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 Kreps, David K. Levine, Harry Pei, Andy Skrypacz, Joel Sobel, and Alex Wolitzky for helpful comments, and NSF grants SES 1643517 and SES 1951056 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 costly signals and cheap-talk messages. For a given signal and strategy profile, a sender type is justified if some conceivable (i.e. undominated) response makes the type weakly prefer to play the signal rather than conform to the strategy profile, and makes all other types weakly prefer to conform. A justified response to a signal 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. As opposed to the equal expected lifetimes of Fudenberg and Levine [1993] and Fudenberg and He [2018, 2020a], relatively long-lived senders fit 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 assumption that the message space is large
enough that, for each signal and subset of sender types, there is a distinct message that claims “I am playing this signal and my type is in this set.” We further assume that receivers are initially trusting, which roughly means that the receivers’ prior leads them to trust such 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 5 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. A given signal can only be a best response for justified types, so receivers are very unlikely to encounter a signal being played by a non-justified type. Initial trust then implies 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. In particular, these speeches are not counterfactual, but are messages that are actually sent, which lets us determine how receivers respond to them. Thus, we sidestep the “Stiglitz critique” (Cho and Kreps [1987], Rabin and Sobel [1996]) of signaling game refinements, which is based on iterated arguments about how players believe their opponent “should” interpret hypothetical deviations.

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 commonly used signaling game refinements, such as the Intuitive Criterion (Cho and Kreps [1987]), D1 (Banks and Sobel [1987]), and NWBR (Cho and Kreps [1987]),\footnote{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].} and the intuitions behind these concepts have a connection to non-equilibrium learning. However, none of the traditional equilibrium refinements
is a necessary condition for stability in our learning model. Indeed, as we show in Example 1, the stable outcomes of our 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 to a single belief. 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 the least-cost separating equilibria or equilibria that approximate it. While existing refinements like D1 and NWBR also make this selection, JCE provides the first learning justification for this selection. Moreover, JCE is a strict subset of the rationality-compatible equilibrium (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 \( \Theta \), 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 \( u_1 : \Theta \times S \times A \rightarrow \mathbb{R} \) and the utility function of the receiver is \( u_2 : \Theta \times S \times A \rightarrow \mathbb{R} \). Both players’ utilities depend on the signal and type of the sender and the action of the receiver; neither utility depends on the message of the sender. The full-support

\[ ^2 \text{Moreover, as far as we know they have not been shown to be necessary in -any- learning model.} \]
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 denote the set of sender behavior strategies by $\Pi_1 = (\Delta(S \times M))^\Theta$, the set of receiver behavior strategies by $\Pi_2 = (\Delta(A))^{S \times M}$, and let $\Pi = \Pi_1 \times \Pi_2$ be the set of strategy profiles. Finally, we write $u_1(\theta, \pi)$ and $u_2(\pi)$ for the expected payoffs from strategy profile $\pi = (\pi_1, \pi_2)$, let $BR(p, s) = \arg \max_{a \in A} u_2(p, s, a)$ denote the best responses for the receiver to signal $s$ under belief $p \in \Delta(\Theta)$, and let $BR(\tilde{\Theta}, s) = \bigcup_{p \in \Delta(\tilde{\Theta})} BR(p, s)$ denote the best responses for the receiver to signal $s$ for some $p$ with support in $\tilde{\Theta}$.

### 2.2 Learning Environment

We consider an overlapping generations learning environment where time is discrete and doubly infinite, $t \in \{\ldots, -2, -1, 0, 1, 2, \ldots\}$, and where, for each $\theta \in \Theta$, there is a continuum of agents of mass $\lambda(\theta)$ in the role of a type $\theta$ 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. In each match, the sender plays a signal $s \in S$ and a message $m \in M$. The receiver observes the chosen $(s, m)$ and responds with an 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 aggregate receiver behavior strategy, while receiver agents have a non-doctrinaire prior $g_2 \in \Delta(\Delta(\Theta \times S \times M))$ over the distribution of sender types, signals, and messages.\(^3\) (To simplify notation, we assume there is a single prior for all agents in a given player role, but all of our results extend to any finite number of priors per role.) 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 $\mathcal{H}_{1,t} = (S \times M \times \Theta)^t$ to be the set of histories that a sender of age $t$ could have observed, and let $\mathcal{H}_1 = \bigcup_{t \in \mathbb{N}} \mathcal{H}_{1,t}$ be the collection of all such histories. 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 $\mathcal{H}_2 = \bigcup_{t \in \mathbb{N}} \mathcal{H}_{2,t}$ be the collection of all such sequences.

All agents are rational Bayesians who maximize their expected discounted payoff. Because the receivers always observe the type of the sender at the end of each match, neither their continuation probability nor their discount factor impacts their play, and they simply choose an action that maximizes their expected payoff in the current match. Receivers use a policy $y : \mathcal{H}_2 \rightarrow Y$ that makes such choices. In contrast, senders’ observations do depend on their play, so they have an incentive to “experiment” with various signal-message pairs that have the potential to lead to an increase in payoff. The size of the senders’ experimentation incentive depends on their continuation probability $\gamma_1$ and their discount factor $\delta \in [0, 1)$. Type $\theta$ senders use an optimal policy $x^{\delta, \gamma_1}_\theta : \mathcal{H}_1 \rightarrow X$. We will focus on the case where both $\delta$ and $\gamma_1$ are near 1, so the senders have maximal incentives to experiment. The force behind our equilibrium refinement is that different sender types will experiment in different ways.

\(^3\)Here “non-doctrinaire” means “has a continuous density function that is strictly positive on the interior of the probability simplex.” Because we allow correlated beliefs, it would be equivalent to define the receivers’ beliefs as elements of $\Delta(\Theta \times \Pi_1)$. 
2.3 Steady States and Aggregate Play

At every period $t$, the state of the system, denoted $\mu_t = (\mu_{1,t}, \mu_{2,t}) \in \Delta(H_1)^{\Theta} \times \Delta(H_2)$, gives the shares of agents in a given player role with the various possible histories. Given $\mu_t$, the profile $x^{\delta, \gamma_1, \gamma_2} = \{x_{\theta}^{\delta, \gamma_1}\}_{\theta \in \Theta}$ of sender policies induces a sender behavior strategy $\sigma_1^{\delta, \gamma_1}((\mu_{1,t}, t)) \in \Pi_1$ that we call the aggregate sender play. Similarly, the receiver policy $y$ induces a receiver behavior strategy $\sigma_2(\mu_{2,t}) \in \Pi_2$ that we call the aggregate receiver play. We call $\sigma_1^{\delta, \gamma_1}(\mu_t) = (\sigma_1^{\delta, \gamma_1}(\mu_{1,t}), \sigma_2(\mu_{2,t})) \in \Pi_1 \times \Pi_2$ the aggregate strategy profile. (Online Appendix Section OA.1 gives formal definitions of the mappings $\sigma_1^{\delta, \gamma_1}$, $\sigma_2$:

A policy profile generates an update rule $f^{\delta, \gamma_1, \gamma_2} : \Delta(H_1)^{\Theta} \times \Delta(H_2) \rightarrow \Delta(H_1)^{\Theta} \times \Delta(H_2)$, taking the state in period $t$ to the state in period $t + 1$, a mapping $R^{\delta, \gamma_1} : \Pi_2 \rightarrow \Pi_1$ that describes the limit of the aggregate play of the senders as $t \rightarrow \infty$ when the aggregate receiver play is fixed at $\pi_2$, and a mapping $R^{\gamma_2} : \Pi_1 \rightarrow \Pi_2$ that describes the limit of the aggregate receiver play as $t \rightarrow \infty$ when the aggregate sender play is fixed at $\pi_1$. We refer to the mapping $R^{\delta, \gamma_1, \gamma_2}(\pi) \equiv (R_1^{\delta, \gamma_1}(\pi_2), R_2^{\gamma_2}(\pi_1))$ as the aggregate response mapping. Straightforward arguments show that these mappings are continuous.

We study this system’s steady states, those $\mu$ satisfying $f^{\delta, \gamma_1, \gamma_2}(\mu) = \mu$, 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 $g = (g_1, g_2)$, $\delta$, and $\gamma_1, \gamma_2$.

The proof follows standard lines and is omitted.

We study the limit of steady state play when $\gamma_1$ and $\gamma_2$ tend to 1, so senders and receivers can acquire enough observations to outweigh their prior. We also assume that $\delta$ goes to 1 to ensure that the senders experiment enough to rule out limits that are not Nash equilibria. In addition, we assume that the receivers play much
less often than the senders. These assumptions fit settings where the senders are institutions who both have an incentive to experiment and, over time, interact with a large number of individuals in the role of the receivers; one example is firms signaling their knowledge about their productivity, future growth, etc. to potential employees via offers of incentive pay. 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.

In particular, we consider the limit of steady state aggregate play in the iterated limit \( \lim_{\gamma_2 \to 1} \lim_{\delta \to 1} \lim_{\gamma_1 \to 1} \Pi^*(g, \delta, \gamma_1, \gamma_2) \). That is, we focus on all strategy profiles that are limits of steady states corresponding to some sequence of parameters \( \delta, \gamma_1, \) and \( \gamma_2 \) satisfying this iterated limit. We will call these the stable profiles.\(^4\) 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 limit, where all agents have the same expected lifetimes, and they first become arbitrarily long-lived and then arbitrarily patient, so that near the limit very few players have incentives to experiment. For the same reason, we require senders to become long-lived (\( \gamma_1 \to 1 \)) before they become patient (\( \delta \to 1 \)).\(^5\) 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.

\(^4\)Formally, strategy profile \( \pi \) is stable if there is a sequence \( \{\gamma_{j,k}\}_{j \in \mathbb{N}} \to 1 \), sequences \( \{\delta_{j,k}\}_{j,k \in \mathbb{N}} \) with \( \lim_{k \to \infty} \delta_{j,k} = 1 \) 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 \( \pi_{j,k,l} \in \Pi^*(g, \delta_{j,k}, \gamma_{1,j,k,l}, \gamma_{2,j}) \).

\(^5\)We do not know what happens when \( \delta \to 1 \) before \( \gamma_1 \to 1 \). The order with which \( \gamma_2 \) and \( \delta \) go to 1 is not crucial. We specify that \( \delta \) converges to 1 before \( \gamma_2 \) because it affords slightly cleaner results and simpler proofs. Section 6 discusses a more general limit.
2.4 Key Assumptions

Our results about stable profiles use two additional assumptions. First, we assume that the sender message space is sufficiently rich.

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

Assumption 1 implies that for each signal \(s \in S\) and subset of sender types \(\tilde{\Theta} \subseteq \Theta\) there is a message \(m_{s,\tilde{\Theta}} \in M\) that can be interpreted as claiming “I am playing \(s\) and my type is in \(\tilde{\Theta}\).”\(^6\) Our next assumption is that when the sender plays \(s\) and sends the message \(m_{s,\tilde{\Theta}}\), the receiver “trusts” the message provided that they have not previously encountered a sender with any other type \(\theta \not\in \tilde{\Theta}\) playing \(s\) and sending \(m_{s,\tilde{\Theta}}\).

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

Without any assumptions on the receiver’s prior, stability offers little predictive power and allows implausible outcomes, as we show by example in B.1. In the example, there are two sender types, \(\theta_1\) and \(\theta_2\), and two signals, \(In\) and \(Out\). \(Out\) is strictly dominant for \(\theta_2\), and \(\theta_1\) prefers to play \(In\) if the receiver responds to \(In\) with the best response to \(\theta_1\), so the reasonable outcome seems to be one where \(\theta_1\) plays \(In\) and \(\theta_2\) plays \(Out\). However, without additional assumptions, there are stable profiles in which both types play \(Out\), and the receiver responds to \(In\) with the best response to \(\theta_2\) regardless of the message sent. In contrast, stability does have substantial predictive power under initial trust. Initial trust, which amounts to an implicit restriction on the receiver prior, says that receivers give the sender the “benefit of the doubt” and act in accordance with certain claims they have not previously seen proved false.\(^7\) It does

\(^6\) “I am playing \(s\) and my type is in \(\tilde{\Theta}\)” need not be the literal content of the message. For instance, \(m_{s,\tilde{\Theta}}\) could represent an argument like e.g. “I am playing signal \(s\) so you should believe my type is in \(\tilde{\Theta}\) because...”

\(^7\) Initial trust is similar in spirit 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.
not require that the receivers are certain that these claims are true, only that they 
give them a sufficiently high probability of being true. Of course, the receiver may 
quickly learn to distrust claims that prove to be false, which is why Assumption 2 is 
only applied to claims for which no direct contradictory evidence exists. We maintain 
Assumptions 1 and 2 throughout the main text; in Section 6 we discuss an alternative 
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(Θ, s). These 
are the undominated responses to s; the other responses are conditionally dominated 
in the sense of Fudenberg and Tirole [1991b]. Thus Δ(BR(Θ, s)) is the set of receiver 
mixed actions that assign probability 1 to undominated responses.

Definition 1 (Fudenberg and He [2018]). Strategy profile π is a perfect Bayesian 
equilibrium with heterogeneous off-path beliefs (PBE-H) if

1. For each θ ∈ Θ, u_1(θ, π) = max_{(s,m)∈S×M} u_1(θ, s, π_2(·|s,m)).
2. For each on-path signal-message pair (s, m), π_2(·|s,m) ∈ Δ(BR(p(·|s,m), s)),
   where p(·|s,m) is the posterior belief given (s,m) obtained through Bayes’ rule.
3. For each off-path signal-message pair \((s, m)\), \(\pi_2(\cdot | s, m) \in \Delta(BR(\Theta, s))\).

Conditions 1 and 2 of Definition 1 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.

Conditions 1–3 together are slightly weaker than perfect Bayesian equilibria (PBE, Fudenberg and Tirole [1991a]). This is because PBE replaces Condition 3 with the requirement that the receiver response to each \((s, m)\) is in the set

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

of mixed best responses to \(s\).\(^8\) \(\Delta(BR(\Theta, s))\) can be strictly larger than \(MBR(\Theta, s)\) because it can include mixtures over actions that are not best responses to the same belief.

Justified communication equilibrium adds the “justified-response” condition to PBE-H. To define this condition, for each \((s, \pi)\) let

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

This is the set of mixtures over receiver responses to \(s\) that are optimal for some belief over \(\Theta\) and would make type \(\theta\) strictly prefer \(s\) to their outcome under \(\pi\).\(^9\) Let

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

be the corresponding set for which type \(\theta\) would be indifferent between \(s\) and their

\(^8\)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.

\(^9\)This set is very similar to the set \(D_\theta\) used by Cho and Kreps [1987] to formulate NWBR; we discuss the differences in Section 3.3.
outcome under $\pi$. For every $s \in S$ and $\pi \in \Pi_1 \times \Pi_2$, let

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

be the set of types $\theta$ where there is some mixed receiver action $\alpha \in \Delta(BR(\Theta(s), s))$ that makes $\theta$ weakly prefer $(s, \alpha)$ to their outcome under $\pi$ and no other type $\theta'$ strictly prefer $(s, \alpha)$ to their outcome under $\pi$.

**Definition 2.** The set of justified types for signal $s$ given profile $\pi$ is

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

A justified response $\alpha \in \Delta(BR(\Theta(s, \pi), s))$ assigns positive probability only to actions that are best responses to beliefs with support in $\Theta(s, \pi)$.$^{10}$

**Definition 3.** The strategy profile $\pi$ 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 $\pi_2(\cdot | s, m) \in \Delta(BR(\Theta(s, \pi), s))$.

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 Condition 2 of Definition 3 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.

$^{10}$We show in A.3 that $\Theta^\dagger(s, \pi) = \emptyset$ only when $s$ is equilibrium dominated for all types, so how to define $\Theta(s, \pi)$ in this case is not important.
3.2 Only Justified Communication Equilibria are Stable

**Theorem 1.** If $\pi$ is stable, then it is a justified communication equilibrium.

Condition 3 of Definition 1 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 1, as well as the additional requirement in Definition 3, using three supporting lemmas, the proofs of which are in C.1.

To establish that every stable profile is a PBE-H, we first show that the aggregate sender play is optimal given the aggregate receiver play.

**Lemma 1.** Suppose that $\pi$ is stable. Then, for each $\theta \in \Theta$, $\pi_1(\cdot | \theta)$ puts support only on those sender signal-message pairs that are optimal for type $\theta$ under the receiver behavior strategy $\pi_2$.

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 follows from combining this 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 $\pi$ is stable. Then for any sender signal-message pair $(s, m)$ that occurs with positive probability under $\pi$, $\pi_2(\cdot | s, m)$ puts support only on receiver actions that are best-responses to $s$ and the posterior belief induced by $\lambda$ 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 with high probability, and since receivers myopically optimize, the lemma follows.
Neither Lemma 1 nor Lemma 2 requires Assumptions 1 or 2. Lemma 1 implies Condition 1 of the definition of PBE-H, 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 $\tilde{\Theta} \subseteq \Theta$, if every type $\theta \notin \tilde{\Theta}$ strictly prefers their payoff under $\pi$ to their payoff from playing $(s, m_{s,\tilde{\Theta}})$ (and having the receiver respond with $\pi_2(\cdot|s, m_{s,\tilde{\Theta}})$), then the aggregate receiver response to $(s, m_{s,\tilde{\Theta}})$ must be supported on $BR(\tilde{\Theta}, s)$.

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, the example in B.1 shows that, without initial trust, there can be stable profiles that are not JCE (and may not even satisfy the Intuitive Criterion). This is why the learning foundation of JCE requires cheap-talk communication and initial trust.

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

To get some intuition for this result, note that, when $u_1(\theta, s, \pi_2(\cdot|s, m_{s,\tilde{\Theta}})) < u_1(\theta, \pi)$ for all $\theta \notin \tilde{\Theta}$, Lemma 2 implies that the aggregate probability that a type outside of $\tilde{\Theta}$ plays $(s, m_{s,\tilde{\Theta}})$ is small if the prevailing aggregate strategy profile is close to $\pi$, $\delta$ is close to 1, and, given $\delta$, $\gamma_1$ 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 $\tilde{\Theta}$ play the signal-message pair $(s, m_{s,\tilde{\Theta}})$ becomes arbitrarily small as the aggregate probability of such play by types outside of $\tilde{\Theta}$ approaches 0. Recall that receivers who have never observed a type outside of $\tilde{\Theta}$ play $(s, m_{s,\tilde{\Theta}})$ would respond to $(s, m_{s,\tilde{\Theta}})$ with some action in $BR(\tilde{\Theta}, s)$. Combining these facts, it follows that the share of receivers who play some action in $BR(\tilde{\Theta}, s)$ in response to $(s, m_{s,\tilde{\Theta}})$ 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.** Let $\pi$ be a stable profile. We have already established that $\pi$ must be a PBE-H. We now show that the additional condition in Definition 3 holds.
Fix $s \in S$. Since $u_1(\theta, s, \pi_2(\cdot | s, m_s, \Theta(s, \pi))) \leq u_1(\theta, \pi)$ holds for all $\theta \in \Theta(s, \pi)$, it must be that $u_1(\theta, s, \pi_2(\cdot | s, m_s, \Theta(s, \pi))) < u_1(\theta, \pi)$ for all $\theta \notin \Theta(s, \pi)$. Lemma 3 then implies that $\pi_2(\cdot | s, m_s, \Theta(s, \pi)) \in \Delta(BR(\Theta(s, \pi), s))$. ■

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 rules out strategy profiles that commonly used refinements like the Intuitive Criterion and D1 preserve. Conversely, the following example shows that stable profiles need not be PBE and a fortiori need not satisfy any refinements of PBE.\footnote{The equilibrium refinements in Fudenberg and He \citeyear{fudenberg2018equilibrium} and Fudenberg and He \citeyear{fudenberg2020refinements} also relax PBE to PBE-H, but those papers do not show that this relaxation is needed.}

**Example 1.** The sender’s type space is $\Theta = \{\theta_1, \theta_2\}$, with $\lambda(\theta_1) = \lambda(\theta_2) = 1/2$. The sender’s signal space is $S = \{In, Out\}$, their message space is $M = \{m_{\theta_1}, m_{\theta_2}, m_{\Theta}\}$, and the receiver’s action space is $A = \{a_1, a_2, a_3\}$. The payoffs are given in Table 1.

<table>
<thead>
<tr>
<th>$\theta_1$</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$a_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$In$</td>
<td>$-2, 2$</td>
<td>$1, -2$</td>
<td>$1, 3$</td>
</tr>
<tr>
<td>$Out$</td>
<td>$0, 0$</td>
<td>$0, 0$</td>
<td>$0, 0$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\theta_2$</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$a_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$In$</td>
<td>$1, -2$</td>
<td>$-2, 2$</td>
<td>$1, 3$</td>
</tr>
<tr>
<td>$Out$</td>
<td>$0, 0$</td>
<td>$0, 0$</td>
<td>$0, 0$</td>
</tr>
</tbody>
</table>

Table 1: The payoffs for Example 1.

This game does not have a PBE 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 $\alpha[a_2] > 0$, 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 $\pi$ 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 JCE, because $\Theta(In, \pi) = \Theta$, and $a_1, a_2 \in BR(\Theta, s)$. Moreover, A.2 shows that both types playing $Out$ can be a stable outcome. □

We give direct proofs of stability in Example 1 and most of our other examples. We also give sufficient conditions for stability in the “strictly monotonic” games analyzed in Appendix C.2. Our general approach is to modify the aggregate response
correspondence so that all the resulting fixed points coincide with the strategy profile of interest in the limit, and then show that these fixed points are also fixed points of the true aggregate response correspondence.

3.3 Relation to Other Equilibrium Refinements

By definition, a justified communication equilibrium need only be a PBE-H rather than a PBE. 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 PBE. 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 profiles satisfying D1 that are not JCE. We provide detailed definitions of these refinements and formal proofs of their relationships with JCE in Appendix A. 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.

To compare NWBR with JCE, we first develop some notation. For every \( \theta \in \Theta \), \( s \in S \), 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
Let \( \Theta \) and would make type \( \theta \) 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 \( \theta \) would be indifferent with their equilibrium outcome. For every \( s \in S \) and \( \pi \in \Pi_1 \times \Pi_2 \), let \( \Theta^\dagger(s, \pi) = \{ \theta \in \Theta : D^0_\theta(s, \pi) \not\subseteq \bigcup_{\theta' \neq \theta} D_{\theta'}(s, \pi) \} \) be the set of types \( \theta \) where there is some mixed receiver best response \( \alpha \in MBR(\Theta, s) \) that makes \( \theta \) indifferent between \( (s, \alpha) \) and their equilibrium outcome, and no other type \( \theta' \) strictly prefers \( (s, \alpha) \) to their equilibrium outcome.

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

\[
\tilde{\Theta}(s, \pi) = \begin{cases} 
\Theta^\dagger(s, \pi) & \text{if } \Theta^\dagger(s, \pi) \neq \emptyset, \\
\Theta & \text{if } \Theta^\dagger(s, \pi) = \emptyset.
\end{cases}
\]

Also, let \( MBR(\tilde{\Theta}, s) = \{ \alpha \in \Delta(A) : \exists p \in \Delta(\tilde{\Theta}) \text{ s.t. } u_2(p, s, \alpha) \geq u_2(p, s, a) \ \forall a \in A \} \) denote the set of mixed best responses to \( s \) for beliefs supported on a given \( \tilde{\Theta} \subseteq \Theta \).
Definition 4 (Kohlberg and Mertens [1986], Cho and Kreps [1987]). Strategy profile $\pi$ satisfies never a weak best response (NWBR) if, for every $s \in S$, there is some $\alpha \in MBR(\tilde{\Theta}(s, \pi), s)$ such that $u_1(\theta, \alpha) \leq u_1(\theta, \pi)$ for all $\theta \in \Theta$.

Up to path-equivalence, JCE selects the same profiles as NWBR would if the mixed best responses $MBR(\tilde{\Theta}, s)$ were replaced with the convex hulls of best responses $\Delta(BR(\bar{\Theta}, s))$. Indeed, as shown in OA.2, it would be equivalent to define JCE by setting $\Theta^\dagger(s, \pi) = \{\theta \in \Theta : \tilde{D}_0^\theta(s, \pi) \not\subseteq \cup_{\theta' \neq \theta} \tilde{D}_{\theta'}(s, \pi)\}$, rather than $\Theta^\dagger(s, \pi) = \{\theta \in \Theta : \tilde{D}_{\theta}(s, \pi) \cup \tilde{D}_0^\theta(s, \pi) \not\subseteq \cup_{\theta' \neq \theta} \tilde{D}_{\theta'}(s, \pi)\}$. In this sense, JCE is a natural adaptation of NWBR with a learning foundation. Moreover, NWBR and JCE are essentially equivalent whenever the sets of mixed best responses and the corresponding convex hulls of best responses are the same. One setting that satisfies this condition is when there are at most two undominated receiver responses to each signal.

Proposition 2. In any game in which $MBR(\bar{\Theta}, s) = \Delta(BR(\bar{\Theta}, s))$ for all $s \in S$ and $\bar{\Theta} \subseteq \Theta$, a strategy profile $\pi$ is a justified communication equilibrium if and only if it is path-equivalent to a PBE-H that satisfies NWBR.

In general, NWBR is a weakly stronger requirement than JCE: Every profile that satisfies NWBR is path-equivalent to a JCE.

Proposition 3. If $\pi$ is a PBE-H that satisfies NWBR, then $\pi$ is path-equivalent to a JCE.

A.1 proves Proposition 3. The converse of the proposition is false: Example OA 3 in Online Appendix Section OA.7.3 shows that there can be JCE that are PBE (and so are PBE-H), but do not satisfy NWBR.

12 Fudenberg and Kreps [1988] and Sobel, Stole, and Zapater [1990] recognized that the convex hull of best responses is more natural in a learning setting, and discussed some of the differences this could make, but neither paper showed that NWBR rules out a profile that is stable in a learning model. 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 Path equivalence is needed in the statement of this result because, unlike JCE, NWBR does not impose requirements about the receiver’s actual responses to off-path play.
4 Co-Monotonic Signaling Games

Co-monotonic signaling games are games in which all sender types share the same preference over mixed receiver actions with support in $BR(\Theta, s)$.

Definition 5. 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 for a signaling game to be co-monotonic is that there be functions $v : S \times A \rightarrow \mathbb{R}$, $\omega : \Theta \times S \rightarrow \mathbb{R}_+$, and $\psi : \Theta \times S \rightarrow \mathbb{R}$ such that $u_1(\theta, s, a) = \omega(\theta, s)v(s, a) + \psi(\theta, s)$ for all $\theta \in \Theta$, $s \in S$, and $a \in A$. Many games, including the following simple economic example, satisfy this condition.

Example 2. The type space is $\Theta = \{1, 2, 3\}$, with $\lambda(1) = \lambda(2) = \lambda(3) = 1/3$. The signal space is $S = \{0, 10, 20, ..., 100\}$, and $A = \{0, 10, 20, ..., 60\}$. 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 firm’s choice of $s$ reflects a payment to the employee that can provide a costly signal of the firm’s knowledge, captured by $\theta$. 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.6 in the Online Appendix shows that JCE selects only equilibria that approximate the least-cost separating equilibrium of this game. □

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 best response $\alpha' \in MBR(\Theta, s)$ such

\(^{15}\)Recall that the convex hull of the mixed best responses, i.e. $\text{co}(MBR(\Theta, s))$, equals $\Delta(BR(\Theta, s))$.\]
that \( u_1(\theta, s, \alpha) = u_1(\theta, s, \alpha') \) for all \( \theta \). This ensures that \( \overline{\Theta}(s, \pi) = \hat{\Theta}(s, \pi) \) for every PBE-H \( \pi \).

**Lemma 4.** In a co-monotonic signaling game, \( \overline{\Theta}(s, \pi) = \hat{\Theta}(s, \pi) \) for all \( s \in S \) and PBE-H \( \pi \in \Pi \).

The proof of Lemma 4 is in A.1.

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 4) 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, every justified communication equilibrium is a PBE-H that satisfies NWBR, and every PBE-H that satisfies NWBR is path-equivalent to a justified communication equilibrium.

**Proof of Proposition 4.** Suppose that \( \pi \) is a PBE-H that satisfies NWBR. Then, by Proposition 3, \( \pi \) is path-equivalent to a JCE.

If \( \pi \) is a JCE, 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 \( a_s \in BR(\overline{\Theta}(s, \pi), s) \) such that \( a_s \in \arg \min_{a \in BR(\overline{\Theta}(s, \pi), s)} u_1(\theta, s, a) \) for all \( \theta \in \Theta \), so \( u_1(\theta, s, a_s) \leq u_1(\theta, \pi) \) for all \( \theta \in \Theta \). Since the game is co-monotonic, Lemma 4 implies that \( a_s \in BR(\hat{\Theta}(s, \pi), s) \), so \( \pi \) is a PBE-H that satisfies NWBR. ■

A consequence of Proposition 4 and the observation that every PBE-H that satisfies NWBR is path-equivalent to a PBE is that, in co-monotonic signaling games, every JCE is path-equivalent to a PBE that satisfies NWBR. Moreover, as shown by Cho and Sobel [1990], NWBR and D1 coincide in monotonic games, so JCE is also path-equivalent to D1 in co-monotonic games.

**Corollary 2.** In a co-monotonic signaling game, every justified communication equilibrium is path-equivalent to a PBE that satisfies NWBR and D1.
Thus, JCE provides a learning foundation for restricting attention to D1 equilibria in co-monotonic games, as in e.g. Nachman and Noe [1994], DeMarzo and Duffie [1999], and DeMarzo, Kremer, and Skrzypacz [2005].

In various co-monotonic games, such as that of DeMarzo and Duffie [1999], JCE selects the least-cost separating equilibrium outcome, often called the “Riley outcome” (Riley [1979]). Moreover, Cho and Sobel [1990] showed that NWBR selects the Riley outcome 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 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.

---

16Technically, the game analyzed in DeMarzo, Kremer, and Skrzypacz [2005] is not a traditional signaling game because of the presence of multiple senders; however, this distinction is not important.

17As noted by e.g. Fudenberg and Tirole [1991b], it may seem odd that adding a type with a small probability $\varepsilon$ can make a large change in the Riley outcome. Stability tracks this change in the Riley outcome because we hold the prior fixed as we take $\delta$, $\gamma_1$, and $\gamma_2$ to 1 in the iterated limit. For smaller $\varepsilon$, $\delta$, $\gamma_1$, and $\gamma_2$ must be closer to their limit values for the set of steady states to approximate its limit.

18Kalai and Lehrer [1993], Lehrer and Solan [2007], Esponda [2013], Battigalli, Francetich, Lanzani, and Marinacci [2019] studied 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
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. It assumed that the senders’ prior beliefs over the aggregate receiver responses are independent across signals, so that the senders’ optimal policy is given by the Gittins index (Gittins [1979]). 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 8 in A.5. 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 this with the assumptions of initially trusting receivers and relatively long-lived senders.

We view initial trust as a plausible and appealingly simple assumption. It 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 refutations provided by other receivers. Initial trust is also related to the restrictions imposed by Rabin [1990], Farrell [1993], and Clark [2020] on how receivers respond to certain “credible” messages in signaling game equilibria. In these papers, common knowledge of the equilibrium to be played figures heavily in determining the credibility of messages; such restrictions do not fit with our model of non-equilibrium learning. Moreover, deriving restrictions
on equilibria from a learning model and assumptions on the receiver’s prior yields more insight than imposing the restrictions directly.

## 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 and non-convex set of mixed best responses.

We can obtain similar solution concepts by replacing initial trust with alternative assumptions. 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.8.1, 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 the senders’ incentives. Any stable profile under this assumption must satisfy a refinement that is similar to, but weaker, than JCE.

We close with comments on some extensions of our results. First, all 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 $(\delta, \gamma_2) \to (1, 1)$, without putting conditions on the relative speed with which $\delta$ and $\gamma_2$ converge to 1. Moreover, as we show in OA.9, the conclusion of Theorem 1 applies under this general limit to all stable profiles satisfying certain conditions, such as on-path strict incentives for the receiver.

Next, an extensive experimental literature shows that a non-trivial share of experimental subjects tell the truth even when this earns less compensation, and thus behave as if they face a cost of lying. Kartik, Ottaviani, and Squintani [2006] and Kartik [2009] incorporated messages with such lying costs into models of strategic communication. OA.8.2 discusses how our analysis can be extended to signaling games with costly lying. Intuitively, lying costs make it less appealing for a non-justified type to falsely represent themself as justified.

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 More on Other Refinements

A.1 NWBR

Proof of Proposition 3: The proof of Proposition 3 uses the following two lemmas.

Lemma 5. $\Theta^1(s, \pi) \subseteq \Theta^1(s, \pi)$ for all $s \in S$ and $\pi \in \Pi$. 
Proof. If $\theta \not\in \Theta^\dagger(s, \pi)$, then by definition $\tilde{D}_0^\theta(s, \pi) \subseteq \cup_{\theta' \neq \theta} \tilde{D}_{\theta'}(s, \pi)$. For $\alpha \in D_0^\theta(s, \pi)$, $\alpha \in MBR(\Theta, s) \subseteq \Delta(BR(\Theta, s))$ and $u_1(\theta, s, \alpha) = u_1(\theta, \pi)$, so $\alpha \in \tilde{D}_0^\theta(s, \pi)$. Since $\tilde{D}_0^\theta(s, \pi) \subseteq \cup_{\theta' \neq \theta} \tilde{D}_{\theta'}(s, \pi)$, 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 $\alpha$ 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 $\theta \not\in \Theta^\dagger(s, \pi)$. 

\[ \]

Lemma 6. If $\pi$ is a PBE-H that satisfies NWBR, 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 $\theta \in \Theta$ and $a \in BR(\Theta, s)$.

Proof. Let $\pi$ be a PBE-H that satisfies NWBR. Fix $s \in S$ and suppose that $\Theta^\dagger(s, \pi) = \emptyset$. 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 $\pi$. 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 $\pi$. $\mathcal{A}_-$ and $\mathcal{A}_+$ are disjoint open subsets of $MBR(\Theta, s)$, and $\mathcal{A}_- \cup \mathcal{A}_+ = MBR(\Theta, s)$ since $\Theta^\dagger(s, \pi) = \emptyset$. 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 $\pi$ is a PBE-H that satisfies NWBR since then, for every $\alpha \in MBR(\Theta(s, \pi), s)$, there is some $\theta$ 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 3. Let $\pi$ be a PBE-H that satisfies NWBR. We will show that, for every off-path $s$, there is a justified response $\alpha_s \in \Delta(BR(\Theta(s, \pi), s))$ such that $u_1(\theta, s, \alpha_s) \leq u_1(\theta, \pi)$ for all $\theta \in \Theta$. This means that the profile $\tilde{\pi} = (\pi_1, \pi_2)$ in which $\pi_2$ coincides with $\pi_2$ for all on-path $s$ and dictates $\alpha_s$ for all off-path $s$ is a justified communication equilibrium that is path-equivalent to $\pi$.

If $\Theta^\dagger(s, \pi) \neq \emptyset$, then by Lemma 5, $\Theta^\dagger(s, \pi) \subseteq \Theta^\dagger(s, \pi)$, so $\Theta(s, \pi) \subseteq \Theta(s, \pi)$. If $\Theta^\dagger(s, \pi) = \emptyset$, then by Lemma 6, $u_1(\theta, s, a) < u_1(\theta, \pi)$ for all $\theta \in \Theta$ and $a \in BR(\Theta, s)$, so $\Theta^\dagger(s, \pi) = \emptyset$ and $\Theta(s, \pi) = \Theta(s, \pi) = \Theta$. Either way, $\Theta(s, \pi) \subseteq \Theta(s, \pi)$, so
MBR(\(\Theta(s, \pi), s\)) \subseteq \Delta(BR(\Theta(s, \pi), s))$. Since \(\pi\) satisfies NWBR, there is an \(\alpha \in MBR(\Theta(s, \pi), s) \subseteq \Delta(BR(\Theta(s, \pi), s))\) such that \(u_1(\theta, s, \alpha) \leq u_1(\theta, \pi)\) for all \(\theta\). \(\blacksquare\)

**Proof of Lemma 4:** Fix \(s \in S\) and PBE-H \(\pi\). We show that, for all \(\theta \in \Theta\),  
\[\tilde{D}_\theta(s, \pi) \cup \tilde{D}_\theta^0(s, \pi) \not\subseteq \cup_{\theta' \neq \theta} \tilde{D}_{\theta'}(s, \pi)\]  
if and only if \(D_\theta^0(s, \pi) \not\subseteq \cup_{\theta' \neq \theta} D_{\theta'}(s, \pi)\). This means that \(\Theta^\dagger(s, \pi) = \Theta^\dagger(s, \pi)\), which implies that \(\Theta(s, \pi) = \tilde{\Theta}(s, \pi)\).

Suppose that \(D_\theta^0(s, \pi) \not\subseteq \cup_{\theta' \neq \theta} D_{\theta'}(s, \pi)\). Then there is some \(\alpha \in MBR(\Theta(s, \pi))\) 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 \(\alpha \in \Delta(BR(\Theta(s, \pi))\), this immediately implies that \(\tilde{D}_\theta(s, \pi) \cup \tilde{D}_\theta^0(s, \pi) \not\subseteq \cup_{\theta' \neq \theta} \tilde{D}_{\theta'}(s, \pi)\).

Suppose that \(\tilde{D}_\theta(s, \pi) \cup \tilde{D}_\theta^0(s, \pi) \not\subseteq \cup_{\theta' \neq \theta} \tilde{D}_{\theta'}(s, \pi)\). Then there is some \(\alpha \in \Delta(BR(\Theta(s, \pi))\) such that \(u_1(\theta, s, \alpha) \geq u_1(\theta, \pi)\) and \(u_1(\theta', s, \alpha) \leq u_1(\theta', \pi)\) for all \(\theta' \neq \theta\). Moreover, since \(\pi\) is a PBE-H, there is some \(\alpha' \in \Delta(BR(\Theta(s, \pi))\) such that \(u_1(\theta, s, \alpha') \leq u_1(\theta, \pi)\). By continuity, there exists some \(\alpha'' \in MBR(\Theta(s, \pi))\) such that \(u_1(\theta, s, \alpha'') = u_1(\theta, \pi)\) \(\leq u_1(\theta, s, \alpha)\). Because the game is co-monotonic, \(u_1(\theta', s, \alpha'') \leq u_1(\theta', s, \alpha) \leq u_1(\theta', \pi)\) holds for all \(\theta' \neq \theta\). Thus, \(D_\theta^0(s, \pi) \not\subseteq \cup_{\theta' \neq \theta} D_{\theta'}(s, \pi)\). \(\blacksquare\)

### A.2 Stability Does Not Imply PBE

**Proposition 5.** The game in Example 1 has stable profiles where both types play Out with probability 1, even though this is not the outcome of a PBE.

The main text already explained why both types playing Out is not a PBE outcome. The proof that the outcome is stable, presented in OA.4, 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 \(\theta\), 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 \((In, m_\theta)\).) We show that, holding \(\gamma_2\) 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. To show this, we modify the aggregate
response correspondence so that the aggregate receiver response to any \((In, m)\) plays \(a_3\) with probability no more than \(1/4\), so it is not weakly optimal for both types to play \(In\) with the same message.\(^{21}\) This means that, in the limit, most receivers never encounter \((In, m)\) being played by both types, and the assumptions on their prior imply that these receivers do not play \(a_3\). Indeed, it turns out that in the limit most receivers have either not encountered a sender play \((In, m)\) and so play either \(a_1\) or \(a_2\) (depending on their prior), or have encountered sender types playing \((In, m)\), which overturned their prior and leads them to play either \(a_1\) or \(a_2\). Thus the constraint on the aggregate receiver response is not binding, and the fixed points of the modified aggregate response correspondence are fixed points of the true aggregate response correspondence. Next we show that 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 \(\theta\) 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 \(u_1(\theta, In, BR(\theta, In)) < 0\).

A.3 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 \(\pi\) in the sense of Cho and Kreps [1987].

**Definition 6** (Cho and Kreps [1987]). *Strategy profile \(\pi\) 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 6.** If \(\pi\) is a justified communication equilibrium, then \(\pi\) is a PBE-H that passes the Intuitive Criterion Test.

The key step of the proof is to show that in a PBE-H, unless \(s\) is equilibrium dominated for every type, \(s\) is not equilibrium dominated for any justified type. So

\(^{21}\)Either \(\pi_2(a_1|In, m) \geq 3/8\) which deters \(\theta_1\), or \(\pi_2(a_2|In, m) \geq 3/8\), which deters \(\theta_2\).
if there is a justified response that deters all types from playing \( s \), then the profile passes the Intuitive Criterion Test. Formally, we use the following lemma to show that \( \Theta(s, \pi) \subseteq E(s, \pi) \) when \( E(s, \pi) \neq \emptyset \).

**Lemma 7.** If \( \pi \) 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 \( \theta \in \Theta \) and \( a \in BR(\Theta, s) \).

The proof of Lemma 7 is analogous to that of Lemma 6 and is given in OA.5.

**Proof of Proposition 6.** Fix \( s \in S \) and suppose that \( E(s, \pi) \neq \emptyset \). This means that there is some \( \theta \) and \( \alpha \in BR(\Theta, s) \) such that \( u_1(\theta, s, a) < u_1(\theta, \pi) \). By Lemma 7, \( \Theta^\dagger(s, \pi) \neq \emptyset \), so \( \overline{\Theta}(s, \pi) = \Theta^\dagger(s, \pi) \). Moreover, \( \overline{\Theta}(s, \pi) \subseteq E(s, \pi) \) since every type \( \theta \) for which \( \max_{a \in BR(\Theta, s)} u_1(\theta, s, a) < u_1(\theta, \pi) \) satisfies \( D_\theta(s, \pi) \cup D_\theta^0(s, \pi) = \emptyset \neq D_{\theta'} \) for every \( \theta' \in \Theta \). 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 \( \alpha \) 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 \( \theta \in E(s, \pi) \). \( \blacksquare \)

### A.4 D1 and co-D1

For every \( s \in S \) and \( \pi \in \Pi_1 \times \Pi_2 \), let \( \Theta^{\dagger, D1}(s, \pi) = \{ \theta \in \Theta : \forall \theta' \neq \theta, D_\theta(s, \pi) \cup D_\theta^0(s, \pi) \nsubseteq D_{\theta'}(s, \pi) \} \) be the set of types \( \theta \) where, for every \( \theta' \neq \theta \), there is some receiver mixed best response \( \alpha \in MBR(\Theta, s) \) that makes \( \theta \) weakly prefer \((s, \alpha)\) to their equilibrium outcome, while \( \theta' \) weakly prefers their equilibrium outcome to \((s, \alpha)\).

Let \( \widehat{\Theta}^{D1}(s, \pi) \subseteq \Theta \) be the set of types given by

\[
\widehat{\Theta}^{D1}(s, \pi) = \begin{cases} 
\Theta^{\dagger, D1}(s, \pi) & \text{if } \Theta^{\dagger, D1}(s, \pi) \neq \emptyset \\
\Theta & \text{if } \Theta^{\dagger, D1}(s, \pi) = \emptyset 
\end{cases}
\]

**Definition 7** (Banks and Sobel [1987]). A PBE \( \pi \) 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 $\Theta^{\dagger,D_1}(s,\pi) = \{ \theta \in \Theta : \forall \theta' \neq \theta, \tilde{D}_\theta(s,\pi) \cup \tilde{D}_\theta(s,\pi) \not\subseteq \tilde{D}_{\theta'}(s,\pi) \}$ be the set of types $\theta$ where, for every $\theta' \neq \theta$, there is some mixed receiver action $\alpha \in \Delta(BR(\Theta, s))$ that makes $\theta$ weakly prefer $(s, \alpha)$ to their equilibrium outcome and $\theta'$ weakly prefer their equilibrium outcome to $(s, \alpha)$. Let $\Theta^{D_1}(s,\pi) \subseteq \Theta$ be the set of types given by

$$
\Theta^{D_1}(s,\pi) = \begin{cases} 
\Theta^{\dagger,D_1}(s,\pi) & \text{if } \Theta^{\dagger,D_1}(s,\pi) \neq \emptyset \\
\emptyset & \text{if } \Theta^{\dagger,D_1}(s,\pi) = \emptyset.
\end{cases}
$$

**Definition 8.** A PBE-H $\pi$ is co-D1 if for every $s \in S$, there is an $\alpha \in \Delta(BR(\Theta, s))$ such that $u_1(\theta, s, \alpha) \leq u_1(\theta, \pi)$ for all $\theta \in \Theta$.

**Proposition 7.** If $\pi$ is a justified communication equilibrium, then $\pi$ is a PBE-H that is co-D1.\(^{22}\)

Co-D1 is a more permissive refinement than JCE because it strikes fewer types. The following proof shows that $\Theta(s,\pi) \subseteq \Theta^{D_1}(s,\pi)$ for all $s$; Example OA 2 shows that the inclusion is sometimes strict.

**Proof.** Fix $s \in S$. We will argue that $\Theta(s,\pi) \subseteq \Theta^{D_1}(s,\pi)$. This, along with the justified response criterion of JCE and the fact that every JCE is a PBE-H, implies that $\pi$ is co-D1.

If $\Theta^{\dagger}(s,\pi) \neq \emptyset$, then $\Theta^{\dagger}(s,\pi) = \Theta^{\dagger}(s,\pi)$. Let $\theta$ be a type such that $\theta \notin \Theta^{D_1}(s,\pi)$. Then there is some type $\theta' \neq \theta$ such that $D_\theta(s,\pi) \cup D_{\theta'}(s,\pi) \not\subseteq D_{\theta'}(s,\pi)$. This implies that $\theta' \notin \Theta(s,\pi)$, so $\Theta^{\dagger}(s,\pi) \subseteq \Theta^{D_1}(s,\pi)$ follows.

If $\Theta^{\dagger}(s,\pi) = \emptyset$, by Lemma 7, $u_1(\theta, s, a) < u_1(\theta, \pi)$ for all $a \in BR(\Theta, s)$. Thus $\Theta^{\dagger,D_1}(s,\pi) = \emptyset$ as $D_\theta(s,\pi) \cup D_{\theta'}(s,\pi) \not\subseteq D_{\theta'}(s,\pi)$ for all $\theta, \theta' \in \Theta$. Thus, $\Theta(s,\pi) = \Theta^{D_1}(s,\pi)$. \(\blacksquare\)

\(^{22}\)Divinity and universal divinity (Banks and Sobel [1987]) resemble D1 and co-D1 but use iterative procedures. As we show in OA.8.3, we can capture a similar iterated procedure by strengthening initial trust so that if the only types outside of $\Theta$ that have played $(s, m_{s,\theta})$ are in $\Theta'$, then the receiver responds to this signal-message pair with some element of $BR(\Theta \cup \Theta', s)$. 28
A.5 Rationality-Compatible Equilibrium

We write $\Pi^*_2 = \times_{(s,m) \in S \times M} \Delta(BR(\Theta, s))$ for the set of receiver strategies that assign probability 0 to conditionally dominated responses.

**Definition 9** (Fudenberg and He [2020a]). Signal $s \in S$ is more rationally-compatible with $\theta'$ than $\theta''$, written as $\theta' \succeq_s \theta''$, if, for every $\pi_2 \in \Pi^*_2$,

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

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

$$\overline{P}(s, \pi) = \begin{cases} 
\Delta(E(s, \pi)) \cap \left( \cap_{\theta' \triangleright s, \theta''} P_{\theta' \triangleright \theta''} \right) & \text{if } E(s, \pi) \neq \emptyset \\
\Delta(\Theta) & \text{if } E(s, \pi) = \emptyset.
\end{cases}$$

**Definition 10** (Fudenberg and He [2020a]). Strategy profile $\pi$ 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 $\theta'$ and $\theta''$ dominates the prior likelihood ratio whenever $\theta' \succeq_s \theta''$. It also requires that the posterior assigns probability 0 to equilibrium-dominated types.

**Proposition 8.** If $\pi$ is a justified communication equilibrium, then $\pi$ is an RCE.\textsuperscript{24}

\textsuperscript{23}Note that this criterion does not depend on the message space in the game: If $s \in S$ is more rationally-compatible with $\theta'$ than $\theta''$ under message space $M$, then it is also more rationally compatible with $\theta'$ than $\theta''$ under any other message space.

\textsuperscript{24}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.
Intuitively, any response that makes a less compatible type weakly prefer to play \( s \) makes more compatible types strictly prefer to play it, so less compatible types are not justified.

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

Since \( \Delta(\Theta(s, \pi)) \subseteq \Delta(\Theta) = P(s, \pi) \) when \( E(s, \pi) = \emptyset \), we need only handle the case where \( E(s, \pi) \neq \emptyset \). In this case by Lemma 7, \( \Theta(s, \pi) = \Theta^I(s, \pi) \) and \( \Delta(\Theta(s, \pi)) \subseteq \Delta(E(s, \pi)) \). Suppose that \( \theta' \) and \( \theta'' \) are two types such that \( \theta' \succ_s \theta'' \). Then Definition 10 implies that \( D_{\theta''}(s, \pi) \cup D_{\theta'}^0(s, \pi) \subseteq \tilde{D}_{\theta'}(s, \pi) \), so \( \theta'' \not\in \Theta^I(s, \pi) \). As a result, \( \Delta(\Theta(s, \pi)) = \Delta(\Theta^I(s, \pi)) \subseteq \cap_{(\theta', \theta'')} \text{s.t. } \theta' \succ_s \theta'' P_{\theta' \theta''} \). We conclude \( \Delta(\Theta(s, \pi)) \subseteq \Delta(E(s, \pi)) \cap \cap_{(\theta', \theta'')} \text{s.t. } \theta' \succ_s \theta'' P_{\theta' \theta''} = P(s, \pi) \). \( \blacksquare \)

**B Additional Examples**

**B.1 Stability without Initially Trusting Receivers**

Example OA 1 in OA.7.1 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, two signals, \( In \) and \( Out \), and two actions, \( a_1 \) and \( a_2 \). 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 only 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 \(\gamma_2\), so there are limits of steady states in which both types play \(Out\).\(^{25}\)

## B.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.7.2. In the example, type space is \(\Theta = \{\theta_1, \theta_2, \theta_3\}\), with \(\lambda(\theta_1) = \lambda(\theta_2) = \lambda(\theta_3) = 1/3\). The signal space is \(S = \{In, Out\}\), and \(A = \{a_1, a_2, a_3\}\). The payoff matrices are:

\[
\begin{array}{ccc}
\theta_1 & a_1 & a_2 & a_3 \\
In & 4, 1 & -1, 0 & -1, -1 \\
Out & 0, 0 & 0, 0 & 0, 0 \\
\end{array}
\quad
\begin{array}{ccc}
\theta_2 & a_1 & a_2 & a_3 \\
In & -1, 0 & 4, 1 & -1, -1 \\
Out & 0, 0 & 0, 0 & 0, 0 \\
\end{array}
\quad
\begin{array}{ccc}
\theta_3 & a_1 & a_2 & a_3 \\
In & 1, 0 & 1, 0 & -1, -1 \\
Out & 0, 0 & 0, 0 & 0, 0 \\
\end{array}
\]

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 \(\theta_3\) 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 \(\theta_3\) cannot be eliminated by D1, and because the receiver wants to play \(a_3\) versus \(\theta_3\), D1 allows equilibria in which every type plays Out. However, whenever the receiver response makes \(\theta_3\) weakly prefer \(In\), one of the other types strictly prefers to play \(In\). This means that \(\theta_3\) is not a justified type, so no equilibrium in which every type plays \(Out\) is a JCE. Intuitively, stability precludes every type playing \(Out\) because for the aggregate receiver response to use \(a_3\) whenever \(In\) is played, it must be that some \(\theta_3\) agents learn to play \(In\). But if \(\theta_3\)

\(^{25}\)Without initial trust, all types playing \(Out\) remains a stable outcome even if the signal space is expanded to include copies of \(In\) with identical payoffs, i.e. \(In_1, In_2, \text{etc.}\)
agents learn that it is weakly optimal to play $In$, then either the $\theta_1$ agents or the $\theta_2$ agents learn that it is strictly optimal to do so.

### B.3 Stability Does Not Imply D1

Example OA 3 in OA.7.3 shows that there can be stable profiles that are PBE but not D1. In the example, there are two types and three actions, $a_1$, $a_2$, and $a_3$. Type $\theta_1$ 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 $\theta_2$ 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 $\theta_1$, which leads the receiver to play $a_1$.

However, there are receiver priors that satisfy initial trust and support stable profiles in which all types play $Out$. A stable profile cannot have either type $\theta_2$ playing $In$ with probability 1 or type $\theta_2$ playing $In$ with positive probability and type $\theta_1$ playing $In$ with probability 0, 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 $\theta_1$ weakly prefers to play some $(In, m)$ than $Out$, $\theta_2$ 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.

### C Proofs Concerning Stability

#### C.1 Supporting Results for Theorem 1

We use the following lemma in several proofs and examples.
Lemma 8. 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 \( \theta \) under the receiver behavior strategy \( \pi_{2, \gamma_2} \).

The proof of Proposition 5 in Fudenberg and He [2018] establishes that, in a learning setting with the same continuation probability \( \gamma \) 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 \( \delta \to 1 \). Precisely the same arguments can be used to establish Lemma 8 by showing that, for fixed \( \gamma_2 \), 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.

Proof of Lemma 1: Let \( \{\pi_{j, k, l} \in \Pi^*(g, \delta_j, \gamma_{1, j, k, l}, \gamma_{2, j})\}_{j, k, l \in \mathbb{N}} \) 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 \( \lim_{j \to \infty} \gamma_{2, j} = 1 \), \( \lim_{k \to \infty} \delta_j = 1 \) for all \( j \), and \( \lim_{l \to \infty} \gamma_{1, j, k, l} = 1 \) for all \( j, k \). By Lemma 8, 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 signal-message pairs that 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 \( \pi_1(\cdot|\theta) = \lim_{j \to \infty} \pi_{1, \gamma_2, j}(\cdot|\theta) \) puts support only on signal-message pairs that are optimal under \( \pi_2 = \lim_{j \to \infty} \pi_{2, \gamma_2, j} \).

Proof of Lemma 2: Let \( q(\theta, s, m) = \lambda(\theta)\pi_1(s, m|\theta) \) be the distribution over sender type, signal, and message induced by \( \lambda \) and \( \pi_1 \). Also, let \( X_{on} \) be the set of sender signal-message pairs that occur with positive probability under \( \pi \), and let \( p(\theta|s, m) \) denote the conditional probability of \( \theta \) given \( (s, m) \in X_{on} \).

For \( \varepsilon > 0 \), let \( Q_{\varepsilon} = \{q' \in \Delta(\Theta \times S \times M) : \max_{(\theta, s, m)} |q'(\theta, s, m) - q(\theta, s, m)| \leq \varepsilon \} \).

\(^{26}\)That proof does not require the paper’s assumptions that the senders and receivers have the same expected lifetime and that the senders’ prior over receiver strategies is independent across signals.
By upper hemicontinuity, there exists some \( \varepsilon > 0 \) such that every receiver whose belief \( \tilde{g}_2 \in \Delta(\Delta(\Theta \times S \times M)) \) puts probability at least \( 1 - \varepsilon \) on \( Q_\varepsilon \) will respond to every \((s, m) \in X^m\) with some action belonging to \( \text{BR}(\rho(\cdot|s, m), s) \).

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

**Proof of Lemma 3:** 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 \( u_1(\theta, s, \pi_2(\cdot|s, m_s, \Theta)) < u_1(\theta, \pi) \) for all \( \theta \notin \tilde{\Theta} \), Lemma 8 implies that there is some \( J \in \mathbb{N} \) such that, for all \( \theta' \notin \tilde{\Theta} \) and \( j > J \), \( \lim_{k \to \infty} \lim_{l \to \infty} \pi_{j,k,l}(s, m_s, \Theta') = 0 \). Receivers who have never observed the signal-message pair \((s, \tilde{\Theta})\) played by a type outside of \( \tilde{\Theta} \) would respond to this pair with an action belonging to \( \text{BR}(\tilde{\Theta}, s) \). Thus if \( \lim_{k \to \infty} \lim_{l \to \infty} \pi_{j,k,l}(s, m_s, \Theta') = 0 \) for all \( \theta' \notin \tilde{\Theta} \) then \( \lim_{k \to \infty} \lim_{l \to \infty} \pi_{2,j,k,l}(\text{BR}(\tilde{\Theta}, s)|s, m_s, \Theta) = 1 \). Since this holds for all \( j > J \), \( \pi_2(\text{BR}(\tilde{\Theta}, s)|s, m_s, \Theta) = \lim_{j \to \infty} \lim_{k \to \infty} \lim_{l \to \infty} \pi_{2,j,k,l}(\text{BR}(\tilde{\Theta}, s)|s, m_s, \Theta) = 1 \).
C.2 A Sufficient Condition for Stability

Definition 11. 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 $\pi$, let $X_{on}$ be the set of on-path signal-message pairs, let $p(\theta|s,m)$ denote the conditional probability of $\theta$ 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 12. The JCE $\pi$ 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 $x = (s, m) \in X_{on}$, 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 $\theta \in \Theta$ and $a \in BR(\Theta(s, \pi), s)$.

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 types are strictly deterred from playing any off-path signal for any justified response.

Proposition 9. If $\pi$ is a 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 non-doctrinaire priors $g_1, g_2$, including those that do not satisfy initial trust.

OA.3 contains the proof of Proposition 9. Because $\pi$ is uniformly justified, there is a receiver behavior strategy which strictly incentivizes every type to play their corresponding signal in $\pi$, and, when each sender type does so, leads to the same distribution over $\Theta \times S \times A$ as $\pi$. The proof modifies the aggregate response correspondences so that the receiver response matches this behavior strategy with high probability whenever the aggregate sender play is such that some type gives their corresponding signal in $\pi$ too little probability. Lemma 8 implies the aggregate sender play given by the fixed points of the modified aggregate response correspondence 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 $\pi$ with high probability. 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 $\pi$, this means that the receiver response to any on-path signal-message pair matches the response in $\pi$. We show that this, along with the fact that $\pi$ is uniformly justified, implies that the limit aggregate sender play signals according to $\pi$ with probability 1. Consequently, the constraint from the modification of the aggregate receiver response is not binding, and the fixed points of the modified response mapping are valid steady state profiles that induce the same distribution over $\Theta \times S \times A$ as $\pi$ in the limit.

References


\footnote{Fudenberg and Levine [2006] and Fudenberg and He [2020a] prove that some strategy profiles are stable by considering priors that assign high probability to a neighborhood of the target profile. Modifying the aggregate response mapping lets us prove stability for a broad class of priors.}


