theta, \pi)$ for all $a \in BR(\Theta, s)$. ■

Proof of Proposition 2. Let $\pi = (\pi_1, \pi_2) \in \Pi_1 \times \Pi_2$ be a PBE-H that satisfies NWBR. We will show that, for every off-path s, there is a justified response $\alpha_s \in \Delta(BR(\overline{\Theta}(s, \pi), s))$ such that $u_1(\theta, s, \alpha) \leq u_1(\theta, \pi)$ for all $\theta \in \Theta$. This means that the profile $\widetilde{\pi} = (\pi_1, \widetilde{\pi}_2)$ in which $\widetilde{\pi}_2$ coincides with π_2 for all on-path s and dictates α_s for all off-path s is a justified communication equilibrium that is path-equivalent to π.

If $\Theta^\ddagger(s, \pi) \neq \emptyset$, then by Lemma 6, $\Theta^\ddagger(s, \pi) \subseteq \Theta^\dagger(s, \pi)$, so $\widehat{\Theta}(s, \pi) \subseteq \overline{\Theta}(s, \pi)$. If instead $\Theta^\ddagger(s, \pi) = \emptyset$, then by Lemma 7, $u_1(\theta, s, a) < u_1(\theta, \pi)$ for all $\theta \in \Theta$, $a \in BR(\Theta, s)$, so $\Theta^\dagger(s, \pi) = \emptyset$ and $\widehat{\Theta}(s, \pi) = \overline{\Theta}(s, \pi) = \Theta$. Therefore, $MBR(\widehat{\Theta}(s, \pi), s) \subseteq \Delta(BR(\overline{\Theta}(s, \pi), s))$, and, since π satisfies NWBR, there exists some $\alpha \in MBR(\widehat{\Theta}(s, \pi), s) \subseteq \Delta(BR(\overline{\Theta}(s, \pi), s))$ such that $u_1(\theta, s, \alpha) \leq u_1(\theta, \pi)$ for all $\theta \in \Theta$. ■

A.3 Proof of Lemma 4

Proof. Fix $s \in S$. We show that, for all $\theta \in \Theta$, $\widetilde{D}^0_\theta(s,\pi) \not\subseteq \cup_{\theta' \neq \theta} \widetilde{D}_{\theta'}(s,\pi)$ if and only if $D^0_\theta(s,\pi) \not\subseteq \cup_{\theta' \neq \theta} D_{\theta'}(s,\pi)$. This means that $\Theta^\dagger(s,\pi) = \Theta^\ddagger(s,\pi)$, which implies that $\overline{\Theta}(s,\pi) = \widehat{\Theta}(s,\pi)$.

Suppose that $D^0_\theta(s,\pi) \not\subseteq \cup_{\theta' \neq \theta} D_{\theta'}(s,\pi)$. Then there is some $\alpha \in MBR(\Theta, s)$ such that $u_1(\theta, s, \alpha) = u_1(\theta, \pi)$ and $u_1(\theta', s, \alpha) \leq u_1(\theta', \pi)$ for all $\theta' \neq \theta$. Since $\alpha \in \Delta(BR(\Theta, s))$, this immediately implies that $\widetilde{D}^0_\theta(s,\pi) \not\subseteq \cup_{\theta' \neq \theta} \widetilde{D}_{\theta'}(s,\pi)$.

Suppose that $\widetilde{D}^0_\theta(s,\pi) \not\subseteq \cup_{\theta' \neq \theta} \widetilde{D}_{\theta'}(s,\pi)$. Then there is some $\alpha \in \Delta(BR(\Theta, s))$ such that $u_1(\theta, s, \alpha) = u_1(\theta, \pi)$ and $u_1(\theta', s, \alpha) \leq u_1(\theta', \pi)$ for all $\theta' \neq \theta$. By continuity, there exists some $\alpha' \in MBR(\Theta, s)$ such that $u_1(\theta, s, \alpha') = u_1(\theta, s, \alpha)$. Because the game is co-monotonic, $u_1(\theta', s, \alpha') = u_1(\theta', s, \alpha)$ holds for all $\theta' \neq \theta$ as well. Thus, $u_1(\theta, s, \alpha') = u_1(\theta, \pi)$ and $u_1(\theta', s, \alpha') \leq u_1(\theta', \pi)$ for all $\theta' \neq \theta$. This gives $\widetilde{D}^0_\theta(s,\pi) \not\subseteq \cup_{\theta' \neq \theta} \widetilde{D}_{\theta'}(s,\pi)$. ■

B A Sufficient Condition for Stability

Definition 7. *A signaling game is* **strictly monotonic** *if, for all $\theta, \theta' \in \Theta$, $s \in S$, and $\alpha, \alpha' \in MBR(\Theta, s)$,*

1. $u_1(\theta, s, \alpha) \geq u_1(\theta, s, \alpha')$ *if and only if* $u_1(\theta', s, \alpha) \geq u_1(\theta', s, \alpha')$*, and*
2. $u_1(\theta, s, \alpha) = u_1(\theta, s, \alpha')$ *implies* $\alpha = \alpha'$*.*

Here the first condition is exactly the monotonicity of Cho and Sobel [1990]. The second condition requires that the sender preference is a strict order on $MBR(\Theta, s)$.

In general signaling games with cheap talk, different sender types may wish to induce different receiver actions while playing the same signal. However, in strictly monotonic games, sender types using the same signal always desire to induce the same receiver mixed best response. This makes it easier to give sufficient conditions for strategy profiles to be stable.

For a given strategy profile π, let X^{on} be the set of on-path signal-message pairs, let $p(\theta|s,m)$ denote the conditional probability of θ given $(s,m) \in X^{on}$, let S^{on} be the set of on-path signals, and let S^{off} be the set of off-path signals.

Definition 8. *The JCE* $\pi = (\pi_1, \pi_2) \in \Pi_1 \times \Pi_2$ *is* **uniformly justified** *if*

1. *For all* $\theta \in \Theta$*, there is some* $s_\theta \in S$ *such that* $\max_{m \in M} u_1(\theta, s_\theta, \pi_2(\cdot|s_\theta, m)) > \max_{s \neq s_\theta, m \in M} u_1(\theta, s, \pi_2(\cdot|s, m))$*,*
2. *For every* $(s,m) \in X^{on}$*, there is some* $a_x \in A$ *such that* $u_2(p(\cdot|s,m), s, a_x) > \max_{a \neq a_x} u_2(p(\cdot|s,m), s, a)$*,*
3. *For all* $s \in S^{off}$*,* $u_1(\theta, s, a) < u_1(\theta, \pi)$ *for all* $\theta \in \Theta$*,* $a \in BR(\overline{\Theta}(s,\pi), s)$*.*

Condition 1 says that every sender type plays exactly one signal and that they have strict incentives to do so. Condition 2 says that the receiver has a strictly optimal action in response to every on-path signal-message pair. Condition 3 says that all sender types are deterred from playing any off-path signal for any justified receiver response.

Proposition 5. *If* π *is uniformly justified JCE in a strictly monotonic signaling game, it induces the same distribution over* $\Theta \times S \times A$ *as a stable profile for all* g_1, g_2*.*

OA.2 contains the proof of Proposition 5. Note that there is a receiver behavior strategy which strictly incentivizes every type to play their corresponding signal in π, and, when each sender type does so, leads to the same distribution over $\Theta \times S \times A$ as π. The proof works by modifying aggregate receiver response to match this behavior strategy with high probability unless each type plays their corresponding signal in π with sufficiently high probability. Lemma 5 implies the aggregate sender play given by the fixed points of the modified aggregate response mapping is optimal in the iterated limit. The modification to the receiver aggregate response thus ensures that the limit aggregate sender strategy uses the signals prescribed by π with high probability.[20]

[20] Fudenberg and Levine [2006] and Fudenberg and He [2020a] prove the stability of certain strategy profiles by considering priors that assign high probability to a neighborhood of the target profile. Our approach of modifying the aggregate response mapping lets us prove stability for a broad class of priors.

Additionally, by strict monotonicity and the optimality of the aggregate sender play, the receiver response to any on-path signal-message pair must only depend on the signal. As receiver incentives are strict in π, this means that the receiver response to any on-path signal-message pair matches the response in π. We show that this, along with the fact that π is uniformly justified, implies that the limit aggregate sender play signals according to π with probability 1. This means that the constraint from the modification of the receiver best response is not binding, so the fixed points of the modified response mapping are valid steady state profiles, and in the limit induce the same distribution over $\Theta \times S \times A$ as π.

C Omitted Details about Other Refinements

C.1 Stability Does Not Imply Sequential Equilibrium

Proposition 6. *For the game in Table 1, there are stable profiles where both types play Out with probability* 1*.*

Proposition 6 shows that in the game in Example 1, there are stable profiles where both types play *Out* with probability 1, which can never happen in a sequential equilibrium. The proof of Proposition 6, which is presented in OA.3, takes the receiver prior to be such that, if a receiver has encountered past play of (In, m) and all such plays have been by senders with the same type θ, then the receiver will respond to the next instance of (In, m) with $BR(\theta, In)$. (To make the receiver initially trusting, the prior is also chosen so that a receiver responds with $BR(\theta, In)$ when they first encounter a sender who plays (In, m_θ).) We show that, holding γ_2 fixed, there are profiles in the iterated limit where $\gamma_1 \to 1$ then $\delta \to 1$ in which, for each message m, at most one sender type plays (In, m) with positive aggregate probability. Then any $\gamma_2 \to 1$ limit of such profiles must result in both types playing In with aggregate probability 0. Otherwise, there would be some message m where only type θ plays (In, m) with positive aggregate probability in the limit. This would lead to the receiver population responding to (In, m) with $BR(\theta, In)$ with aggregate probability 1, which would contradict the optimality of aggregate sender play in the limit since $u_1(\theta, In, BR(\theta, In)) < 0$.

C.2 Intuitive Criterion

Let $E(s, \pi) = \{\theta \in \Theta : \max_{a \in BR(\Theta, s)} u_1(\theta, s, a) \geq u_1(\theta, \pi)\}$. This is the set of sender types for whom s is not equilibrium dominated by profile π in the sense of Cho and Kreps [1987].

Definition 9 (Cho and Kreps [1987]). *Strategy profile π* **passes the Intuitive Criterion Test** *if, for every $s \in S$ and $\theta \in E(s, \pi)$, $\min_{a \in BR(E(s,\pi),s)} u_1(\theta, s, a) \leq u_1(\theta, \pi)$.*

Proposition 7. *If $\pi = (\pi_1, \pi_2) \in \Pi_1 \times \Pi_2$ is a justified communication equilibrium, then π is a PBE-H that passes the Intuitive Criterion Test.*

The key step of the proof is to show that in any PBE-H π where $E(s, \pi) \neq \emptyset$ for $s \in S$, we have $\overline{\Theta}(s, \pi) \subseteq E(s, \pi)$. This is a consequence of the following lemma.

Lemma 8. *If $\pi = (\pi_1, \pi_2) \in \Pi_1 \times \Pi_2$ is a JCE, then, for every $s \in S$, either*

1. *$\Theta^\dagger(s, \pi) \neq \emptyset$, or*
2. *$u_1(\theta, s, a) < u_1(\theta, \pi)$ for all $\theta \in \Theta$, $a \in BR(\Theta, s)$.*

The proof of Lemma 8 is analogous to that of Lemma 7 and is given in OA.4.

Proof of Proposition 7. Fix $s \in S$ and suppose that $E(s, \pi) \neq \emptyset$. This means that there is some θ and $\alpha \in BR(\Theta, s)$ such that $u_1(\theta, s, a) \geq u_1(\theta, \pi)$. By Lemma 8, $\Theta^\dagger(s, \pi) \neq \emptyset$, so $\overline{\Theta}(s, \pi) = \Theta^\dagger(s, \pi)$. Moreover, $\overline{\Theta}(s, \pi) \subseteq E(s, \pi)$ since every type θ for which $\max_{a \in BR(\Theta, s)} u_1(\theta, s, a) < u_1(\theta, \pi)$ satisfies $\widetilde{D}^0_\theta(s, \pi) = \emptyset \subseteq \widetilde{D}_{\theta'}$ for every $\theta' \in E(s, \pi)$. Thus, any $\alpha \in \Delta(BR(\overline{\Theta}(s, \pi), s))$ is also an element of $\Delta(BR(E(s, \pi), s))$. Since there is some such α such that $u_1(\theta, s, \alpha) \leq u_1(\theta, \pi)$ for all $\theta \in \Theta$, we conclude that $\min_{a \in BR(E(s,\pi),s)} u_1(\theta, s, a) \leq u_1(\theta, \pi)$ for all $\theta \in E(s, \pi)$. ■

C.3 D1 and co-D1

For every $s \in S$ and $\pi \in \Pi_1 \times \Pi_2$, let $\Theta^{\ddagger,D1}(s,\pi) = \{\theta \in \Theta : \forall \theta' \neq \theta,\ D_\theta(s,\pi) \cup D^0_\theta(s,\pi) \not\subseteq D_{\theta'}(s,\pi)\}$ be the set of types θ where, for every $\theta' \neq \theta$, there is some receiver mixed best response $\alpha \in MBR(\Theta, s)$ that makes θ weakly prefer (s,α) to their equilibrium outcome, while θ' weakly prefers their equilibrium outcome to (s,α). Let $\widehat{\Theta}^{D1}(s,\pi) \subseteq \Theta$ be the set of types given by

$$\widehat{\Theta}^{D1}(s,\pi) = \begin{cases} \Theta^{\ddagger,D1}(s,\pi) & \text{if } \Theta^{\ddagger,D1}(s,\pi) \neq \emptyset \\ \Theta & \text{if } \Theta^{\ddagger,D1}(s,\pi) = \emptyset \end{cases}.$$

Definition 10 (Banks and Sobel [1987])**.** *A PBE π* **satisfies D1** *if for every $s \in S$, there is an $\alpha \in MBR(\widehat{\Theta}^{D1}(s,\pi), s))$ such that $u_1(\theta, s, \alpha) \leq u_1(\theta, \pi)$ for all $\theta \in \Theta$.*

For every $s \in S$ and $\pi \in \Pi_1 \times \Pi_2$, let $\Theta^{\dagger,D1}(s,\pi) = \{\theta \in \Theta : \forall \theta' \neq \theta,\ \widetilde{D}_\theta(s,\pi) \cup \widetilde{D}^0_\theta(s,\pi) \not\subseteq \widetilde{D}_{\theta'}(s,\pi)\}$ be the set of types θ where, for every $\theta' \neq \theta$, there is some mixed receiver action $\alpha \in \Delta(BR(\Theta, s))$ that makes θ weakly prefer between (s,α) to their equilibrium outcome and θ' weakly prefers their equilibrium outcome to (s,α). Let $\overline{\Theta}^{D1}(s,\pi) \subseteq \Theta$ be the set of types given by

$$\overline{\Theta}^{D1}(s,\pi) = \begin{cases} \Theta^{\dagger,D1}(s,\pi) & \text{if } \Theta^{\dagger,D1}(s,\pi) \neq \emptyset \\ \Theta & \text{if } \Theta^{\dagger,D1}(s,\pi) = \emptyset \end{cases}.$$

Definition 11. *A PBE-H π is* **co-D1** *if for every $s \in S$, there is an $\alpha \in \Delta(BR(\overline{\Theta}^{D1}(s,\pi), s))$ such that $u_1(\theta, s, \alpha) \leq u_1(\theta, \pi)$ for all $\theta \in \Theta$.*

Proposition 8. *If $\pi = (\pi_1, \pi_2) \in \Pi_1 \times \Pi_2$ is a justified communication equilibrium, then π is a PBE-H that is co-D1.*[21]

[21] *Divinity* and *universal divinity* (Banks and Sobel [1987]) resemble D1 and co-D1 but use iterative procedures. As we show in OA.7.3, we can capture a similar iterated procedure with an additional assumption about the receiver policy.

Co-D1 is a more permissive refinement than JCE because it strikes fewer types. The following proof shows that $\overline{\Theta}(s,\pi) \subseteq \overline{\Theta}^{D1}(s,\pi)$ for all s; Example OA 2 shows that the inclusion is sometimes strict.

Proof. Fix $s \in S$. We will argue that $\overline{\Theta}(s,\pi) \subseteq \overline{\Theta}^{D1}(s,\pi)$. This, along with the justified response criterion of JCE and the fact that every JCE is a PBE-H, implies that π is co-D1.

If $\Theta^{\dagger}(s,\pi) \neq \emptyset$, then $\overline{\Theta}(s,\pi) = \Theta^{\dagger}(s,\pi)$. Let θ be a type such that $\theta \notin \overline{\Theta}^{D1}(s,\pi)$. Then there is some type $\theta' \neq \theta$ such that $D_{\theta}(s,\pi) \cup D^0_{\theta}(s,\pi) \subseteq D_{\theta'}(s,\pi)$. This implies that $\theta' \notin \overline{\Theta}(s,\pi)$, so $\overline{\Theta}(s,\pi) \subseteq \overline{\Theta}^{D1}(s,\pi)$ follows.

If $\Theta^{\dagger}(s,\pi) = \emptyset$, by Lemma 8, $u_1(\theta,s,a) < u_1(\theta,\pi)$ for all $a \in BR(\Theta,s)$. Thus $\Theta^{\dagger,D1}(s,\pi) = \emptyset$ as $D_{\theta}(s,\pi) \cup D^0_{\theta}(s,\pi) \subseteq D_{\theta'}(s,\pi)$ for all $\theta,\theta' \in \Theta$. Thus, $\overline{\Theta}(s,\pi) = \Theta = \overline{\Theta}^{D1}(s,\pi)$. ∎

C.4 Rationality-Compatible Equilibrium

We write $\Pi_2^{\bullet} = \times_{(s,m)\in S\times M}\Delta(BR(\Theta,s))$ for the set of receiver strategies that assign probability 0 to conditionally dominated responses.

Definition 12 (Fudenberg and He [2020a]). *Signal $s \in S$ is* **more rationally-compatible** *with θ' than θ'', written as $\theta' \succsim_s \theta''$,*[22] *if, for every $\pi_2 \in \Pi_2^{\bullet}$,*

$$\max_{m\in M} u_1(\theta'',s,\pi_2(\cdot|s,m)) \geq \max_{s'\neq s, m\in M} u_1(\theta'',s',\pi_2(\cdot|s',m)) \text{ implies that}$$

$$\max_{m\in M} u_1(\theta',s,\pi_2(\cdot|s,m)) > \max_{s'\neq s, m\in M} u_1(\theta',s',\pi_2(\cdot|s',m)).$$

In words, this says that type θ' is more rationally-compatible with signal s than is θ'' if any undominated receiver strategy that makes θ'' willing to play s makes θ' strictly prefer to play it. Let $P_{\theta'\rhd\theta''} = \{p \in \Delta(\Theta) : \lambda(\theta'')p(\theta') \geq \lambda(\theta')p(\theta'')\}$ be the set of probability distributions over sender type where the odds ratio of θ' to θ'' exceed their

[22]Note that this criterion does not depend on the message space in the game: If $s \in S$ is more rationally-compatible with θ' than θ'' with message space M, then it is also more rationally compatible with θ' than θ'' under any other message space.

odds ratio under the prior distribution. For $s \in S$ and $\pi \in \Pi_1 \times \Pi_2$, let $\overline{P}(s,\pi) \subseteq \Delta(\Theta)$ be the set of beliefs over the sender type given by

$$\overline{P}(s,\pi) = \begin{cases} \Delta(E(s,\pi)) \cap \left(\cap_{(\theta',\theta'') \text{ s.t. } \theta' \succsim_s \theta''} P_{\theta' \triangleright \theta''}\right) & \text{if } E(s,\pi) \neq \emptyset \\ \Delta(\Theta) & \text{if } E(s,\pi) = \emptyset \end{cases}.$$

Definition 13 (Fudenberg and He [2020a])**.** *Strategy profile π is a* **rationality-compatible equilibrium (RCE)** *if it is a PBE-H where, for every $s \in S$, there is an $\alpha \in \Delta(BR(\overline{P}(s,\pi),s))$ such that $u_1(\theta,s,\alpha) \leq u_1(\theta,\pi)$ for all $\theta \in \Theta$.*

This definition requires that the receiver's posterior likelihood ratio for types θ' and θ'' dominates the prior likelihood ratio whenever $\theta' \succsim_s \theta''$. It also requires that the posterior assigns probability 0 to equilibrium-dominated types.

Proposition 9. *If $\pi = (\pi_1, \pi_2) \in \Pi_1 \times \Pi_2$ is a justified communication equilibrium, then π is an RCE.*[23]

Intuitively, any response that makes a less compatible type indifferent about playing s makes more compatible types strictly prefer to play it.

Proof. Fix $s \in S$. We will argue that $\Delta(\overline{\Theta}(s,\pi)) \subseteq \overline{P}(s,\pi)$. Thus any $\alpha \in \Delta(BR(\overline{\Theta}(s,\pi),s))$ also belongs to $\Delta(BR(\overline{P}(s,\pi),s))$. Consequently, the justified response criterion of JCE along with the fact that every JCE is a PBE-H implies that π is an RCE.

Since $\Delta(\overline{\Theta}(s,\pi)) \subseteq \Delta(\Theta) = \overline{P}(s,\pi)$ when $E(s,\pi) = \emptyset$, we need only handle the case where $E(s,\pi) \neq \emptyset$. In this case by Lemma 8, $\overline{\Theta}(s,\pi) = \Theta^{\dagger}(s,\pi)$ and $\Delta(\overline{\Theta}(s,\pi)) \subseteq \Delta(E(s,\pi))$. Suppose that θ' and θ'' are two types such that $\theta' \succsim_s \theta''$. Then Definition 13 implies that $\widetilde{D}^0_{\theta''}(s,\pi) \subseteq \widetilde{D}_{\theta'}(s,\pi)$, so $\theta'' \notin \Theta^{\dagger}(s,\pi)$. As a result, $\Delta(\overline{\Theta}(s,\pi)) = \Delta(\Theta^{\dagger}(s,\pi)) \subseteq \cap_{(\theta',\theta'') \text{ s.t. } \theta' \succsim_s \theta''} P_{\theta' \triangleright \theta''}$. We conclude $\Delta(\overline{\Theta}(s,\pi)) \subseteq \Delta(E(s,\pi)) \cap (\cap_{(\theta',\theta'') \text{ s.t. } \theta' \succsim_s \theta''} P_{\theta' \triangleright \theta''}) = \overline{P}(s,\pi)$. ■

[23] The partially pooling equilibrium in Example 2 that is ruled out by JCE can be shown to be an RCE, so JCE is strictly more demanding.

D Additional Examples

D.1 Stability without Initially Trusting Receivers

Example OA 1 in OA.6.1 shows that when receiver play is not initially trusting, stable profiles may not even satisfy the Intuitive Criterion Test, let alone the stronger requirements of JCE. In the example, there are two types and two actions, a_1, a_2. One of the types never wants to play In; the other wants to play In when the receiver responds with a_1. Moreover, a_1 is the unique best response for the receiver when this type plays In, so all types playing Out cannot pass the Intuitive Criterion Test. However, if the receiver prior is such that they play a_2 in response to a first encounter of (In, m) for each message m, then all types playing Out is a stable outcome. The reason for this is that, if the aggregate receiver response to every (In, m) is to play a_2 with high probability, then the aggregate probability that any sender type plays (In, m) converges to 0 in the iterated limit where $\gamma_1 \to 1$ then $\delta \to 1$. This drives the aggregate receiver response to play a_2 with probability 1 in response to every (In, m) in the iterated limit where $\gamma_1 \to 1$ then $\delta \to 1$ for fixed γ_2. Thus there are stable profiles in which both types play Out.

D.2 D1 Does Not Imply JCE

Even though D1 is typically thought of as a strong refinement, there are equilibria that satisfy D1 (and so also satisfy co-D1) that are not JCE, as shown by example in OA.6.2. In the example, the sender's type space is $\Theta = \{\theta_1, \theta_2, \theta_3\}$, and the receiver's prior is that all types are equally likely. The sender's signal space is $S = \{In, Out\}$, and the receiver's action space is $A = \{a_1, a_2, a_3\}$. The payoffs to the sender and receiver are given in the table below.

θ_1	a_1	a_2	a_3
In	$4, 1$	$-1, 0$	$-1, -1$
Out	$0, 0$	$0, 0$	$0, 0$

θ_2	a_1	a_2	a_3
In	$-1, 0$	$4, 1$	$-1, -1$
Out	$0, 0$	$0, 0$	$0, 0$

θ_3	a_1	a_2	a_3
In	$1, 0$	$1, 0$	$-1, 4$
Out	$0, 0$	$0, 0$	$0, 0$

None of the types want to play In when the receiver responds with a_3, but each type would play In under some other responses. Moreover, if the receiver plays a_3 with probability 0, type θ_3 strictly prefers to play In. For the other two types, there are some mixtures over a_1 and a_2 at which In is strictly preferred to Out and others where Out is strictly preferred to In. This means that θ_3 cannot be eliminated by D1, and because the receiver wants to play a_3 versus θ_3, D1 allows equilibria in which every type plays Out. However, the payoffs of the sender types are such that, whenever the receiver response makes θ_3 indifferent between In and Out, one of the other types strictly prefers to play In. This means that θ_3 is not a justified type, so no equilibrium in which every type plays Out is a JCE.

D.3 Stability Does Not Imply D1

Example OA 3 in OA.6.3 shows that there can be stable profiles that are sequential equilibria but not D1. In the example, there are two types and three actions, a_1, a_2, and a_3. Type θ_1 strictly prefers to play In when the receiver plays either a_1 or a_3 and strictly prefers to play Out when the receiver plays a_2. Type θ_2 strictly prefers to play In when the receiver plays a_1, is indifferent when the receiver plays a_3, and strictly prefers to play Out when the receiver plays a_2. Moreover, the receiver's mixed best responses to In are all the mixed actions which do not put positive probability on both a_1 and a_2. D1 then rules out the equilibrium in which all types play Out since D1 requires the receiver's posterior after In to concentrate on θ_1, which leads the receiver to play a_1.

However, there are receiver priors with stable profiles in which all types play Out. A stable profile cannot have type θ_2 play In with probability 1, as otherwise the receiver would respond to some on-path (In, m) with a_2, which would strictly deter both types from playing (In, m). Moreover, if the aggregate receiver response to each (In, m) plays a_3 with probability less than 1/10, then whenever type θ_1 weakly prefers to play some (In, m) than Out, θ_2 strictly prefers to play In. We show that there are receiver priors compatible with steady state profiles in which the receiver's aggregate behavior strategy satisfies this condition, so all types playing Out is a stable outcome.

References

J. S. Banks and J. Sobel. Equilibrium selection in signaling games. *Econometrica*, 55: 647–661, 1987.

P. Battigalli. Comportamento razionale ed equilibrio nei giochi e nelle situazioni sociali. Undergraduate Thesis, 1987.

P. Battigalli, S. Cerreia-Vioglio, F. Maccheroni, and M. Marinacci. Self-confirming equilibrium and model uncertainty. *American Economic Review*, 105:646–677, 2015.

P. Battigalli, A. Francetich, G. Lanzani, and M. Marinacci. Learning and self-confirming long-run biases. *Journal of Economic Theory*, 183:740–785, 2019.

K. Binmore and L. Samuelson. Evolutionary drift and equilibrium selection. *Review of Economic Studies*, 66:363–393, 1999.

A. Blume and J. Sobel. Communication-proof equilibria in cheap-talk games. *Journal of Economic Theory*, 65:359–382, 1995.

Y. Chen, N. Kartik, and J. Sobel. Selecting cheap-talk equilibria. *Econometrica*, 76: 117–136, 2008.

I-K. Cho and D. M. Kreps. Signaling games and stable equilibria. *The Quarterly Journal of Economics*, 102:179–221, 1987.

I-K. Cho and J. Sobel. Strategic stability and uniqueness in signaling games. *Journal of Economic Theory*, 50:381–413, 1990.

D. Clark. Robust neologism proofness in signaling games. Working Paper, 2020.

E. Dekel, D. Fudenberg, and D. K. Levine. Payoff information and self-confirming equilibrium. *Journal of Economic Theory*, 89:165–185, 1999.

P. Diaconis and D. Freedman. On the uniform consistency of bayes estimates for multinomial probabilities. *The Annals of Statistics*, 18:1317–1327, 1990.

I. Esponda. Rationalizable conjectural equilibrium: A framework for robust predictions. *Theoretical Economics*, 8:467–501, 2013.

J. Farrell. Meaning and credibility in cheap-talk games. *Games and Economic Behavior*, 5:514–531, 1993.

D. Fudenberg and K. He. Learning and type compatibility in signaling games. *Econometrica*, 86:1215–1255, 2018.

D. Fudenberg and K. He. Payoff information and learning in signaling games. *Games and Economic Behavior*, 120:96–120, 2020a.

D. Fudenberg and K. He. Player-compatible learning and player-compatible equilibrium. Working Paper, 2020b.

D. Fudenberg and Y. Kamada. Rationalizable partition-confirmed equilibrium. *Theoretical Economics*, 10:775–806, 2015.

D. Fudenberg and Y. Kamada. Rationalizable partition-confirmed equilibrium with heterogeneous beliefs. *Games and Economic Behavior*, 109:364–381, 2018.

D. Fudenberg and D. Kreps. A theory of learning, experimentation, and equilibrium in games. Working Paper, 1988.

D. Fudenberg and D. Kreps. Learning in extensive-form games, ii: Experimentation and nash equilibrium. Working Paper, 1994.

D. Fudenberg and D. K. Levine. Steady state learning and nash equilibrium. *Econometrica*, 61:547–573, 1993.

D. Fudenberg and D. K. Levine. Superstition and rational learning. *American Economic Review*, 96:630–651, 2006.

D. Fudenberg and J. Tirole. Perfect bayesian equilibrium and sequential equilibrium. *Journal of Economic Theory*, 53:236–260, 1991a.

D. Fudenberg and J. Tirole. *Game Theory*. The MIT Press, 1991b.

J. C. Gittins. Bandit processes and dynamic allocation indices. *Journal of the Royal Statistical Society, Series B*, 41:148–164, 1979.

M. Goltsman and G. Pavlov. How to talk to multiple audiences. *Games and Economic Behavior*, 72:100–122, 2011.

S. Hart. Evolutionary dynamics and backward induction. *Games and Economic Behavior*, 41:227–264, 2002.

P. Jehiel and D. Samet. Learning to play games in extensive form by valuation. *Journal of Economic Theory*, 124:129–148, 2005.

E. Kalai and E. Lehrer. Rational learning leads to nash equilibrium. *Econometrica*, 61:1019–1045, 1993.

N. Kartik. Strategic communication with lying costs. *Review of Economic Studies*, 76: 1359–1395, 2009.

N. Kartik, M. Ottaviani, and F. Squintani. Credulity, lies, and costly talk. *Journal of Economic Theory*, 134:93–116, 2006.

E. Kohlberg and J-F. Mertens. On the strategic stability of equilibria. *Econometrica*, 54:1003–1037, 1986.

E. Lehrer and E. Solan. Learning to play partially-specified equilibrium. Working Paper, 2007.

B. L. Lipman and D. J. Seppi. Robust inference in communication games with partial provability. *Journal of Economic Theory*, 66:370–405, 1995.

S. A. Matthews, M. Okuno-Fujiwara, and A. Postlewaite. Refining cheap talk equilibria. *Journal of Economic Theory*, 55:247–273, 1991.

G. Nöldeke and L. Samuelson. An evolutionary analysis of backward and forward induction. *Games and Economic Behavior*, 5:425–454, 1993.

W. Olszewski. Rich language and the refinements of cheap-talk equilibria. *Journal of Economic Theory*, 128:164–186, 2006.

M. Rabin. Communication between rational agents. *Journal of Economic Theory*, 51: 144–170, 1990.

J. G. Riley. Informational equilibrium. *Econometrica*, 47:331–359, 1979.

A. Rubinstein and A. Wolinsky. Rationalizable conjectural equilibrium: Between nash and rationalizability. *Games and Economic Behavior*, 6:299–311, 1994.

J. Sobel, L. Stole, and I. Zapater. Fixed-equilibrium rationalizability in signaling games. *Journal of Economic Theory*, 52:304–331, 1990.

I. Zapater. Credible proposals in communication games. *Journal of Economic Theory*, 72:173–197, 1997.