# Justified Communication Equilibrium

*By* Daniel Clark and Drew Fudenberg*

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 receivers are* initially trusting *and senders play many more times than receivers. In the learning 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 resembles the earlier refinements D1 and NWBR, and it coincides with them in* co-monotonic *signaling games.*

Cheap-talk communication is available in many of the settings signaling games are intended to model, and 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.

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) and D1 (Banks and Sobel, 1987), but JCE can make very different predictions in economically relevant settings. Unlike those refinements, JCE has a foundation in the theory of learning in games. We provide this foundation by analyzing the limits of steady states in an overlapping generations learning environment where agents are patient, have long expected lifetimes, and the senders on average 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

* Department of Economics, MIT, 77 Massachusetts Avenue, Cambridge, MA, 02139. Clark: dgclark@mit.edu. Fudenberg: drewf@mit.edu. We thank Glenn Ellison, Kevin He, David Kreps, David K. Levine, Harry Pei, Andy Skrypacz, Joel Sobel, Alex Wolitzky, two anonymous referees, and an editor for helpful comments, and NSF grants SES 1643517 and SES 1951056 for financial support.

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 V.C 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 add cheap talk to the extensive form of the signaling game, our analysis can and does specify how receivers respond to each possible message, including to the "null message" of saying nothing at all, so 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, and address the possible complications in adding explicit communication to the signaling game discussed in Fudenberg and Tirole (1991*a*).

Our results can be seen as both a validation of and a correction to previous signaling game refinements, which are only roughly in line with the implications of non-equilibrium learning. Specifically, none of the traditional equilibrium refinements is a necessary condition for stability in our learning model.[1] Indeed, as shown by Example 3, the stable outcomes of our learning model need not be perfect Bayesian equilibria (Fudenberg and Tirole, 1991*b*), 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 preview a few results here. As the left-hand box in Figure 1 illustrates, every JCE passes the "Intuitive Criterion Test," and every JCE is a rationality-compatible equilibrium (RCE, Fudenberg and He (2020)), which is the strongest previous equilibrium refinement for signaling games that has a learning

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

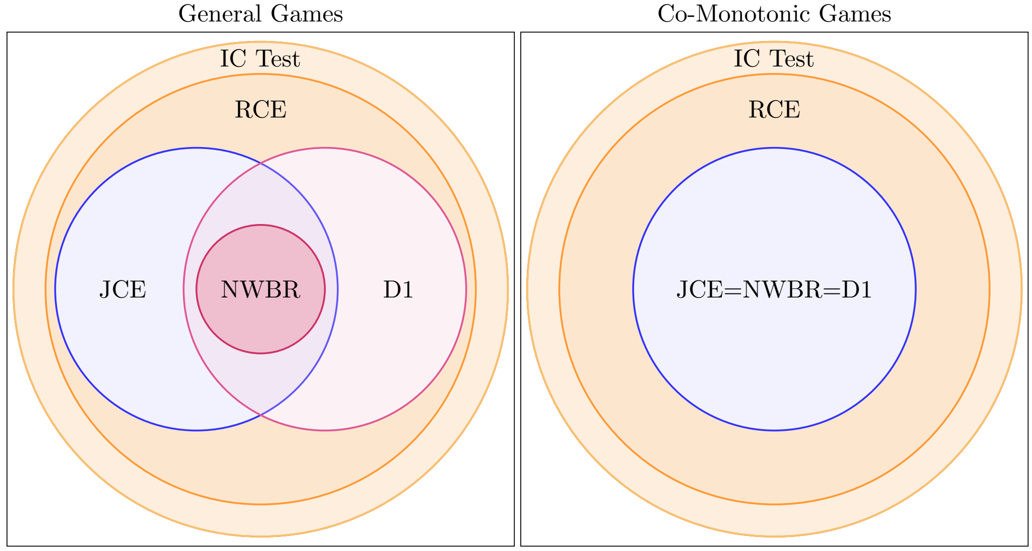


FIGURE 1. JCE AND OTHER RESTRICTIONS OF PBE-H.

foundation. JCE is not nested with D1, but every PBE-H that satisfies NWBR is path-equivalent to a JCE.[2] The right-hand box in Figure 1 depicts the fact that JCE, NWBR, and D1 are essentially equivalent in the special but important class of *co-monotonic* signaling games, which provides a learning justification for selecting the least-cost separating equilibria in many of these games.

## I. Preliminaries

### A. *Signaling Games with Communication*

In a signaling game with communication, the sender (player 1) has a type space $\Theta$, a signal space $S$, and a message space $M$. The sender observes their type, which is drawn from a full-support distribution $\lambda \in \Delta(\Theta)$, and then chooses a signal $s \in S$ and a message $m \in M$.[3] The receiver (player 2) 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. We assume that all of these sets are finite. 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*.

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 $u_2 : \Theta \times S \times A \to \mathbb{R}$. Each player's utility depends on the sender's

[2] 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).
[3] Throughout, we use $\Delta(\Omega)$ to denote the set of (Borel) probability distributions over a set $\Omega$.

signal and type and the receiver's action; neither utility depends on the message of the sender. We will abuse notation slightly and write $u_1(\theta, \pi)$ and $u_2(\pi)$ for the expected payoffs from strategy profile $\pi = (\pi_1, \pi_2)$, as well as $u_1(\theta, s, \alpha)$ for the expected utility of the type $\theta$ sender from playing signal $s$ when the receiver responds according to $\alpha \in \Delta(A)$, and $u_2(p, s, \alpha)$ for the expected utility of the receiver from playing $\alpha$ when the sender plays $s$ and the probability distribution over their type is $p \in \Delta(\Theta)$. Finally, $BR(\theta, s) = \arg\max_{a \in A} u_2(\theta, s, a)$ denotes the pure best responses for the receiver to signal $s$ when the sender's type is $\theta$, $BR(p, s) = \arg\max_{a \in A} u_2(p, s, a)$ denotes the pure best responses for the receiver to signal $s$ under belief $p \in \Delta(\Theta)$, and $BR(\widetilde{\Theta}, s) = \cup_{p \in \Delta(\widetilde{\Theta})} BR(p, s)$ denotes the pure best responses to signal $s$ for some $p$ with support in $\widetilde{\Theta}$.

### B. *Definition of Justified Communication Equilibrium*

The set of actions that are a best response to some belief about $\theta$ is $BR(\Theta, s)$. These are the *undominated responses* to $s$; the other responses are *conditionally dominated* in the sense of Fudenberg and Tirole (1991*a*). Thus $\Delta(BR(\Theta, s))$ is the set of receiver mixed actions that assign probability 1 to undominated responses.

DEFINITION 1 (Fudenberg and He, 2018): *Strategy profile $\pi$ 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 $(s, m)$, $\pi_2(\cdot|s, m) \in \Delta(BR(p_{(s,m)}, s))$, 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 (1991*b*)). 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 $s$.[4] $\Delta(BR(\Theta, s))$ 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 type $\theta$, signal $s$, and strategy profile $\pi$, let

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

[4]Recall that PBE and sequential equilibrium are equivalent in signaling games.

This is the set of mixtures over undominated receiver responses to $s$ that would make type $\theta$ strictly prefer $s$ to their outcome under $\pi$.[5] 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 $\theta$ would be indifferent between $s$ and their outcome under $\pi$. For every $s \in S$ and $\pi \in \Pi_1 \times \Pi_2$, let

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

be the set of types $\theta$ for which there is some mixed receiver action $\alpha \in \Delta(BR(\Theta, s))$ 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*

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

*A* **justified response** *to signal $s$ given profile $\pi$ is a distribution $\alpha \in \Delta(BR(\overline{\Theta}(s,\pi), s))$ that assigns positive probability only to actions that are best responses to beliefs with support in $\overline{\Theta}(s,\pi)$.*[6]

Note that in a PBE-H, each type is justified for every signal it plays with positive probability. This is because every signal-message pair the type is willing to play must give them their equilibrium payoff, while no other type can get strictly more than their equilibrium payoff by playing it.

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 $\pi_2(\cdot|s,m) \in \Delta(BR(\overline{\Theta}(s,\pi), s))$.*

Every JCE must be a PBE-H. The second condition requires that the receiver's response to each signal is justified for at least one message. Since the equilibrium response to on-path signal-message pairs is justified in any PBE-H, the substance of JCE comes from the requirement that there be a justified response to every off-path signal. As we will see, this conclusion only follows from our learning model when the message space is sufficiently large. However, the definition of JCE applies for any non-null message space, including the case without cheap talk, where the message space is singleton.

[5]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 III.

[6]Appendix A1 shows that if $\pi$ is a PBE-H, $\Theta^\dagger(s,\pi) = \emptyset$ only when $s$ is equilibrium dominated for all types, so how to define $\overline{\Theta}(s,\pi)$ in this case is not important.

### C. Hiring a Worker

This subsection presents two stylized examples of a firm (the sender) potentially hiring a worker (the receiver) for a particular job. In both examples, the firm's signal $s \in \{Hire, Pass\}$ is its choice of whether to hire the worker. The worker's choice of action $a \in \{e_H, e_M, e_L\}$ represents how hard they work; $e_H$ represents high effort, $e_M$ medium effort, and $e_L$ low effort. The firm has three possible types, $\Theta = \{\theta_H, \theta_M, \theta_L\}$: type $\theta_H$ represents high quality, $\theta_M$ medium quality, and $\theta_L$ low quality.[7] The payoffs to both parties are normalized to 0 when the firm does not hire. The examples differ only in their payoff functions when the firm hires. In the first example JCE rules out an equilibrium that satisfies D1, and in the second JCE preserves an equilibrium ruled out by D1 (and *a fortiori* ruled out by NWBR). Both of these possibilities can happen in more general settings; our goal here is to illustrate the logic of JCE in a simple and economically sensible setting.

EXAMPLE 1:

| $\theta_H$ | $e_H$ | $e_M$ | $e_L$ |
|---|---|---|---|
| $Hire$ | $16, 2$ | $1, 0$ | $-2, -1$ |
| $Pass$ | $0, 0$ | $0, 0$ | $0, 0$ |

| $\theta_M$ | $e_H$ | $e_M$ | $e_L$ |
|---|---|---|---|
| $Hire$ | $8, 0$ | $6, 1$ | $-4, 0$ |
| $Pass$ | $0, 0$ | $0, 0$ | $0, 0$ |

| $\theta_L$ | $e_H$ | $e_M$ | $e_L$ |
|---|---|---|---|
| $Hire$ | $4, -1$ | $1, 0$ | $-1, 1$ |
| $Pass$ | $0, 0$ | $0, 0$ | $0, 0$ |

In this example, a hired worker wishes to adjust their costly effort with the quality of the firm because the worker gains when the firm does well, and firm quality and worker effort are complements in determining the likelihood of success. Moreover, the return to effort varies with type so much that the intermediate effort level is strictly dominated when probability of the intermediate type $\theta_M$ is small.

All firm types have the same ordinal ranking over outcomes, $(Hire, e_H) \succ (Hire, e_M) \succ Pass \succ (Hire, e_L)$, but they do not have the same ranking of outcome distributions. For instance, there are mixtures over $e_H$ and $e_L$ that make $\theta_H$ strictly prefer to $Hire$ and $\theta_L$ strictly prefer to $Pass$, while there are mixtures over $e_M$ and $e_L$ that make $\theta_L$ strictly prefer to $Hire$ while $\theta_H$ strictly prefers to $Pass$. For motivation, suppose that, relative to the low quality $\theta_L$ firm, the high quality $\theta_H$ firm can very efficiently capitalize on a worker exerting high effort, but does not benefit much from medium effort, and is harmed by a poor worker exerting low effort. Similarly, there are mixtures over $e_M$ and $e_L$ that make $\theta_M$ strictly prefer to $Hire$ and $\theta_L$ strictly prefer to $Pass$, while there are mixtures

[7]The conclusions in these examples do not depend on the distribution over types, so we omit $\lambda$.

over $e_H$ and $e_L$ that make $\theta_L$ strictly prefer to $Hire$ and $\theta_M$ strictly prefer to $Pass$. This is the case when, relative to the low quality firm, the medium-quality firm gains significantly from medium effort, does not gain much from high effort, and is greatly hurt by low effort.

In every JCE there is a positive probability that the worker is hired. To see why, consider a strategy profile $\pi$ in which $\pi_1(Hire|\theta) = 0$ for all $\theta$. Observe that when $\alpha \in \Delta(\{e_H, e_M, e_L\})$ satisfies $4\alpha(e_H) + \alpha(e_M) - \alpha(e_L) \geq 0$, either $16\alpha(e_H) + \alpha(e_M) - 2\alpha(e_L) > 0$ or $8\alpha(e_H) + 6\alpha(e_M) - 4\alpha(e_L) > 0$, so either $\theta_H$ or $\theta_M$ strictly prefers $Hire$ whenever $\theta_L$ weakly prefers $Hire$. Thus $\widetilde{D}_{\theta_L}(Hire, \pi) \cup \widetilde{D}^0_{\theta_L}(Hire, \pi) \subseteq \widetilde{D}_{\theta_H}(Hire, \pi) \cup \widetilde{D}_{\theta_M}(Hire, \pi)$, so $\theta_L \notin \Theta^\dagger(Hire, \pi)$. Moreover, $\Theta^\dagger(Hire, \pi)$ is not empty, because some effort distributions make $\theta_H$ prefer $Hire$ and the other types prefer $Pass$, so $\theta_L$ is not a justified type.[8] Since it is optimal for a hired worker to play $e_L$ only when they put positive probability on the firm being $\theta_L$, no justified response can use $e_L$ with positive probability, so all firm types strictly prefer to $Hire$.

Unlike JCE, D1 and weaker equilibrium refinements such as the Intuitive Criterion allow equilibria in which all types $Pass$.[9] The Intuitive Criterion allows this equilibrium because $\theta_L$ would obtain a higher payoff from $Hire$ if the worker responds with either $e_H$ or $e_M$, both of which are undominated. Consequently, the Intuitive Criterion allows the worker respond to $Hire$ with $e_L$, since it is the best response to $\theta_L$. Similarly, D1 allows the worker to respond to $Hire$ with $e_L$, because there is no single type that strictly prefers to $Hire$ whenever $\theta_L$ weakly prefers to do so. In particular, $\theta_L$ strictly prefers to play $Hire$ when the worker responds with $(1/7)e_H + (6/7)e_L$, though this makes $\theta_M$ strictly prefer to $Pass$. Likewise, $\theta_L$ strictly prefers to $Hire$ when the worker responds with $(2/5)e_M + (3/5)e_L$, though this makes $\theta_H$ strictly prefer to $Pass$. $\square$

Example 1 shows that there are some sensible economic environments where JCE makes stronger predictions than D1.[10] The reverse can also be true, as shown in the following example, where JCE allows an outcome that D1 and the stronger NWBR condition rule out. Because JCE, unlike D1 or NWBR, has a learning foundation, this highlights the subtlety of the implications of learning foundations for equilibrium play.

EXAMPLE 2:

As before, a hired worker wishes to exert high effort when hired by a high quality firm and low effort when hired by a low quality firm. However, here a hired worker also wishes to exert high effort when hired by a medium quality firm. Moreover, there is no belief over the sender's type that makes both high and low

[8] In fact, the set of justified types for $Hire$ given $\pi$ is $\overline{\Theta}(Hire, \pi) = \Theta^\dagger(Hire, \pi) = \{\theta_H, \theta_M\}$.
[9] Section III reviews the formal definitions of the Intuitive Criterion and D1.
[10] OA.7.2 in the Online Appendix provides a qualitatively different example concerning job assignment and corporate culture where JCE is again stronger than D1.

| $\theta_H$ | $e_H$ | $e_M$ | $e_L$ |
|---|---|---|---|
| $Hire$ | $16,2$ | $1,0$ | $-5,-3$ |
| $Pass$ | $0,0$ | $0,0$ | $0,0$ |

| $\theta_M$ | $e_H$ | $e_M$ | $e_L$ |
|---|---|---|---|
| $Hire$ | $8,1$ | $6,0$ | $-4,-2$ |
| $Pass$ | $0,0$ | $0,0$ | $0,0$ |

| $\theta_L$ | $e_H$ | $e_M$ | $e_L$ |
|---|---|---|---|
| $Hire$ | $4,-8$ | $1,0$ | $-1,4$ |
| $Pass$ | $0,0$ | $0,0$ | $0,0$ |

effort levels best responses due to concavity in the worker's payoff.

The firm's payoffs are similar to Example 1, except here the payoff of the $\theta_H$ firm from a worker exerting low effort is reduced. This guarantees that, whenever a type $\theta_H$ or $\theta_M$ firm weakly prefers to $Hire$ a worker whose effort concentrates on $e_H$ and $e_L$, type $\theta_L$ strictly prefers to $Hire$.

Every type playing $Pass$ is both a PBE and a JCE outcome. It is a PBE outcome because all types are deterred from playing $Hire$ when the receiver responds with $e_L = BR(\theta_L, Hire)$. Moreover, $\theta_L$ is a justified type for $Hire$ under a strategy profile $\pi$ where all types pass since there are effort distributions which make $\theta_L$ prefer to play $Hire$ and the other types prefer to $Pass$. Thus, $e_L$ is a justified response to $Hire$, so it is a JCE outcome for every type to $Pass$.

However, every type playing $Pass$ is not a D1 outcome. This is because no mixed best response to $Hire$ puts positive probability on both $e_H$ and $e_L$. Consequently, every mixed best response that makes $\theta_L$ weakly prefer to $Hire$, makes $\theta_M$ strictly prefer to do so. Likewise, for type $\theta_H$. The only response to $Hire$ allowed by D1 is then $e_H = BR(\theta_M, Hire)$, which deters no type from hiring. □

Both Examples 1 and 2 use the setting of a firm hiring a worker, but the point that JCE and D1 are not nested holds more generally, as we explain in Section III.

## II. The Learning Model

### A. Model Overview

Now we sketch the structure of the learning model we use to provide a foundation for JCE, and then explain why the learning model generates the predictions we saw in the previous examples. (Later subsections provide the remaining details and formal results of the model, as well as some alternative models with the same implications.) The model is an overlapping generations learning environment where time is discrete and doubly infinite, $t \in \{..., -2, -1, 0, 1, 2, ...\}$. For for each $\theta$, 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, and then 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 $s$ and a message $m$. The receiver observes the chosen $(s, m)$ and responds with an action $a$. 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.

All agents are rational Bayesians who choose *policies* (maps from past observations to current play) that maximize their expected discounted payoff. At every period $t$, the *state* of the system is the shares of agents in a given player role with the various possible histories. The state and the optimal policies induce an *aggregate sender strategy* and an *aggregate receiver strategy*, and thus an *update rule* that maps states in period $t$ to states in period $t+1$. We study this system's steady states, which are the fixed points of the update rule.

Because the receivers observe the type of the sender at the end of each match, neither their continuation probability nor their discount factor impacts their play, and their optimal dynamic programming policy is to simply choose an action that maximizes their expected payoff in the current match. Senders' observations do depend on their play, so their optimal policies incorporate a value for "experimenting" 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, their discount factor $\delta \in [0, 1)$, and how much they have already learned: Inexperienced senders have more incentive to experiment, and senders cease experimenting when they have enough data. Moreover, different types of sender will choose to experiment in different ways.

We focus on the limits 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, and $\gamma_1$ tends to 1 quicker than $\gamma_2$, so that the typical sender plays many more times than the typical receiver. This means that most receivers only ever match with senders who have substantially more experience than them. We also assume that $\delta$ goes to 1, to ensure that the senders experiment enough to rule out limits that are not Nash equilibria. We call the profiles that correspond to this limit the *stable profiles*.[11] This limit provides an idealized version of long-run behavior in 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 workers via offers of incentive pay. While workers 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.

Preliminary lemmas show that every stable profile must be a PBE-H. The

[11] As Section V.A explains, our results hold under other models of the population structure that also have relatively experienced senders.

optimality of sender play follows from the fact that patient and long-lived senders eventually stop experimenting and play a best response to the aggregate receiver strategy. The beliefs of long-lived receivers are almost entirely driven by their data as opposed to their priors, ensuring that they respond optimally to on-path signal-message pairs, and because receivers are myopic, their off-path play is always a best response to some beliefs, as PBE-H requires.

Our main result, which shows that all stable profiles are JCE, uses two additional assumptions. First, we assume that the sender message space is sufficiently large that for each signal $s \in S$ and subset of sender types $\widetilde{\Theta} \subseteq \Theta$, there is a distinct message $m_{s,\widetilde{\Theta}}$ that can be interpreted as "I am playing $s$ and my type is in $\widetilde{\Theta}$."[12] We also assume that the receiver "trusts" the message provided that they have not previously encountered a sender with any other type $\theta \notin \widetilde{\Theta}$ playing $s$ and sending $m_{s,\widetilde{\Theta}}$. We discuss these assumptions in more detail in Section II.D. With them we prove the following result:

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

### B. *Hiring a Worker, Revisited*

We now discuss our learning model in the context of the two "hiring a worker" examples. In particular, we explain why the model rules out the "All $Pass$" outcome in Example 1, where it is consistent with D1 but not JCE, while the model does allow "All $Pass$" in Example 2, where it is consistent with JCE but not D1.

EXAMPLE 1 CONTINUED:

To see why there is no stable profile where all firms Pass, recall that with enough experience, firms learn the aggregate effort distribution and exhaust the option value of continued experimentation. Experienced firms then either hire and optimally communicate with workers or $Pass$. If the stable outcome is for all types of firm to $Pass$, it must be that the aggregate effort distribution puts positive probability on effort $e_L$ regardless of how a hiring firm communicates. However, low effort is only optimal for a worker if they assign positive probability to the hiring firm being type $\theta_L$. Initially-trusting workers will exert high or low effort when a hiring firm claims to not be type $\theta_L$, unless they have previously experienced deception by type $\theta_L$ firms. Since the typical firm has many more interactions over its lifetime than the typical worker, most workers only ever match with experienced firms. Thus, in order for a significant share of workers to experience deception by $\theta_L$ firms, $\theta_L$ firms must learn that it is optimal to $Hire$ and play $m_{In,\{\theta_H,\theta_M\}}$. However, either $\theta_H$ or $\theta_M$ type firms strictly prefer to $Hire$ whenever a $\theta_L$ weakly prefers to $Hire$, so one of these types will not $Pass$ because it will learn it is strictly optimal to $Hire$. □

[12] The literal content of $m_{s,\widetilde{\Theta}}$ need not be "I am playing $s$ and my type is in $\widetilde{\Theta}$." Instead, $m_{s,\widetilde{\Theta}}$ might be a statement like "I am playing signal $s$ so you should believe my type is in $\widetilde{\Theta}$ because..."

EXAMPLE 2 CONTINUED:

Online Appendix Section OA.6.1 shows that the outcome where every type plays $Pass$ is stable by demonstrating that there are steady-state profiles in which most sufficiently experienced firms play $Pass$, but type $\theta_L$ firms experiment with $Hire$ much longer than the other types do. This means that the vast majority of workers who have previously been employed were only hired by low quality firms, which leads them to exert effort $e_L$ the next time they are hired. In contrast, initially-trusting workers who have not previously experienced employment will exert high effort when first hired by a firm claiming to be of high or medium quality, so the aggregate worker effort distributions will concentrate on $e_H$ and $e_L$. As observed earlier, under such effort distributions, whenever type $\theta_H$ or $\theta_M$ weakly prefers to $Hire$, type $\theta_L$ strictly prefers to do so. This is what drives type $\theta_L$ firms to experiment with $Hire$ much more than the other types, which supports the desired steady states.

D1 eliminates the "All $Pass$" outcome while our learning model allows it because D1 only considers receiver mixed best responses. However, in a learning model, there is no reason that the prevailing aggregate receiver strategy must be a mixed best response, and the steady states described above have aggregate worker responses that put positive probability on both both $e_H$ and $e_L$: Inexperienced workers exert effort $e_H$, while most of the experienced workers learn that it is optimal to exert effort $e_L$. $\square$

### C. *Details of the Learning Model*

We now fill in the remaining details about the learning environment we study, provide formal statements of our assumptions, and prove our main result. We also discuss alternative interpretations of the stable profiles, and related versions of stability that correspond to different ways of passing to the limit. Readers who are more interested in the implications of JCE than its learning foundation can skip ahead to Section III.

At the beginning of their lives, senders have a non-doctrinaire prior $g_1 \in \Delta(\Pi_2)$ over the aggregate receiver behavior strategy $\pi_2$, and receivers have a non-doctrinaire prior $g_2 \in \Delta(\Delta(\Theta \times S \times M))$ over $q \in \Delta(\Theta \times S \times M)$, the prevailing distribution of sender types, signals, and messages.[13] (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

[13]Here "non-doctrinaire" means "has a continuous density function that is strictly positive on the interior of the probability simplex." Since $q(\theta, s, m) = \lambda(\theta)\pi_1(s, m|\theta)$ is the distribution over $(\theta, s, m)$ induced by the sender type distribution $\lambda$ and the aggregate sender behavior strategy $\pi_1$, it would be equivalent to define the receivers' beliefs as elements of $\Delta(\Delta(\Theta) \times \Pi_1)$.

to any finite sequence of observations. Define $\mathcal{H}_{1,t} = (S \times M \times A)^t$ to be the histories that a sender of age $t$ could have observed, with the convention that $(S \times M \times A)^0 = \emptyset$, and let $\mathcal{H}_1 = \cup_{t\in\mathbb{N}}\mathcal{H}_{1,t}$ 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 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 maximize their expected discounted payoff. The receivers use a policy $\mathbf{y} : \mathcal{H}_2 \to A^{S\times M}$ which maps their histories to pure strategies to maximize

$$\mathbb{E}_{g_2}\left[\sum_{t=0}^{\infty}\sum_{\theta,s,m,a}\gamma_2^t q(\theta,s,m)u_2(\theta,s,\mathbf{y}(s,m|h_{2,t}))\right].$$

Type $\theta$ senders use an optimal policy $\mathbf{x}_\theta^{\delta,\gamma_1} : \mathcal{H}_1 \to S\times M$ that maps their histories to signal-message pairs to maximize

$$\mathbb{E}_{g_1}\left[\sum_{t=0}^{\infty}\sum_{s,m,a}(\delta\gamma_1)^t\pi_2(a|\mathbf{x}_\theta^{\delta,\gamma_1}(h_{1,t}))u_1(\theta,\mathbf{x}_\theta^{\delta,\gamma_1}(h_{1,t}),a)\right].$$[14]

We will focus on the case where both $\delta$ and $\gamma_1$ are near 1, so the senders have maximal incentives to experiment.[15]

At every period $t$, the *state* of the system, denoted $\mu_t = (\mu_{1,t}, \mu_{2,t}) \in (\Delta(\mathcal{H}_1))^\Theta \times \Delta(\mathcal{H}_2)$, gives the shares of agents in a given player role with the various possible histories. Given $\mu_t$, the profile $\mathbf{x}^{\delta,\gamma_1} = \{\mathbf{x}_\theta^{\delta,\gamma_1}\}_{\theta\in\Theta}$ of sender policies induces a sender behavior strategy $\sigma_1^{\delta,\gamma_1}(\mu_{1,t}) \in \Pi_1$ 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 $\sigma^{\delta,\gamma_1}(\mu_t) = (\sigma_1^{\delta,\gamma_1}(\mu_{1,t}), \sigma_2(\mu_{2,t})) \in \Pi_1 \times \Pi_2$ the *aggregate strategy profile.* (Appendix D gives formal definitions of the mappings $\sigma_1^{\delta,\gamma_1} : (\Delta(\mathcal{H}_1))^\Theta \to \Pi_1$ 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* $\mathbf{f}^{\delta,\gamma_1,\gamma_2} : (\Delta(\mathcal{H}_1))^\Theta \times \Delta(\mathcal{H}_2) \to (\Delta(\mathcal{H}_1))^\Theta \times \Delta(\mathcal{H}_2)$, taking the state in period $t$ to the state in period $t+1$, a mapping $\mathscr{R}_1^{\delta,\gamma_1} : \Pi_2 \to \Pi_1$ that describes the limit of the aggregate play of the senders as $t \to \infty$ when the aggregate receiver play is fixed at $\pi_2$, and a mapping $\mathscr{R}_2^{\gamma_2} : \Pi_1 \to \Pi_2$ that describes the limit of the aggregate receiver play as $t \to \infty$ when the aggregate sender play is fixed at $\pi_1$. We refer to the mapping $\mathscr{R}^{\delta,\gamma_1,\gamma_2}(\pi) \equiv (\mathscr{R}_1^{\delta,\gamma_1}(\pi_2), \mathscr{R}_2^{\gamma_2}(\pi_1))$ as the *aggregate response mapping.* OA.2.1 verifies that this mapping is continuous.

[14] Here we slightly abuse notation by having both components of $\mathbf{x}_\theta^{\delta,\gamma_1}$ enter the utility function, though it does not depend on the sender's message.

[15] Recall from Fudenberg and Kreps (1988) and Fudenberg and Levine (1993) that with impatient players learning need not lead to Nash equilibrium, let alone to refinements of it.

We study this system's steady states, those $\mu$ satisfying $\mathbf{f}^{\delta,\gamma_1,\gamma_2}(\mu) = \mu$. We call the corresponding aggregate strategy profiles the *steady-state profiles*, and denote them by $\Pi^*(g, \delta, \gamma_1, \gamma_2) \subseteq \Pi_1 \times \Pi_2$. As OA.2.2 shows, these are the fixed points of the aggregate response mapping. Continuity of the aggregate response mapping, along with Brouwer's fixed point theorem, then implies that steady-state profiles always exist.

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* $\gamma_1, \gamma_2$.

We consider 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 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*.[16] A corollary of Proposition 1 is that there are stable strategy profiles.

COROLLARY 1: *Stable strategy profiles exist.*

### D. 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 requires that the message space is large enough to have a distinct element, $m_{s,\widetilde{\Theta}} \in M$, for each signal $s \in S$ and subset of sender types $\widetilde{\Theta} \subseteq \Theta$. This allows $m_{s,\widetilde{\Theta}}$ to be interpreted as "I am playing $s$ and my type is in $\widetilde{\Theta}$." 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 a sender with any other type $\theta \notin \widetilde{\Theta}$ playing $s$ and sending $m_{s,\widetilde{\Theta}}$.

ASSUMPTION 2: **(Initial Trust)** *For every* $s \in S$ *and* $\widetilde{\Theta} \subseteq \Theta$, *there is some* $m_{s,\widetilde{\Theta}} \in M$ *such that* $\mathbf{y}(s, m_{s,\widetilde{\Theta}} | h_2) \in BR(\widetilde{\Theta}, s)$ *for every* $h_2 \in \mathcal{H}_2$ *in which* $(\theta', s, m_{s,\widetilde{\Theta}})$ *has not been observed for any* $\theta' \notin \widetilde{\Theta}$.

Initial trust 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.[17] It does not require that the receivers are certain that these claims are true, only

[16] Formally, strategy profile $\pi$ 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 $\lim_{k\to\infty} \delta_{j,k} = 1$ 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 $\pi_{j,k,l} \in \Pi^*(g, \delta_{1,j,k}, \gamma_{1,j,k,l}, \gamma_{2,j})$.

[17] 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 V.C.

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.[18]

We maintain Assumptions 1 and 2 throughout the main text. Section V.D discusses alternatives to initial trust that give refinements of PBE-H that are similar to JCE.

Without communication, or with communication but no assumptions on the receivers' prior beliefs about the meaning of previously unobserved messages, stability offers little predictive power and the theorem is false. In particular, it then allows implausible outcomes, as shown by example in OA.7.1. 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 $Out$. Indeed, this is the unique JCE as well as the unique equilibrium outcome that satisfies weaker refinements such as the Intuitive Criterion. However, if the receivers are "initially skeptical" so that when they first witness $(In, m)$ they believe it probably came from $\theta_2$ regardless of $m$, there are stable profiles in which both types play $Out$. This is because, if very few senders play $In$, the aggregate receiver response to $In$ concentrates on the best response to $\theta_2$, which ensures that almost all senders in the population learn that it is optimal to play $Out$.[19] Intuitively, cheap talk has no effect in "babbling" equilibria where messages are meaningless, and effective communication requires some restrictions on how people interpret messages they have never seen before.

### *E. Proof of Theorem 1*

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

To prove this theorem we first show that a stable profile is a PBE-H. Condition 3 of the definition of PBE-H 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 of JCE given in Definition 3, using three supporting lemmas, whose proofs are in Appendix B.

The following lemma shows that stable profiles satisfy Condition 1 of Definition 1.

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 $\pi_2$.*

[18] Initial trust implicitly places restrictions on the receivers' prior $g_2$. For simplicity, we state it directly on receiver behavior.

[19] The same argument shows that this outcome is also stable when cheap talk is not feasible.

The proof of Lemma 1 shows that for fixed $\gamma_2 \in [0, 1)$, aggregate sender play is optimal given the aggregate receiver play when first $\gamma_1 \to 1$ and 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 correspondence (in the underlying two-player game) has a closed graph.

The next lemma shows that stable profiles satisfy Condition 2 of Definition 1.

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 $\{\pi_1(s, m|\theta)\}_{\theta\in\Theta}$.*

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. 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 $\pi$ 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* in settings where cheap-talk messages are not available. Moreover, the example in OA.7.1 shows that without initial trust, there can be stable profiles that are not JCE.

LEMMA 3: *Suppose that $\pi$ 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 $\pi_2(BR(\widetilde{\Theta}, s)|s, m_{s,\widetilde{\Theta}}) = 1$.*

Here we give some intuition for this result. When $u_1(\theta, s, \pi_2(\cdot|s, m_{s,\widetilde{\Theta}})) < u_1(\theta, \pi)$ for all $\theta \notin \widetilde{\Theta}$, the proof of Lemma 1 shows that, for fixed $\gamma_2$, the aggregate probability that a type outside of $\widetilde{\Theta}$ plays $(s, m_{s,\widetilde{\Theta}})$ is small when first $\gamma_1 \to 1$ and then $\delta \to 1$. For any fixed receiver continuation probability, 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.

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 justified response condition in Definition 3 holds. Fix $s \in S$. Since $\pi$ is a PBE-H, $u_1(\theta, s, \pi_2(\cdot|s, m_{s,\overline{\Theta}(s,\pi)})) \leq u_1(\theta, \pi)$ holds for all $\theta \in \overline{\Theta}(s, \pi)$. By definition, whenever a receiver response weakly deters all justified types from playing a given signal, it must strictly deter every non-justified type. Thus, $u_1(\theta, s, \pi_2(\cdot|s, m_{s,\overline{\Theta}(s,\pi)})) < u_1(\theta, \pi)$ for all $\theta \notin \overline{\Theta}(s, \pi)$. Applying Lemma 3 to $\widetilde{\Theta} = \overline{\Theta}(s, \pi)$ 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. Not all JCE are stable, because non-doctrinaire priors prevent receiver agents from ever using weakly dominated strategies, and there can be JCE using weakly dominated receiver strategies. It is difficult to give an exact characterization of stable profiles for general games, because all non-doctrinaire initially-trusting priors must be considered to show that a given profile is not stable. Instead, we use direct proofs to show that certain equilibria or classes of equilibria are stable. Proposition C1 in Appendix C gives a partial converse to Theorem 1: It shows that all *uniformly justified* JCE in *strictly monotonic* games are stable for all non-doctrinaire priors, including those that do not satisfy initial trust. We also give direct proofs of stability in Example 2 and most of our other examples. The general approach in these proofs is to modify the aggregate response mapping so that its fixed points coincide with the target strategy profile in the limit, and then show that these fixed points are also fixed points of the true aggregate response mapping.

## III. Relation to Other Equilibrium Refinements

We have seen by example that JCE and D1 are not nested. We now study their relationship in more detail, as well as the relationship between JCE and other refinements. As a preliminary step, we show that stable profiles need not be PBE, and *a fortiori* need not satisfy any refinements of PBE.[20] This is the reason that JCE is defined as a refinement of PBE-H.

EXAMPLE 3:

The type space is $\Theta = \{\theta_1, \theta_2\}$, the signal space is $S = \{In, Out\}$, and the action space is $A = \{a_1, a_2, a_3\}$. The payoffs are given by these tables:

| $\theta_1$ | $a_1$ | $a_2$ | $a_3$ |
|---|---|---|---|
| $In$ | $-2, 1$ | $1, .1$ | $1, -1$ |
| $Out$ | $0, 0$ | $0, 0$ | $0, 0$ |

| $\theta_2$ | $a_1$ | $a_2$ | $a_3$ |
|---|---|---|---|
| $In$ | $1, -1$ | $1, .1$ | $-2, 1$ |
| $Out$ | $0, 0$ | $0, 0$ | $0, 0$ |

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

This game does not have a PBE in which both types play $Out$, because $\theta_1$ prefers to play $Out$ only if there is positive probability that the receiver responds to $In$ with $a_1$, while $\theta_2$ prefers to play $Out$ only if the receiver's response to $In$ uses $a_3$ with positive probability, yet there is no mixed best response to $In$ where the receiver assigns positive probability to both $a_1$ and $a_3$. Nevertheless, the profile $\pi$ in which both sender types play $Out$ and the receiver always responds to $In$ with $(1/2)a_1 + (1/2)a_3$ is a JCE, because both sender types are justified and so $a_1 = BR(\theta_1, In)$ and $a_3 = BR(\theta_2, In)$ are each justified responses.

Moreover, Online Appendix Section OA.6.2 shows that both types playing $Out$ can be a stable outcome, because there are steady-state profiles in which the aggregate receiver strategy plays $a_2$ with probability less than $1/4$ in response to $In$ combined with any message. Under such receiver play, for every message $m$, it can be optimal for at most one sender type to play $(In, m)$. Thus, if in the limit the aggregate strategy of type $\theta_1$ plays $(In, m)$ with positive probability, then the aggregate strategy of type $\theta_2$ must play $(In, m)$ with 0 probability, and the receivers must learn to respond to $(In, m)$ with $a_1 = BR(\theta_1, In)$. But this response strictly deters type $\theta_1$ from playing $(In, m)$, and an analogous argument applies for the type $\theta_2$ senders. $\square$

Unlike JCE, 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, which lets us more naturally compare the predictions of the modified versions of these refinements with JCE. As we will see, JCE is stronger than the modified version of the Intuitive Criterion. JCE and D1 are not nested, although JCE is nested inside the set of equilibria that satisfy a modified version of D1 we call *co-D1*. JCE and NWBR are particularly similar, and in some sense JCE is an adaptation of NWBR with a learning foundation.

We begin by showing that JCE is stronger than a modified version of the Intuitive Criterion we call the *Intuitive Criterion Test*. Let $E(s, \pi) = \{\theta \in \Theta : \max_{a \in BR(\Theta, s)} u_1(\theta, s, a) \geq u_1(\theta, \pi)\}$. These are the types for whom $s$ is not equilibrium dominated by profile $\pi$ in the sense of Cho and Kreps (1987).

DEFINITION 4 (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 2: *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, i.e. $\overline{\Theta}(s, \pi) \subseteq E(s, \pi)$ when $E(s, \pi) \neq \emptyset$. This implies that if there is a justified response that deters all types from playing $s$, then the profile passes the Intuitive Criterion Test. The proof of Proposition 2 is given in Appendix A1.

To compare JCE with D1 and NWBR, 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 receiver mixed best responses to $s$ that give type $\theta$ strictly more than their equilibrium payoff, and let $D^0_\theta(s,\pi) = \{\alpha \in MBR(\Theta, s) : u_1(\theta, s, \alpha) = u_1(\theta, \pi)\}$ be the mixed best responses that give type $\theta$ their equilibrium payoff. These are the analogs of the sets $\widetilde{D}_\theta(s,\pi)$ and $\widetilde{D}^0_\theta(s,\pi)$ when $\Delta(BR(\Theta, s))$ is replaced by $MBR(\Theta, s)$.

To define D1, let $\Theta^{\ddagger,D1}(s,\pi) = \{\theta \in \Theta : \forall \theta' \neq \theta,\ D_\theta(s,\pi) \cup D^0_\theta(s,\pi) \not\subseteq D_{\theta'}(s,\pi)\}$, and let $\widehat{\Theta}^{D1}(s,\pi) \subseteq \Theta$ be the set of types given by

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

Also, let $MBR(\widetilde{\Theta}, s) = \{\alpha \in \Delta(A) : \exists p \in \Delta(\widetilde{\Theta}) \text{ s.t. } u_2(p,s,\alpha) \geq u_2(p,s,a)\ \forall a \in A\}$ denote the set of mixed best responses to $s$ for beliefs supported on a given $\widetilde{\Theta} \subseteq \Theta$.

DEFINITION 5 (Banks and Sobel, 1987): *Strategy profile $\pi$* **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 $\theta \in \Theta$.*

D1 can be stronger than JCE (and rule out some stable profiles) because it only considers receiver mixed-best responses, both in finding possible responses to off-path signal-message pairs and in the construction of the sets of sender types to which the receiver must be best-responding. As we have seen, however, the larger convex hull of receiver best responses emerges in our learning model rather than the receiver mixed best responses.[21]

To see the difference this makes, 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 $\theta$ 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 $\overline{\Theta}^{D1}(s,\pi) \subseteq \Theta$ be the set

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

DEFINITION 6: *A PBE-H $\pi$ 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 3: *If $\pi$ is a justified communication equilibrium, then $\pi$ is a PBE-H that is co-D1.*

[21] 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, but neither paper showed that restricting attention to the receiver mixed best responses rules out a profile that is stable in a learning model.

Co-D1 is more permissive than JCE because it strikes fewer types. Appendix A2 gives the proof of Proposition 3, which shows that $\overline{\Theta}(s,\pi) \subseteq \overline{\Theta}^{D1}(s,\pi)$ for all $s$; Example 1 shows that the inclusion is sometimes strict.

To define NWBR, let $\Theta^{\ddagger}(s,\pi) = \{\theta \in \Theta : D^0_\theta(s,\pi) \not\subseteq \cup_{\theta' \neq \theta} D_{\theta'}(s,\pi)\}$, which are the $\theta$ for which there is a mixed receiver best response $\alpha \in MBR(\Theta, s)$ that makes $\theta$ indifferent between $(s,\alpha)$ and their equilibrium outcome and every other type weakly prefer their equilibrium outcome to $(s,\alpha)$. Let $\widehat{\Theta}(s,\pi) \subseteq \Theta$ be the set

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

DEFINITION 7 (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 $\alpha \in MBR(\widehat{\Theta}(s,\pi), s)$ 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 $MBR(\widetilde{\Theta}, s)$ were replaced with the convex hulls of best responses $\Delta(BR(\widetilde{\Theta}, s))$. Indeed, as shown in OA.1, it would be equivalent to define JCE by setting $\Theta^{\dagger}(s,\pi) = \{\theta \in \Theta : \widetilde{D}^0_\theta(s,\pi) \not\subseteq \cup_{\theta' \neq \theta} \widetilde{D}_{\theta'}(s,\pi)\}$, rather than $\Theta^{\dagger}(s,\pi) = \{\theta \in \Theta : \widetilde{D}_\theta(s,\pi) \cup \widetilde{D}^0_\theta(s,\pi) \not\subseteq \cup_{\theta' \neq \theta} \widetilde{D}_{\theta'}(s,\pi)\}$. Thus JCE modifies NWBR in much the same way that co-D1 modifies D1, so NWBR is a stronger refinement than JCE.

PROPOSITION 4: *Any PBE-H that satisfies NWBR is path-equivalent to a PBE that is a JCE.*[22]

A PBE-H $\pi$ that satisfies NWBR is not necessarily a PBE, since the receiver's response to off-path play need not be a best reply to any single belief over the sender's type. However, every such profile is path-equivalent to a PBE, since the receiver's response to a given off-path $(s,m)$ can always be replaced by some $\alpha \in MBR(\widehat{\Theta}(s,\pi), s)$ that deters the sender types from playing it. Appendix A3 completes the proof of Proposition 4 by showing that an "NWBR type" is always a justified type. That is, for a given signal and PBE-H, $\widehat{\Theta}(s,\pi) \subseteq \overline{\Theta}(s,\pi)$ for all $s$ and all PBE-H $\pi$.

The converse of Proposition 4 is in general false, as shown earlier by Example 2. However, there are important settings in which NWBR and JCE are path-equivalent. One is when there are at most two undominated receiver responses to each signal, because then mixed best responses and convex hulls of best responses are the same. We now explore a different class of games where this equivalence holds.

[22]Path equivalence is needed in this statement because, unlike JCE, NWBR does not impose requirements about the receiver's actual responses to off-path play.

## IV. Co-Monotonic Signaling Games

This section highlights an important class of commonly studied games in which JCE, D1, and NWBR are path-equivalent, so the learning foundation for JCE applies to D1 and NWBR as well. In these *co-monotonic* signaling games, all sender types share the same preference over mixtures over $BR(\Theta, s)$.

DEFINITION 8: *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.

A sufficient condition 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 $\psi : \Theta \times S \to \mathbb{R}$ 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 4:

Like Examples 1 and 2, this example concerns a firm hiring a worker, except here the firm offers incentive pay to their prospective employee. The firm is better informed about the productivity of the worker's effort; their information is represented by their type $\theta \in \Theta = \{1, 2, 3\}$, with each type equally likely. The firm's signal $s = (s_1, s_2) \in \{0, 1/4, 1/2, 3/4, 1\} \times \{0, 1, 2, ..., 100\}$ consists of a share of profits $s_1$ and a base wage $s_2$ which the worker is offered, and the action $a \in \{0, 5, 10, ..., 60\}$ represents the worker's choice of effort level. The expected profit given the firm's type $\theta$ and the worker's effort $a$ is $\theta a$. Thus, the payoffs to the sender and receiver are $u_1(\theta, s, a) = \theta(1 - s_1)a - s_2$ and $u_2(\theta, s, a) = \theta s_1 a + s_2 - a^2/40$, which satisfy the sufficient condition for co-monotonic signaling games given above. OA.6.3 in the Online Appendix shows that JCE selects 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 that $u_1(\theta, s, \alpha) = u_1(\theta, s, \alpha')$ for all $\theta$. This ensures that $\overline{\Theta}(s, \pi) = \widehat{\Theta}(s, \pi)$ for every PBE-H $\pi$.

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

The proof of Lemma 4 is in Appendix A3.

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 7) select the same profiles up to path-equivalence. JCE thus provides a learning foundation for the predictions of NWBR in the class of co-monotonic games.

PROPOSITION 5: *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:

Suppose that $\pi$ is a PBE-H that satisfies NWBR. Then, by Proposition 4, $\pi$ 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 $a_s \in BR(\widehat{\Theta}(s,\pi), s)$, so $\pi$ is a PBE-H that satisfies NWBR. ■

Combining Proposition 5 with the observation that every PBE-H that satisfies NWBR is path-equivalent to a PBE shows 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, and every PBE-H that satisfies NWBR or D1 is path-equivalent to a justified communication equilibrium.*

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).[23]

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 5 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

[23]Technically, the game analyzed in DeMarzo, Kremer and Skrzypacz (2005) is not a traditional signaling game because of the presence of multiple senders, but this distinction is not important.

to the Riley outcome when the action space is a sufficiently fine finite grid, as in Example 4.[24]

## V. Discussion

### A. Alternate Models

The key to our analysis is that we consider a limit where most senders and receivers have substantial experience, but typical senders have significantly more experience, so that most receivers never encounter inexperienced senders. Because it is inexperienced senders who are the most likely to "experiment" with signal-message pairs that depart from the limit strategy profile, most receivers have little experience with off-path play by the senders, which facilitates the analysis of the stable profiles.

We can obtain this situation with many different specifications of the populations of agents and how they interact. For example, suppose that senders and receivers have geometric lifetimes with common continuation probability $\gamma \in [0,1)$, so that they all have expected lifetime $T = 1/(1-\gamma)$. Every period, each sender is matched with a receiver, but each receiver only gets matched with some i.i.d. probability $p \in (0,1)$.[25] A given receiver is expected to have $N_2 = pT$ matches over their lifetime, while a sender is expected to have $N_1 = T$ matches. For every steady state in this alternate model, there is a corresponding steady state in our main model with the same aggregate strategy profile when the receiver's continuation probability is $\tilde{\gamma}_2 = (1-1/T)N_2/(1+(1-1/T)N_2)$, which we demonstrate in Online Appendix Section OA.10. Since $\tilde{\gamma}_2 \to N_2/(1+N_2) \in [0,1)$ as $T \to \infty$ for any fixed $N_2 \in \mathbb{R}_+$ and $N_2/(1+N_2) \to 1$ as $N_2 \to \infty$, the iterated limit where first $T \to \infty$ (so that both sender and receiver agents become long-lived) then $\delta \to 1$ (so that sender agents become patient) then $N_2 \to \infty$ (so that receiver agents become experienced) generates precisely the same predictions as our notion of stability.

Moreover, we can also obtain the same set of stable profiles in models where agents do not have geometric lifetimes: To illustrate, suppose that agents have deterministic lifetimes, and that sender agents are matched every period, while receiver agents are matched every $K$ periods during their life. Suppose that sender agents are involved in $N_1$ matches over the course of their lifetime, while receiver agents are involved in $N_2$. Focusing on the profiles that emerge in the limit where first $N_1 \to \infty$ then $\delta \to 1$ then $N_2 \to \infty$ generates exactly the same predictions as stability in the geometric lifetime models. Thus, the unequal lifetimes of our baseline model are simply a modeling convenience, and not an essential feature.

However, we do need some sort of asymmetry in the interaction structure to

[24] As noted by e.g. Fudenberg and Tirole (1991*a*), 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 the iterated limit.

[25] Correspondingly set the population mass of senders to be $p$ times that of the receivers.

derive our results. When both populations have the same expected number of interactions, we have not been able to derive interesting restrictions on the stable profiles. When the receivers have many more interactions (as would be the case in our baseline model in the limit where first $\gamma_2 \to 1$ then $\gamma_1 \to 1$), all stable profiles must be PBE and not only PBE-H, but we do not know whether stability has additional implications.

### B. *Orders of Limits*

We require senders to become long-lived ($\gamma_1 \to 1$) before they become patient ($\delta \to 1$) so that near the limit very few senders will choose to experiment. (This is why we need the model to have a discount factor parameter.) As in Fudenberg and Levine (1993, 2006) and Fudenberg and He (2018, 2020, 2021), it seems difficult to establish in general what happens when a player's patience level $\delta$ goes to 1 before their lifetime become long, as in this case we do not know how to show that most players stop experimenting. 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. 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 $(\delta, \gamma_2) \to (1, 1)$. Moreover, OA.9 shows that Theorem 1's conclusion applies under this general limit to all stable profiles satisfying an additional condition, such as on-path strict incentives for the receiver.

### C. *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.[26] Fudenberg and Levine (2006) used 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

[26] Kalai and Lehrer (1993), Lehrer and Solan (2007), Esponda (2013), Battigalli et al. (2019) studied rational learning without assuming that agents are patient. Battigalli (1987), Rubinstein and Wolinsky (1994), Dekel, Fudenberg and Levine (1999), Esponda (2013), Battigalli et al. (2015), and Fudenberg and Kamada (2015, 2018), among other papers, studied equilibrium concepts motivated by rational learning without analyzing an explicit learning model, and 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.

policy is given by the Gittins index (Gittins, 1979), and used this to derive restrictions on equilibria. Fudenberg and He (2020) 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). If we treat the signal-message pair $(s, m)$ as a signal when evaluating the conditions of RCE, then RCE collapses to the Intuitive Criterion when the message space is not singleton because messages have no effect on payoffs. If we instead compare RCE to JCE in a game with a singleton message space, RCE is again weaker. For example, the "All *Pass*" outcome of Example 1 is consistent with RCE but not JCE.[27] Moreover, OA.3 shows that every JCE is an RCE, because types that are "less compatible" with a given signal in the sense of RCE can never be justified. This paper obtains a stronger refinement than RCE 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 "credible" messages in signaling games.[28] 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 yields more insight than imposing the restrictions directly.

### D. *Extensions*

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 (2020), then receivers who are long-lived may feel that they have acquired a good sense of each sender type's equilibrium payoff. In 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. OA.8.2 shows that we can capture an iterated procedure similar to that of *divinity* and *universal divinity* (Banks and Sobel, 1987) by strengthening initial trust: If the only types

[27] RCE also permits equilibria ruled out by JCE and D1 in co-monotonic games like Example 4.

[28] 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), Chen, Kartik and Sobel (2008), and Gordon et al. (2021) also studied refinements in cheap-talk games.

who have lied about being in $\widetilde{\Theta}$ are elements of $\widetilde{\Theta}'$, then the receiver responds to a claim of $\widetilde{\Theta}$ as if the sender's type belongs to $\widetilde{\Theta} \cup \widetilde{\Theta}'$.

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. (See the papers surveyed in Abeler, Nosenzo and Raymond (2019).) Kartik, Ottaviani and Squintani (2006) and Kartik (2009) incorporated messages with such lying costs into models of strategic communication. OA.8.3 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. Developing learning foundations for refinements in these games is a promising area for future research, and could lead to learning-based refinements for settings with cheap talk and multiple audiences, as in Goltsman and Pavlov (2011).

## VI. Conclusion

Adding cheap-talk communication to signaling games let us provide a learning-theoretic foundation for the concept of justified communication equilibrium. We recovered some of the intuitions that underlie traditional equilibrium refinements for signaling games, whose predictions were by and large sensible in the games where they were used. We also 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.[29]

Of course, there are multiple ways to 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, which may make it more appealing to use.

## References

**Abeler, Johannes, Daniele Nosenzo, and Collin Raymond.** 2019. "Preferences for Truth-Telling." *Econometrica*, 87(4): 1115–1153.

[29] The similarity of stable profiles and those that survive NWBR does not hold in general outside of signaling games; see Footnote 10 of Fudenberg and Levine (1993).

**Banks, Jeffrey S., and Joel Sobel.** 1987. "Equilibrium Selection in Signaling Games." *Econometrica*, 55(3): 647–661.

**Battigalli, Pierpaolo.** 1987. "Comportamento razionale ed equilibrio nei giochi e nelle situazioni sociali." http://didattica.unibocconi.it/mypage/upload/48808_20120403_091802_1987THESIS.PDF.

**Battigalli, Pierpaolo, Alejandro Francetich, Giacomo Lanzani, and Massimo Marinacci.** 2019. "Learning and Self-Confirming Long-Run Biases." *Journal of Economic Theory*, 183: 740–785.

**Battigalli, Pierpaolo, Simone Cerreia-Vioglio, Fabio Maccheroni, and Massimo Marinacci.** 2015. "Self-Confirming Equilibrium and Model Uncertainty." *American Economic Review*, 105(2): 646–677.

**Binmore, Ken, and Larry Samuelson.** 1999. "Evolutionary Drift and Equilibrium Selection." *Review of Economic Studies*, 66(2): 363–393.

**Blume, Andreas, and Joel Sobel.** 1995. "Communication-Proof Equilibria in Cheap-Talk Games." *Journal of Economic Theory*, 65(2): 359–382.

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

**Cho, In-Koo, and David M. Kreps.** 1987. "Signaling Games and Stable Equilibria." *Quarterly Journal of Economics*, 102(2): 179–221.

**Cho, In-Koo, and Joel Sobel.** 1990. "Strategic Stability and Uniqueness in Signaling Games." *Journal of Economic Theory*, 50(2): 381–413.

**Clark, Daniel.** 2020. "Robust Neologism Proofness in Signaling Games." http://economics.mit.edu/files/16758.

**Dekel, Eddie, Drew Fudenberg, and David K. Levine.** 1999. "Payoff Information and Self-Confirming Equilibrium." *Journal of Economic Theory*, 89(2): 165–185.

**DeMarzo, Peter, and Darrell Duffie.** 1999. "A Liquidity-Based Model of Security Design." *Econometrica*, 67(1): 65–99.

**DeMarzo, Peter M., Ilan Kremer, and Andrzej Skrzypacz.** 2005. "Bidding with Securities: Auctions and Security Design." *American Economic Review*, 95(4): 936–959.

**Diaconis, Persi, and David Freedman.** 1990. "On the Uniform Consistency of Bayes Estimates for Multinomial Probabilities." *The Annals of Statistics*, 18(3): 1317–1327.

**Esponda, Ignacio.** 2013. "Rationalizable Conjectural Equilibrium: A Framework for Robust Predictions." *Theoretical Economics*, 8(2): 467–501.

**Farrell, Joseph.** 1993. "Meaning and Credibility in Cheap-Talk Games." *Games and Economic Behavior*, 5(4): 514–531.

**Fudenberg, Drew, and David K. Levine.** 1993. "Steady State Learning and Nash Equilibrium." *Econometrica*, 61(3): 547–573.

**Fudenberg, Drew, and David K. Levine.** 2006. "Superstition and Rational Learning." *American Economic Review*, 96(3): 630–651.

**Fudenberg, Drew, and David M. Kreps.** 1988. "A Theory of Learning, Experimentation, and Equilibrium in Games." http://economics.mit.edu/files/20982.

**Fudenberg, Drew, and David M. Kreps.** 1994. "Learning in Extensive-Form Games, II: Experimentation and Nash Equilibrium." http://economics.mit.edu/files/20979.

**Fudenberg, Drew, and Jean Tirole.** 1991*a*. *Game Theory.* The MIT Press.

**Fudenberg, Drew, and Jean Tirole.** 1991*b*. "Perfect Bayesian Equilibrium and Sequential Equilibrium." *Journal of Economic Theory*, 53(2): 236–260.

**Fudenberg, Drew, and Kevin He.** 2018. "Learning and Type Compatibility in Signaling Games." *Econometrica*, 86(4): 1215–1255.

**Fudenberg, Drew, and Kevin He.** 2020. "Payoff Information and Learning in Signaling Games." *Games and Economic Behavior*, 120: 96–120.

**Fudenberg, Drew, and Kevin He.** 2021. "Player-Compatible Learning and Player-Compatible Equilibrium." *Journal of Economic Theory*, 194: 105238.

**Fudenberg, Drew, and Yuichiro Kamada.** 2015. "Rationalizable Partition-Confirmed Equilibrium." *Theoretical Economics*, 10(3): 775–806.

**Fudenberg, Drew, and Yuichiro Kamada.** 2018. "Rationalizable Partition-Confirmed Equilibrium with Heterogeneous Beliefs." *Games and Economic Behavior*, 109: 364–381.

**Gittins, John. C.** 1979. "Bandit Processes and Dynamic Allocation Indices." *Journal of the Royal Statistical Society, Series B*, 41(2): 148–164.

**Goltsman, Maria, and Gregory Pavlov.** 2011. "How to Talk to Multiple Audiences." *Games and Economic Behavior*, 72(1): 100–122.

**Gordon, Sidartha, Navin Kartik, Melody Pei-yu Lo, Wojciech Olszewski, and Joel Sobel.** 2021. "Effective Communication in Cheap Talk Games." http://www.columbia.edu/~nk2339/Papers/GKLOS-EffectiveCommunication.pdf.

**Hart, Sergiu.** 2002. "Evolutionary Dynamics and Backward Induction." *Games and Economic Behavior*, 41(2): 227–264.

**Jehiel, Philippe, and Dov Samet.** 2005. "Learning to Play Games in Extensive Form by Valuation." *Journal of Economic Theory*, 124(2): 129–148.

**Kalai, Ehud, and Ehud Lehrer.** 1993. "Rational Learning Leads to Nash Equilibrium." *Econometrica*, 61(5): 1019–1045.

**Kartik, Navin.** 2009. "Strategic Communication with Lying Costs." *Review of Economic Studies*, 76(4): 1359–1395.

**Kartik, Navin, Marco Ottaviani, and Francescoo Squintani.** 2006. "Credulity, Lies, and Costly Talk." *Journal of Economic Theory*, 134(1): 93–116.

**Kohlberg, Elon, and Jean-Francois Mertens.** 1986. "On the Strategic Stability of Equilibria." *Econometrica*, 54(5): 1003–1037.

**Lehrer, Ehud, and Eilon Solan.** 2007. "Learning to Play Partially-Specified Equilibrium." http://www.math.tau.ac.il/~lehrer/Papers/LearningPSCE-web.pdf.

**Lipman, Barton L., and Duane J. Seppi.** 1995. "Robust Inference in Communication Games with Partial Provability." *Journal of Economic Theory*, 66(2): 370–405.

**Matthews, Steven A., Masahiro Okuno-Fujiwara, and Andrew Postlewaite.** 1991. "Refining Cheap Talk Equilibria." *Journal of Economic Theory*, 55(2): 247–273.

**Nachman, David C., and Thomas H. Noe.** 1994. "Optimal Design of Securities under Asymmetric Information." *Review of Financial Studies*, 7(1): 1–44.

**Nöldeke, Georg, and Larry Samuelson.** 1993. "An Evolutionary Analysis of Backward and Forward Induction." *Games and Economic Behavior*, 5(3): 425–454.

**Olszewski, Wojciech.** 2006. "Rich Language and the Refinements of Cheap-Talk Equilibria." *Journal of Economic Theory*, 128(1): 164–186.

**Rabin, Matthew.** 1990. "Communication between Rational Agents." *Journal of Economic Theory*, 51(1): 144–170.

**Rabin, Matthew, and Joel Sobel.** 1996. "Deviations, Dynamics, and Equilibrium Refinements." *Journal of Economic Theory*, 68(1): 1–25.

**Riley, John G.** 1979. "Informational Equilibrium." *Econometrica*, 47(2): 331–359.

**Rubinstein, Ariel, and Asher Wolinsky.** 1994. "Rationalizable Conjectural Equilibrium: Between Nash and Rationalizability." *Games and Economic Behavior*, 6(2): 299–311.

**Sobel, Joel, Lars Stole, and Iñigo Zapater.** 1990. "Fixed-Equilibrium Rationalizability in Signaling Games." *Journal of Economic Theory*, 52(2): 304–331.

**Zapater, Iñigo.** 1997. "Credible Proposals in Communication Games." *Journal of Economic Theory*, 72(1): 173–197.

## Appendix A: Other Refinements

### A1. Intuitive Criterion

LEMMA A1: *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 $a \in BR(\Theta,s)$.*

PROOF:
Let $\pi$ be a JCE. Fix $s \in S$ and suppose that $\Theta^\dagger(s,\pi) = \emptyset$. Let $\mathcal{A}_- = \{\alpha \in \Delta(BR(\Theta,s)) : u_1(\theta,s,\alpha) < u_1(\theta,\pi)\ \forall\theta \in \Theta\}$ be the set of mixtures over receiver best responses that make playing $s$ strictly worse for every type than their outcome under $\pi$. Similarly, let $\mathcal{A}_+ = \{\alpha \in \Delta(BR(\Theta,s)) : \exists\theta \in \Theta \text{ s.t. } u_1(\theta,s,\alpha) > u_1(\theta,\pi)\}$ be the set of mixtures over receiver best responses that make some type strictly better off by playing $s$ than under $\pi$. $\mathcal{A}_-$ and $\mathcal{A}_+$ are disjoint open subsets of $\Delta(BR(\Theta,s))$, and $\mathcal{A}_- \cup \mathcal{A}_+ = \Delta(BR(\Theta,s))$ since $\Theta^\dagger(s,\pi) = \emptyset$. As $\Delta(BR(\Theta,s))$ is connected, either $\Delta(BR(\Theta,s)) = \mathcal{A}_-$ or $\Delta(BR(\Theta,s)) = \mathcal{A}_+$. $\Delta(BR(\Theta,s)) = \mathcal{A}_+$ is not possible when $\pi$ is a JCE since then, for every $\alpha \in \Delta(BR(\overline{\Theta}(s,\pi),s))$, there is a $\theta$ such that $u_1(\theta,s,\alpha) > u_1(\theta,\pi)$. Thus $\Delta(BR(\Theta,s)) = \mathcal{A}_-$, so $u_1(\theta,s,a) < u_1(\theta,\pi)$ for all $a \in BR(\Theta,s)$. ■

PROOF OF PROPOSITION 2:
If $E(s,\pi) \neq \emptyset$, there is some $\theta$ and $\alpha \in BR(\Theta,s)$ such that $u_1(\theta,s,a) \geq u_1(\theta,\pi)$. By Lemma A1, $\Theta^\dagger(s,\pi) \neq \emptyset$, so $\overline{\Theta}(s,\pi) = \Theta^\dagger(s,\pi)$. Moreover, $\overline{\Theta}(s,\pi) \subseteq E(s,\pi)$, because $\max_{a \in BR(\Theta,s)} u_1(\theta,s,a) < u_1(\theta,\pi)$ implies $\widetilde{D}_\theta(s,\pi) \cup \widetilde{D}^0_\theta(s,\pi) = \emptyset \subseteq \widetilde{D}_{\theta'}$ for any $\theta' \in \Theta$. Thus, $BR(\overline{\Theta}(s,\pi),s) \subseteq BR(E(s,\pi),s)$. Hence, for all $\theta \in \Theta$, $\min_{a \in BR(E(s,\pi),s)} u_1(\theta,s,a) \leq \min_{a \in BR(\overline{\Theta}(s,\pi),s)} u_1(\theta,s,a) \leq u_1(\theta,\pi)$. ■

### A2. Co-D1

PROOF OF PROPOSITION 3:
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 $\pi$ is co-D1.

If $\Theta^\dagger(s,\pi) \neq \emptyset$, then $\overline{\Theta}(s,\pi) = \Theta^\dagger(s,\pi)$. Let $\theta$ 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 A1, $u_1(\theta,s,a) < u_1(\theta,\pi)$ 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, $\overline{\Theta}(s,\pi) = \Theta = \overline{\Theta}^{D1}(s,\pi)$. ■

### A3. NWBR

LEMMA A2: $\Theta^\ddagger(s,\pi) \subseteq \Theta^\dagger(s,\pi)$ *for all* $s \in S$ *and* $\pi \in \Pi$.

PROOF:

If $\theta \notin \Theta^{\dagger}(s,\pi)$, then by definition $\widetilde{D}^0_\theta(s,\pi) \subseteq \cup_{\theta'\neq\theta}\widetilde{D}_{\theta'}(s,\pi)$. 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 $\alpha \in \widetilde{D}^0_\theta(s,\pi)$. 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 $\alpha$ 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 $\theta \notin \Theta^{\ddagger}(s,\pi)$. ∎

LEMMA A3: *If $\pi$ 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 $\theta \in \Theta$ and $a \in BR(\Theta,s)$.*

The proof of Lemma A3 is analogous to that of Lemma A1, and is given in Online Appendix Section OA.5.

PROOF OF PROPOSITION 4:

Let $\pi$ be a PBE-H that satisfies NWBR, and for every off-path $s$, let $\alpha_s \in MBR(\widehat{\Theta}(s,\pi),s)$ be such that $u_1(\theta,s,\alpha) \leq u_1(\theta,\pi)$ for all $\theta \in \Theta$. We will show that $\widehat{\Theta}(s,\pi) \subseteq \overline{\Theta}(s,\pi)$ for all $s$, so the profile $\widetilde{\pi} = (\pi_1,\widetilde{\pi}_2)$ in which $\widetilde{\pi}_2$ coincides with $\pi_2$ for all on-path $s$ and dictates $\alpha_s$ for all off-path $s$ is a JCE that is path-equivalent to $\pi$.

If $\Theta^{\ddagger}(s,\pi) \neq \emptyset$, then by Lemma A2, $\Theta^{\ddagger}(s,\pi) \subseteq \Theta^{\dagger}(s,\pi)$, so $\widehat{\Theta}(s,\pi) \subseteq \overline{\Theta}(s,\pi)$. If $\Theta^{\ddagger}(s,\pi) = \emptyset$, then by Lemma A3, $u_1(\theta,s,a) < u_1(\theta,\pi)$ for all $\theta \in \Theta$ and $a \in BR(\Theta,s)$, so $\Theta^{\dagger}(s,\pi) = \emptyset$ and $\widehat{\Theta}(s,\pi) = \overline{\Theta}(s,\pi) = \Theta$. ∎

PROOF OF LEMMA 4:

Fix PBE-H $\pi$. We show that, for all $s \in S$ and $\theta \in \Theta$, $\widetilde{D}_\theta(s,\pi) \cup \widetilde{D}^0_\theta(s,\pi) \not\subseteq \cup_{\theta'\neq\theta}\widetilde{D}_{\theta'}(s,\pi)$ 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 $\widetilde{D}_\theta(s,\pi) \cup \widetilde{D}^0_\theta(s,\pi) \not\subseteq \cup_{\theta'\neq\theta}\widetilde{D}_{\theta'}(s,\pi)$.

Suppose that $\widetilde{D}_\theta(s,\pi) \cup \widetilde{D}^0_\theta(s,\pi) \not\subseteq \cup_{\theta'\neq\theta}\widetilde{D}_{\theta'}(s,\pi)$. Then there is some $\alpha \in \Delta(BR(\Theta,s))$ such that $u_1(\theta,s,\alpha) \geq u_1(\theta,\pi)$ 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 $\alpha' \in \Delta(BR(\Theta,s))$ such that $u_1(\theta,s,\alpha') \leq u_1(\theta,\pi)$. By continuity, there exists some $\alpha'' \in MBR(\Theta,s)$ 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, $D^0_\theta(s,\pi) \not\subseteq \cup_{\theta'\neq\theta}D_{\theta'}(s,\pi)$. ∎

## Appendix B: Supporting Results for Theorem 1

We use the following lemma in several proofs and examples. We omit its proof, which closely follows that of Proposition 5 in Fudenberg and He (2018).

LEMMA B1: *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 $\lim_{k\to\infty} \delta_k = 1$ and $\lim_{l\to\infty} \gamma_{1,k,l} = 1$ for all $k$. Then, for each $\theta \in \Theta$, $\pi_{1,\gamma_2}(\cdot|\theta)$ puts support only on the $(s,m)$ that are optimal for type $\theta$ against $\pi_{2,\gamma_2}$.*

PROOF OF LEMMA 1:

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 B1, 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 best replies to $\pi_{2,\gamma_{2,j}} = \lim_{k\to\infty} \lim_{l\to\infty} \pi_{2,j,k,l}$. 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 best replies to $\pi_2 = \lim_{j\to\infty} \pi_{2,\gamma_{2,j}}$. ■

PROOF OF LEMMA 2:

Let $q(\theta, s, m) = \lambda(\theta)\pi_1(s, m|\theta)$ be the distribution over $(\theta, s, m)$ induced by $\lambda$ and $\pi_1$, let $X^{\text{on}}$ be the set of sender signal-message pairs that occur with positive probability under $\pi$, and let $p_{(s,m)}(\theta)$ denote the conditional probability of $\theta$ given $(s,m) \in X^{\text{on}}$. For $\varepsilon > 0$, let $Q_\varepsilon = \{q' \in \Delta(\Theta \times S \times M) : \max_{(\theta,s,m)} |q'(\theta, s, m) - q(\theta, s, m)| \le \varepsilon\}$. Because best response correspondences are upper hemicontinuous, there is an $\varepsilon > 0$ such that every receiver whose belief $\widetilde{g}_2 \in \Delta(\Delta(\Theta \times S \times M))$ puts probability at least $1-\varepsilon$ on $Q_\varepsilon$ will respond to every $(s,m) \in X^{\text{on}}$ with some $a \in BR(p_{(s,m)}, s)$.

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 whatever empirical distribution they have observed. Moreover, by the law of large numbers, for any $\eta > 0$ we can take this $T$ to be such that, with probability at least $1-\eta/2$, a receiver who has lived more than $T$ periods assigns probability of at least $1-\varepsilon$ to $Q_{\varepsilon/2}$.

Fix sequences $\{\delta_n\}_{n\in\mathbb{N}}$, $\{\gamma_{1,n}\}_{n\in\mathbb{N}}$, and $\{\gamma_{2,n}\}_{n\in\mathbb{N}}$, and let $\pi_n = (\pi_{1,n}, \pi_{2,n}) \in \Pi^*(g, \delta_n, \gamma_{1,n}, \gamma_{2,n})$ be 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 is $\gamma_{2,n}^T$, which 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 $(s,m) \in X^{\text{on}}$ and $\eta > 0$, there exists some $N \in \mathbb{N}$ such that $\pi_{2,n}(BR(p_{(s,m)}, s)|s, m) \ge 1-\eta$ 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 $u_1(\theta, s, \pi_2(\cdot|s, m_{s,\widetilde{\Theta}})) <$

$u_1(\theta, \pi)$ for all $\theta \notin \widetilde{\Theta}$, Lemma B1 implies that there is some $J \in \mathbb{N}$ such that, for all $\theta' \notin \widetilde{\Theta}$ and $j > J$, $\lim_{k\to\infty} \lim_{l\to\infty} \pi_{j,k,l}(s, m_{s,\widetilde{\Theta}}|\theta') = 0$. 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 $\lim_{k\to\infty} \lim_{l\to\infty} \pi_{2,j,k,l}(BR(\widetilde{\Theta}, s)|s, m_{s,\widetilde{\Theta}}) = 1$ if $\lim_{k\to\infty} \lim_{l\to\infty} \pi_{j,k,l}(s, m_{s,\widetilde{\Theta}}|\theta') = 0$ for all $\theta' \notin \widetilde{\Theta}$. Since this holds for all $j > J$, $\pi_2(BR(\widetilde{\Theta}, s)|s, m_{s,\widetilde{\Theta}}) = 1$. ■

## Appendix C: A Sufficient Condition for Stability

DEFINITION C1: *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 $MBR(\Theta, s)$.

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

DEFINITION C2: *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 $u_2(p_{(s,m)}, s, a_x) > \max_{a\neq a_x} u_2(p_{(s,m)}, s, a)$,*
*3) For all $s \in S^{off}$, $u_1(\theta, s, a) < u_1(\theta, \pi)$ for all $\theta \in \Theta$ and $a \in BR(\overline{\Theta}(s, \pi), s)$.*

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

PROPOSITION C1: *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.4 in the Online Appendix contains the proof of Proposition C1. Because $\pi$ is uniformly justified, there is a receiver behavior strategy that makes each type strictly prefer to play their corresponding signal in $\pi$, and, when each 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 B1 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.[30] 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, and because receivers strictly prefer to conform to $\pi$, 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, in the limit, each sender type uses the same distribution over signals as in $\pi_1$. Consequently, the modified aggregate receiver response matches the true aggregate receiver response, and the fixed points of the modified response mapping are valid steady-state profiles that in the limit induce the same distribution over $\Theta \times S \times A$ as $\pi$.

## Appendix D: Details Omitted from Section II

***Strategy Mapping:*** The map $\sigma^{\delta,\gamma_1} : (\Delta(\mathcal{H}_1))^\Theta \times \Delta(\mathcal{H}_2) \to \Pi_1 \times \Pi_2$ taking the state in period $t$ to the aggregate strategy profile has component mappings $\sigma_1^{\delta,\gamma_1} : (\Delta(\mathcal{H}_1))^\Theta \to \Pi_1$ and $\sigma_2 : \Delta(\mathcal{H}_2) \to \Pi_2$ given by $\sigma_1^{\delta,\gamma_1}(\mu_1)[s,m|\theta] = \sum_{h_1:\mathbf{x}_\theta^{\delta,\gamma_1}(h_1)=(s,m)} \mu_\theta[h_1]$ and $\sigma_2(\mu_2)[a|s,m] = \sum_{h_2:\mathbf{y}(s,m|h_2)=a} \mu_2[h_2]$.

***Update Rule:*** The rule that maps the state in period $t$ to the state in period $t+1$, $\mathbf{f}^{\delta,\gamma_1,\gamma_2} : (\Delta(\mathcal{H}_1))^\Theta \times \Delta(\mathcal{H}_2) \to (\Delta(\mathcal{H}_1))^\Theta \times \Delta(\mathcal{H}_2)$, has the following components: The mapping $\mathbf{f}_\theta^{\delta,\gamma_1} : (\Delta(\mathcal{H}_1))^\Theta \times \Delta(\mathcal{H}_2) \to \Delta(\mathcal{H}_1)$ is given by $\mathbf{f}_\theta^{\delta,\gamma_1}(\mu)[\emptyset] = 1-\gamma_1$, and $\mathbf{f}_\theta^{\delta,\gamma_1}(\mu)[(h_1,(s,m,a))] = \gamma_1\mu_\theta[h_1] i_\theta^{\delta,\gamma_1}(h_1,s,m)\sigma_2(\mu)[a|s,m]$, where $(h_1,(s,m,a)) \in \mathcal{H}_1$ is the concatenation of the history $h_1 \in \mathcal{H}_1$ with a period where the sender plays $(s,m)$ and the receiver responds with $a$, and $i_\theta^{\delta,\gamma_1}(h_1,s,m)$ equals 1 if a type $\theta$ sender with history $h_1$ plays $(s,m)$ under policy $\mathbf{x}_\theta$ and equals 0 otherwise. Likewise, $\mathbf{f}_2^{\delta,\gamma_1,\gamma_2} : (\Delta(\mathcal{H}_1))^\Theta \times \Delta(\mathcal{H}_2) \to \Delta(\mathcal{H}_1)$ is given by $\mathbf{f}_2^{\delta,\gamma_1,\gamma_2}(\mu)[\emptyset] = 1-\gamma_2$, and $\mathbf{f}_2^{\delta,\gamma_1,\gamma_2}(\mu)[(h_2,(\theta,s,m))] = \gamma_2\mu_2[h_2]\lambda(\theta)\sigma_1^{\delta,\gamma_1}(\mu)[s,m|\theta]$, where $(h_2,(\theta,s,m)) \in \mathcal{H}_2$ is the concatenation of the history $h_2 \in \mathcal{H}_2$ with a period where the receiver is matched with a type $\theta$ sender who plays $(s,m)$.

***Aggregate Response Mapping:*** To define the aggregate response mapping, we first define mappings $\mathscr{L}_1^{\delta,\gamma_1} : \Pi_2 \to (\Delta(\mathcal{H}_1))^\Theta$ and $\mathscr{L}_2^{\gamma_2} : \Pi_1 \to \Delta(\mathcal{H}_2)$, which output the resulting $t \to \infty$ limit of the distribution of histories in the sender and receiver populations when the aggregate play of the opposing population is

[30] Fudenberg and Levine (2006) and Fudenberg and He (2020) proved 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.

held fixed at a given behavior strategy. For each $\theta \in \Theta$, $\mathscr{L}_{\theta}^{\delta,\gamma_1}(\pi_2)[\emptyset] = 1 - \gamma_1$ is the share of type $\theta$ senders with the null history. The share of type $\theta$ senders with histories $h_{1,t}$ of length $t > 0$ is defined by induction: For each $h_{1,t-1}$, we pair the signal-message pair a type $\theta$ sender with that history would use with the aggregate receiver strategy $\pi_2$ to compute the distribution of period-$t$ outcomes $(s_t, m_t, a_t)$ these senders observe, and assign the corresponding probabilities to the concatenation of these period-$t$ outcomes and $h_{1,t-1}$. (We do this formally in OA.2.1 in the Online Appendix.) Likewise, $\mathscr{L}_2^{\gamma_2}(\pi_1)[\emptyset] = 1 - \gamma_2$ is the share of receiver agents with the null history. A similar induction procedure gives the share of receiver agents with various histories of length $t > 0$: For each history of length $t - 1$, we take the strategy these agents would use, pair this with the distribution of sender types $\lambda$ and the aggregate sender strategy $\pi_1$ to compute the distribution of period-$t$ outcomes $(\theta, s_t, m_t)$ these agents observe, and assign the corresponding probability to the concatenation of the period-$t$ outcomes and the previous history.

The components of the aggregate response mapping $\mathscr{R}^{\delta,\gamma_1,\gamma_2}(\pi) = (\mathscr{R}_1^{\delta,\gamma_1}(\pi_2), \mathscr{R}_2^{\gamma_2}(\pi_1))$ are then found by composing $\mathscr{L}_1^{\delta,\gamma_1}$ and $\mathscr{L}_2^{\gamma_2}$ with strategy mapping $\sigma^{\delta,\gamma_1}$: The aggregate sender response mapping is given by $\mathscr{R}_1^{\delta,\gamma_1}(\pi_2) = \sigma_1^{\delta,\gamma_1}(\mathscr{L}_1^{\delta,\gamma_1}(\pi_2))$, and the aggregate receiver response mapping is given by $\mathscr{R}_2^{\gamma_2}(\pi_1) = \sigma_2(\mathscr{L}_2^{\gamma_2}(\pi_2))$.

# Online Appendix for "Justified Communication Equilibrium"

Daniel Clark and Drew Fudenberg

## OA.1 Equivalent Definition of JCE

We show that it would be equivalent to define JCE by setting $\Theta^{\dagger}(s,\pi) = \{\theta \in \Theta : \widetilde{D}^0_\theta(s,\pi) \not\subseteq \cup_{\theta' \neq \theta} \widetilde{D}_{\theta'}(s,\pi)\}$, rather than $\Theta^{\dagger}(s,\pi) = \{\theta \in \Theta : \widetilde{D}_\theta(s,\pi) \cup \widetilde{D}^0_\theta(s,\pi) \not\subseteq \cup_{\theta' \neq \theta} \widetilde{D}_{\theta'}(s,\pi)\}$.

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

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

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

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

**Proposition OA 1.** *If $\pi$ is a PBE-H, then $\overline{\Theta}(s,\pi) = \overline{\Theta}'(s,\pi)$ for all $s \in S$.*

*Proof.* Fix PBE-H $\pi$. We will argue that $\Theta^{\dagger'}(s,\pi) = \Theta^{\dagger}(s,\pi)$, which gives $\overline{\Theta}'(s,\pi) = \overline{\Theta}(s,\pi)$.

First, suppose that $\theta \in \Theta^{\dagger'}(s,\pi)$. Then by definition, $\widetilde{D}^0_\theta(s,\pi) \not\subseteq \cup_{\theta'\neq\theta}\widetilde{D}_{\theta'}(s,\pi)$. Hence, $\widetilde{D}_\theta(s,\pi) \cup \widetilde{D}^0_\theta(s,\pi) \not\subseteq \cup_{\theta'\neq\theta}\widetilde{D}_{\theta'}(s,\pi)$, so $\theta \in \Theta^{\dagger}(s,\pi)$.

Now, suppose that $\theta \in \Theta^{\dagger}(s,\pi)$. Then by definition, $\widetilde{D}_\theta(s,\pi)\cup\widetilde{D}^0_\theta(s,\pi) \not\subseteq \cup_{\theta'\neq\theta}\widetilde{D}_{\theta'}(s,\pi)$. Thus, there is some $\alpha \in \Delta(BR(\Theta,s))$ 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 $\theta' \neq \theta$. Since $\pi$ is a PBE-H, there is also some $\alpha' \in \Delta(BR(\Theta,s))$ such that $u_1(\theta',s,\alpha') \leq u_1(\theta,\pi)$ for all $\theta' \in \Theta$. By continuity, there is some $\nu \in [0,1]$ and $\alpha'' = \nu\alpha + (1-\nu)\alpha'$ such that $u_1(\theta,s,\alpha'') = u_1(\theta,\pi)$, while $u_1(\theta',s,\alpha'') \leq u_1(\theta',\pi)$ for all $\theta' \neq \theta$. As $\alpha'' \in \Delta(BR(\Theta,s))$, it follows that $\widetilde{D}^0_\theta(s,\pi) \not\subseteq \cup_{\theta'\neq\theta}\widetilde{D}_{\theta'}(s,\pi)$, so $\theta \in \Theta^{\dagger'}(s,\pi)$. ■

## OA.2 Omitted Analysis of Learning Model

### OA.2.1 Continuity of Aggregate Response Mapping

We begin by formally defining the auxiliary maps $\mathscr{L}_1^{\delta,\gamma_1} : \Pi_2 \to (\Delta(\mathcal{H}_1))^{\Theta}$ and $\mathscr{L}_2^{\gamma_2} : \Pi_1 \to \Delta(\mathcal{H}_2)$ introduced in Appendix D. For each $\theta \in \Theta$, let

$$\mathscr{L}_\theta^{\delta,\gamma_1}(\pi_2)[\emptyset] = 1-\gamma_1,$$
$$\mathscr{L}_\theta^{\delta,\gamma_1}(\pi_2)[(h_1,(s,m,a))] = \gamma_1\mathscr{L}_\theta^{\delta,\gamma_1}(\pi_2)[h_1]i_\theta^{\delta,\gamma_1}(h_1,s,m)\pi_2[a|s,m],$$

for all $h_1 \in \mathcal{H}_1$, $s \in S$, $m \in M$, and $a \in A$. To define $\mathscr{L}_2^{\gamma_2}$, let

$$\mathscr{L}_2^{\gamma_2}(\pi_1)[\emptyset] = 1-\gamma_2,$$
$$\mathscr{L}_2^{\gamma_2}(\pi_1)[(h_2,(\theta,s,m))] = \gamma_2\mathscr{L}_2^{\gamma_2}(\pi_1)[h_2]\lambda(\theta)\pi_1[s,m|\theta],$$

for all $h_2 \in \mathcal{H}_2$, $\theta \in \Theta$, $s \in S$, and $m \in M$.

We now establish the continuity of various mappings involving distributions over histories, which we endow with the sup-norm topology.

**Claim OA 1.** *The aggregate strategy mapping* $\sigma^{\delta,\gamma_1} : (\Delta(\mathcal{H}_1))^\Theta \times \Delta(\mathcal{H}_2) \to \Pi_1 \times \Pi_2$ *is continuous.*

*Proof.* We prove that $\sigma_1^{\delta,\gamma_1} : (\Delta(\mathcal{H}_1))^\Theta \to \Pi_1$ is continuous. An analogous argument handles $\sigma_2 : \Delta(\mathcal{H}_2) \to \Pi_2$.

To show that $\sigma_1^{\delta,\gamma_1}$ is continuous, we establish that $\lim_{\mu_1' \to \mu_1} \sigma_1^{\delta,\gamma_1}(\mu_1')[s,m|\theta] = \sigma_1^{\delta,\gamma_1}(\mu_1)[s,m|\theta]$ for all $s \in S$, $m \in M$, $\theta \in \Theta$, and $\mu_1 \in (\Delta(\mathcal{H}_1))^\Theta$. Since $\sum_{s,m} \sigma_1^{\delta,\gamma_1}(\mu_1')[s,m|\theta] = 1$ for all $\mu_1 \in (\Delta(\mathcal{H}_1))^\Theta$, it suffices to show that $\liminf_{\mu_1' \to \mu_1} \sigma_1^{\delta,\gamma_1}(\mu_1')[s,m|\theta] \geq \sigma_1^{\delta,\gamma_1}(\mu_1)[s,m|\theta]$ for all $s$, $m$, and $\theta$. For any $\varepsilon > 0$, let $\mathcal{H}_{1,\varepsilon}$ be a finite set of sender histories such that $\sum_{h_1 \in \mathcal{H}_{1,\varepsilon} : \mathbf{x}_\theta^{\delta,\gamma_1}(h_1) = (s,m)} \mu_\theta[h_1] \geq \sigma_1^{\delta,\gamma_1}(\mu_1)[s,m|\theta] - \varepsilon$. By the nature of the sup-norm topology, $\lim_{\mu_1' \to \mu_1} \sum_{h_1 \in \mathcal{H}_{1,\varepsilon} : \mathbf{x}_\theta^{\delta,\gamma_1}(h_1) = (s,m)} \mu_\theta'[h_1] = \sum_{h_1 \in \mathcal{H}_{1,\varepsilon} : \mathbf{x}_\theta^{\delta,\gamma_1}(h_1) = (s,m)} \mu_\theta[h_1]$. Since $\mu_\theta'[h_1] \geq 0$ for all $h_1 \in \mathcal{H}_1$ and $\mu_1' \in (\Delta(\mathcal{H}_1))^\Theta$, it follows that $\liminf_{\mu_1' \to \mu_1} \sigma_1^{\delta,\gamma_1}(\mu_1')[s,m|\theta] = \liminf_{\mu_1' \to \mu_1} \sum_{h_1 : \mathbf{x}_\theta^{\delta,\gamma_1}(h_1) = (s,m)} \mu_\theta'[h_1] \geq \lim_{\mu_1' \to \mu_1} \sum_{h_1 \in \mathcal{H}_{1,\varepsilon} : \mathbf{x}_\theta^{\delta,\gamma_1}(h_1) = (s,m)} \mu_\theta'[h_1] \geq \sigma_1^{\delta,\gamma_1}(\mu_1)[s,m|\theta] - \varepsilon$. As this holds for arbitrary $\varepsilon > 0$, the desired conclusion follows. ■

**Claim OA 2.** *Both* $\mathscr{L}_1^{\delta,\gamma_1} : \Pi_2 \to (\Delta(\mathcal{H}_1))^\Theta$ *and* $\mathscr{L}_2^{\gamma_2} : \Pi_1 \to \Delta(\mathcal{H}_2)$ *are continuous.*

*Proof.* We prove that $\mathscr{L}_1^{\delta,\gamma_1} : \Pi_2 \to (\Delta(\mathcal{H}_1))^\Theta$ is continuous. An analogous argument handles $\mathscr{L}_2^{\gamma_2} : \Pi_1 \to \Delta(\mathcal{H}_2)$.

For all $\pi_2 \in \Pi_2$, $\mathscr{L}_1^{\delta,\gamma_1}(\pi_2)[h_1] \leq (1-\gamma_1)\gamma_1^t$ for every history $h_1$ of length $t$. Since $\lim_{t\to\infty}(1-\gamma_1)\gamma_1^t = 0$, to establish that $\mathscr{L}_1^{\delta,\gamma_1}(\pi_2)$ is a continuous function of $\pi_2$, it thus suffices to show that $\mathscr{L}_1^{\delta,\gamma_1}(\pi_2)[h_1]$ is continuous for every history $h_1 \in \mathcal{H}_1$. We show this inductively over sender histories. For the null sender history $h_1 = \emptyset$, $\mathscr{L}_1^{\delta,\gamma_1}(\pi_2)[\emptyset]$ for all $\pi_2 \in \Pi_2$ and is thus continuous. Assuming that $\mathscr{L}_1^{\delta,\gamma_1}(\pi_2)[h_1]$ is a continuous function of $\pi_2$, it follows that $\mathscr{L}_1^{\delta,\gamma_1}(\pi_2)[(h_1,(s,m,a))]$ is a continuous function of $\pi_2$ for all $s$, $m$, and $a$, as can be seen from the expression for $\mathscr{L}_1^{\delta,\gamma_1}$ given earlier. This completes the inductive argument. ■

**Corollary OA 1.** *The aggregate response mapping* $\mathscr{R}^{\delta,\gamma_1,\gamma_2} : \Pi_1 \times \Pi_2 \to \Pi_1 \times \Pi_2$ *is continuous.*

*Proof.* By Claims OA 1 and OA 2, $\sigma_1^{\delta,\gamma_1}$ and $\mathscr{L}_1^{\delta,\gamma_1}$ are continuous. Thus $\mathscr{R}_1^{\delta,\gamma_1}(\pi_2) = \sigma_1^{\delta,\gamma_1}(\mathscr{L}_1^{\delta,\gamma_1}(\pi_2))$ is a continuous function of $\pi_2$. Likewise, since $\sigma_2$ and $\mathscr{L}_2^{\gamma_2}$ are continuous, $\mathscr{R}_2^{\gamma_2}(\pi_1) = \sigma_2(\mathscr{L}_2^{\gamma_2}(\pi_1))$ is a continuous function of $\pi_1$. ∎

### OA.2.2 Characterization of Steady State Profiles

**Proposition OA 2.** *Strategy profile $\pi$ is a fixed point of $\mathscr{R}^{\delta,\gamma_1,\gamma_2}$ if and only if there is some steady state $\mu$ such that $\sigma^{\delta,\gamma_1}(\mu) = \pi$.*

*Proof.* Suppose that $\mu$ is a steady state satisfying $\sigma^{\delta,\gamma_1}(\mu) = \pi$. Since $\mu$ is a steady state, the aggregate receiver play in every period is fixed at $\pi_2 = \sigma_2(\mu)$. By definition, $\mathscr{L}_1^{\delta,\gamma_1}(\pi_2)$ is the $t \to \infty$ limit of the distribution over histories in the sender population when the aggregate receiver play is fixed at $\pi_2$. Since $\mu$ is a steady state, it follows that $\mathscr{L}_1^{\delta,\gamma_1}(\pi_2) = \mu_1$. From this, we obtain $\mathscr{R}_1^{\delta,\gamma_1}(\pi_2) = \sigma_1^{\delta,\gamma_1}(\mathscr{L}_1^{\delta,\gamma_1}(\pi_2)) = \sigma_1^{\delta,\gamma_1}(\mu_1) = \pi_1$. An almost identical argument shows that $\mathscr{R}_2^{\gamma_2}(\pi_1) = \pi_2$. We conclude that $\mathscr{R}^{\delta,\gamma_1,\gamma_2}(\pi) = \pi$.

Conversely,suppose that $\pi$ is a fixed point of $\mathscr{R}^{\delta,\gamma_1,\gamma_2}$. Let $\mu$ be the state given by $\mu_1 = \mathscr{L}_1^{\delta,\gamma_1}(\pi_2)$ and $\mu_2 = \mathscr{L}_2^{\gamma_2}(\pi_1)$. Observe that $\sigma_1^{\delta,\gamma_1}(\mu_1) = \sigma_1^{\delta,\gamma_1}(\mathscr{L}_1^{\delta,\gamma_1}(\pi_2)) = \mathscr{R}_1^{\delta,\gamma_1}(\pi_2) = \pi_1$ and $\sigma_2(\mu_2) = \sigma_2(\mathscr{L}_2^{\gamma_2}(\pi_1)) = \mathscr{R}_2^{\gamma_2}(\pi_1) = \pi_2$, so $\pi = \sigma^{\delta,\gamma_1}(\mu)$ is the aggregate strategy profile for state $\mu$. All that remains is to establish that $\mu$ is a steady state, which amounts to showing that $\mathbf{f}_\theta^{\delta,\gamma_1}(\mathscr{L}_1^{\delta,\gamma_1}(\pi_2))[h_1] = \mathscr{L}_1^{\delta,\gamma_1}(\pi_2)[h_1]$ for all $h_1 \in \mathcal{H}_1$ and $\theta \in \Theta$ and $\mathbf{f}_2^{\delta,\gamma_1,\gamma_2}(\mathscr{L}_2^{\gamma_2}(\pi_1))[h_2] = \mathscr{L}_2^{\gamma_2}(\pi_1)[h_2]$ for all $h_2 \in \mathcal{H}_2$. We argue inductively over sender histories that $\mathbf{f}_\theta^{\delta,\gamma_1}(\mathscr{L}_1^{\delta,\gamma_1}(\pi_2))[h_1] = \mathscr{L}_1^{\delta,\gamma_1}(\pi_2)[h_1]$ for all $h_1 \in \mathcal{H}_1$. (A similar inductive argument shows that $\mathbf{f}_2^{\delta,\gamma_1,\gamma_2}(\mathscr{L}_2^{\gamma_2}(\pi_1))[h_2] = \mathscr{L}_2^{\gamma_2}(\pi_1)[h_2]$ for all $h_2 \in \mathcal{H}_2$.) For the null sender history $h_1 = \emptyset$, the equality holds since $\mathbf{f}_\theta^{\delta,\gamma_1}(\mathscr{L}_1^{\delta,\gamma_1}(\pi_2))[\emptyset] = 1-\gamma_1 = \mathscr{L}_1^{\delta,\gamma_1}(\pi_2)[\emptyset]$. Assuming that $\mathbf{f}_\theta^{\delta,\gamma_1}(\mathscr{L}_1^{\delta,\gamma_1}(\pi_2))[h_1] = \mathscr{L}_1^{\delta,\gamma_1}(\pi_2)[h_1]$ holds, it necessarily follows that $\mathbf{f}_\theta^{\delta,\gamma_1}(\mathscr{L}_1^{\delta,\gamma_1}(\pi_2))[(h_1,(s,m,a))] = \mathscr{L}_1^{\delta,\gamma_1}(\pi_2)[(h_1,(s,m,a))]$ for all $s$, $m$, and $a$ since $\sigma_2(\mu_2) = \pi_2$. This completes the inductive argument. ∎

## OA.3 Comparison with RCE

In this section, we restrict attention to signaling games without communication, i.e. $M$ is singleton. We write $\Pi_2^\bullet = \times_{s\in S}\Delta(BR(\Theta, s))$ for the set of receiver strategies that assign probability 0 to conditionally dominated responses.

**Definition OA 1** (Fudenberg and He, 2020). *Signal $s \in S$ is* **more rationally-compatible** *with $\theta'$ than $\theta''$, written as $\theta' \succsim_s \theta''$,*

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

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

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

and let $BR(\overline{P}(s,\pi), s) = \cup_{p\in\overline{P}(s,\pi)} BR(p, s)$ be the set of receiver best responses to signal $s$ for some $p \in \overline{P}(s, \pi)$.

**Definition OA 2** (Fudenberg and He, 2020). *Strategy profile $\pi$ is a* **rationality-compatible equilibrium (RCE)** *if it is a PBE-H where, for every $s \in S$, $\pi_2(\cdot|s) \in \Delta(BR(\overline{P}(s,\pi), s))$.*

This definition requires that the receiver's posterior likelihood ratio for types $\theta'$ and

$\theta''$ dominates the prior likelihood ratio whenever $\theta' \succsim_s \theta''$. It also requires that the posterior assigns probability 0 to equilibrium-dominated types.

**Proposition OA 3.** *If $\pi$ is a justified communication equilibrium, then $\pi$ is an RCE.*

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 $\Delta(\overline{\Theta}(s,\pi)) \subseteq \overline{P}(s,\pi)$. 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 $\pi$ 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 A1, $\overline{\Theta}(s,\pi) = \Theta^\dagger(s,\pi)$ and $\Delta(\overline{\Theta}(s,\pi)) \subseteq \Delta(E(s,\pi))$. Suppose that $\theta'$ and $\theta''$ are two types such that $\theta' \succsim_s \theta''$. Then Definition OA 2 implies that $\widetilde{D}_{\theta''}(s,\pi) \cup \widetilde{D}^0_{\theta''}(s,\pi) \subseteq \widetilde{D}_{\theta'}(s,\pi)$, so $\theta'' \notin \Theta^\dagger(s,\pi)$. As a result, $\Delta(\overline{\Theta}(s,\pi)) = \Delta(\Theta^\dagger(s,\pi)) \subseteq \cap_{(\theta',\theta'')\text{ s.t. }\theta' \succsim_s \theta''} P_{\theta' \rhd \theta''}$. We conclude $\Delta(\overline{\Theta}(s,\pi)) \subseteq \Delta(E(s,\pi)) \cap (\cap_{(\theta',\theta'')\text{ s.t. }\theta' \succsim_s \theta''} P_{\theta' \rhd \theta''}) = \overline{P}(s,\pi)$. ■

## OA.4 Proof of Proposition C1

**Proposition C1.** *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.*

*Proof.* Because $\pi$ is a uniformly justified JCE in a strictly monotonic signaling game, $\pi_2(\cdot|s,m) = \pi_2(\cdot|s,m')$ for all $s \in S$ and $m, m' \in M$ such that $(s,m),(s,m') \in X^{\text{on}}$. Thus, for every $s \in S^{\text{on}}$, there is some $a_s \in A$ such that $\pi_2(a_s|s,m) = 1$ for all $(s,m) \in X^{\text{on}}$. For all $s \in S^{\text{off}}$, fix some $a_s \in BR(\overline{\Theta}(s,\pi),s)$.

Our construction modifies the aggregate receiver response so that the response to any $s$ is $a_s$ with high probability unless the aggregate sender play is such that each type $\theta \in \Theta$ uses $s_\theta$ with sufficiently high probability. We show that the fixed points of this modified aggregate response mapping correspond to fixed points of the true aggregate response mapping in the iterated limit where $\gamma_1 \to 1$ then $\delta \to 1$ then $\gamma_2 \to 1$. Moreover, we show that the limit of these steady state profiles induce the same distribution over $\Theta \times S \times A$ as $\pi$.

Because $\pi$ is a uniformly justified JCE in a strictly monotonic signaling game, there is an $\varepsilon > 0$ such that the following two properties hold. First, when $\pi_2(a_s|s,m) \geq 1-\varepsilon$ for all $s$, playing $s_\theta$ paired with message $m$ is strictly better for type $\theta$ than playing any other $s' \neq s_\theta$ paired with any $m'$. Second, if $\pi_1(s_\theta, m|\theta) \geq 1-\varepsilon$ for every $\theta \in \Theta$, it is strictly optimal for the receiver to respond to $(s,m)$ with $a_s$ for every $s \in S^{\text{on}}$. Fix such an $\varepsilon$.

Let $\kappa : \mathbb{R} \to [0,1]$ be a continuous function such that $\kappa(z) = 0$ for all $z \leq 0$ and $\kappa(z) = 1$ for all $z \geq 1$. Also, let $\phi : \Pi_1 \times \Pi_2 \to \Pi_2$ be the mapping

$$\phi(\pi_1,\pi_2)(\cdot|s,m) = \left(1 - \kappa\left(\frac{2}{\varepsilon}(\min_{\theta\in\Theta} \pi_1(s_\theta|\theta) - 1 + \varepsilon)\right)\right)\mathbb{1}_{a_s}(\cdot) + \kappa\left(\frac{2}{\varepsilon}(\min_{\theta\in\Theta} \pi_1(s_\theta|\theta) - 1 + \varepsilon)\right)\pi_2(\cdot|s,m)$$

for all $s \in S$ and $m \in M$. Note that $\phi$ is continuous. Additionally, $\phi(\pi_1,\pi_2)(a_s|s,m) = 1$ when $\pi_1(s_\theta|\theta) \leq 1-\varepsilon$ for some $\theta \in \Theta$, and $\phi(\pi_1,\pi_2) = \pi_2$ when $\pi_1(s_\theta|\theta) \geq 1-\varepsilon/2$ for all $\theta \in \Theta$.

Consider the correspondence $\widetilde{\mathscr{R}}^{\delta,\gamma_1,\gamma_2} : \Pi_1 \times \Pi_2 \to \Pi_1 \times \Pi_2$ given by $\widetilde{\mathscr{R}}^{\delta,\gamma_1,\gamma_2}(\pi) = (\mathscr{R}_1^{\delta,\gamma_1}(\pi_2), \phi(\pi_1, \mathscr{R}_2^{\gamma_2}(\pi_1)))$. Since $\widetilde{\mathscr{R}}^{\delta,\gamma_1,\gamma_2}$ is continuous, Brouwer's fixed point theorem guarantees the existence of a fixed point $(\pi_1^{\delta,\gamma_1,\gamma_2}, \pi_2^{\delta,\gamma_1,\gamma_2})$. We will establish that, in the iterated limit where $\gamma_1 \to 1$ then $\delta \to 1$ then $\gamma_2 \to 1$, $\pi^{\delta,\gamma_1,\gamma_2} = (\pi_1^{\delta,\gamma_1,\gamma_2}, \pi_2^{\delta,\gamma_1,\gamma_2})$ induces the same distribution over $\Theta \times S \times A$ as $\pi$. Towards this end, consider a sequence $\{\gamma_{2,j}\}_{j\in\mathbb{N}}$, sequences $\{\delta_{j,k}\}_{j,k\in\mathbb{N}}$, and sequences $\{\gamma_{1,j,k,l}\}_{j,k,l\in\mathbb{N}}$ such that (1) $\lim_{j\to\infty} \gamma_{2,j} = 1$, (2) $\lim_{k\to\infty} \delta_{j,k} = 1$ for all $j$, (3) $\lim_{l\to\infty} \gamma_{1,j,k,l} = 1$ for all $j,k$, and (4) $\lim_{j\to\infty} \lim_{k\to\infty} \lim_{l\to\infty} \pi^{\delta_{j,k},\gamma_{1,j,k,l},\gamma_{2,j}} = \pi'$ for some $\pi' = (\pi_1', \pi_2') \in \Pi_1 \times \Pi_2$.

We first establish that $\pi_1'(s_\theta|\theta) \geq 1-\varepsilon$ for all $\theta \in \Theta$. If instead there were some $\theta \in \Theta$ such that $\pi'(s_\theta|\theta) < 1-\varepsilon$, then by construction, $\pi_2'(a_s|s,m) \geq 1-\varepsilon$ for all

$s \in S$ and $m \in M$. Lemma B1 thus requires that $\pi_1'(s_\theta|\theta) = 1$ for all $\theta \in \Theta$, which is a contradiction.

Next we show that $\pi_2'(a_s|s,m) = 1$ for all $s \in S^{\text{on}}$ and $m \in M$ such that $\pi_1'(s,m|\theta) > 0$ for some $\theta \in \Theta$. Fix $s \in S^{\text{on}}$. Consider $m, m' \in M$ such that $\pi_1'(s,m|\theta) > 0$ and $\pi_1'(s,m'|\theta') > 0$ for some $\theta, \theta' \in \Theta$. The construction of $\widetilde{\mathscr{R}}^{\delta,\gamma_1,\gamma_2}$, along with an argument almost identical to the proof of Lemma 2, implies that there exists some $\xi \in [0,1]$ and $\alpha, \alpha' \in MBR(\Theta, s)$ such that $\pi_2'(\cdot|s,m) = (1-\xi)\mathbb{1}_{a_s}(\cdot) + \xi\alpha$ and $\pi_2'(\cdot|s,m') = (1-\xi)\mathbb{1}_{a_s}(\cdot) + \xi\alpha'$. In fact, $\alpha$ and $\alpha'$ must be optimal responses to $s$ under the posterior distributions obtained by updating $\lambda$ using $\{\pi_1'(s,m|\theta)\}_{\theta\in\Theta}$ and $\{\pi_1'(s,m'|\theta)\}_{\theta\in\Theta}$, respectively. Because the game is strictly monotonic, Lemma B1 implies that $\alpha = \alpha'$. Thus, for a given $s$, $\pi_2'(\cdot|s,m)$ is the same for all $m \in M$ for which there is a $\theta' \in \Theta$ such that $\pi_1'(s,m|\theta') > 0$. Combining this with the fact that $\pi_1'(s_\theta|\theta) \geq 1 - \varepsilon$ for all $\theta$, it follows that $\pi_2'(a_s|s,m) = 1$ for all $m \in M$ such that $\pi_1'(s,m|\theta) > 0$ for some $\theta \in \Theta$.

Since $\pi_2'(a_s|s,m) = 1$ for all $s \in S^{\text{on}}$ and $m \in M$ such that $\pi_1'(s,m|\theta) > 0$ for some $\theta \in \Theta$, it follows from Lemma B1 that $\pi_1'(s|\theta) = 0$ whenever $s \in S^{\text{on}}$ and $s \neq s_\theta$. We now show that for all $\theta \in \Theta$, $\pi_1'(s|\theta) = 0$ for all $s \in S^{\text{off}}$. Note that, because $\pi_1(s_\theta|\theta) > 0$ for all $\theta \in \Theta$ and $\pi_2(a_{s_\theta}|s_\theta, m) = 1$ for all $\theta \in \Theta$ and $m \in M$ where $\pi_1(s_\theta, m|\theta) > 0$, Lemma B1 implies that $u_1(\theta, \pi') = u_1(\theta, s_\theta, a_{s_\theta}) = u_1(\theta, \pi)$ for all $\theta \in \theta$. Additionally, Lemma B1 requires that $u_1(\theta, s, \pi_2'(\cdot|s,m)) \leq u_1(\theta, \pi') = u_1(\theta, \pi)$ for all $\theta \in \Theta$, $s \in S$, and $m \in M$. Now, suppose that there is some $s \in S^{\text{off}}$ and $m \in M$ such that $\pi_1'(s,m|\theta) > 0$ for some $\theta \in \Theta$. There are two possible cases: (1) There is some $\theta \notin \overline{\Theta}(s,\pi)$ such that $\pi_1'(s,m|\theta) > 0$, and (2) All $\theta$ with $\pi_1'(s,m|\theta) > 0$ belong to $\overline{\Theta}(s,\pi)$. In Case (1), because $\pi_2'(\cdot|s,m) \in \Delta(BR(\Theta, s))$, there must be some $\theta' \in \overline{\Theta}(s,\pi)$ such that $u_1(\theta', s, \pi_2'(\cdot|s,m)) > u_1(\theta', \pi)$, which is a contradiction. In Case (2), the construction of $\mathscr{R}^{\delta,\gamma_1,\gamma_2}$, combined with an almost identical argument to the one behind Lemma 2, implies that $\pi_2'(\cdot|s,m) \in \Delta(BR(\overline{\Theta}(s,\pi), s))$. Since $\pi$ is a uniformly justified JCE, it follows that $u_1(\theta, s, \pi_2'(\cdot|s,m)) < u_1(\theta, \pi)$ for all $\theta \in \Theta$, but this, along with Lemma B1, implies that $\pi_1'(s,m|\theta) = 0$ for all $\theta \in \Theta$, a contradiction.

It follows that $\pi_1'(s_\theta|\theta) = 1$ for all $\theta$ and $\pi_2'(a_s|s,m) = 1$ for all $s \in S^{\text{on}}$ and $m \in M$ such that $\pi_1'(s,m|\theta) > 0$ for some $\theta \in \Theta$. Thus, $\pi^{\delta,\gamma_1,\gamma_2}$ induces the same distribution over $\Theta \times S \times A$ as $\pi$ in the iterated limit where first $\gamma_1 \to 1$ then $\delta \to 1$ then $\gamma_2 \to 1$. Moreover, since $\pi_1'(s_\theta|\theta) = 1$ for all $\theta \in \Theta$, $\pi_2^{\delta,\gamma_1,\gamma_2} = \phi(\pi_1^{\delta,\gamma_1,\gamma_2}, \mathscr{R}_2^{\gamma_2}(\pi_1^{\delta,\gamma_1,\gamma_2})) = \mathscr{R}_2^{\gamma_2}(\pi_1^{\delta,\gamma_1,\gamma_2})$ in the iterated limit. Thus, $\pi^{\delta,\gamma_1,\gamma_2}$ is a fixed point of $\mathscr{R}^{\delta,\gamma_1,\gamma_2}$ in the iterated limit, which means that $\pi'$ is a stable profile. ■

## OA.5 Proof of Lemma A3

**Lemma A3.** *If $\pi$ 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* $\theta \in \Theta$ *and* $a \in BR(\Theta,s)$.

*Proof.* Let $\pi$ 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 $\pi$. 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 $\Theta^\ddagger(s,\pi) = \emptyset$. 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 $\pi$ is a PBE-H that satisfies NWBR since then, for every $\alpha \in MBR(\widehat{\Theta}(s,\pi),s)$, 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 $a \in BR(\Theta,s)$. ■

## OA.6 Omitted Analysis of Examples

### OA.6.1 Analysis of Example 2

**Proposition OA 4.** *The game in Example 2 has stable profiles where all types play Pass with probability* 1.

*Proof.* We specify that the worker prior $g_2$ is a Dirichlet distribution. For $m \in \{m_{Hire,\theta_H}, m_{Hire,\{\theta_H,\theta_M\}}\}$, it has initial weight 1 on $(\theta_H, Hire, m)$, 1/2 on $(\theta_M, Hire, m)$, and 1/4 on $(\theta_L, Hire, m)$. For $m = m_{Hire,\theta_M}$, it has initial weight 3/5 on $(\theta_H, Hire, m)$, 1 on $(\theta_M, Hire, m)$, and 1/4 on $(\theta_L, Hire, m)$. For all other messages $m$, it has initial weight 1/4 on $(\theta_H, Hire, m)$, 1/4 on $(\theta_M, Hire, m)$, and 1 on $(\theta_L, Hire, m)$. Note that initial trust is satisfied: For instance, when a worker first encounters a firm who plays $(Hire, m_{In,\theta_H})$, the probability they place on the firm having type $\theta_H$ is 4/7, $\theta_M$ is 2/7, and $\theta_L$ is 1/7, so $e_H = BR(\theta_H, Hire)$ is optimal.

We observe that $e_L$ is the worker's unique best response to $Hire$ under any distribution that puts probability strictly higher than 3/7 on $\theta_L$. Additionally, if a worker has encountered past play of $(Hire, m)$ and all such plays have been by firms with type $\theta_L$, then the worker will respond to the next instance of $(Hire, m)$ with $e_L$. To see that this holds for the case $m = m_{Hire,\theta_H}$, note that the worker's conditional distribution over the firm's type after $(Hire, m_{Hire,\theta_H})$ must put probability at least 5/11 on $\theta_L$. Analogous arguments handle the other cases.

We focus on steady state profiles in which, for every $m \in M$, the aggregate probability that a worker responds to $(Hire, m)$ with $e_M$ is less than 1/4. Under such responses, whenever it is weakly optimal for $\theta_H$ or $\theta_M$ to play $Hire$, it must be strictly optimal for $\theta_L$ to do so. To see this, note that

$$u_1(\theta_H, Hire, \alpha) = 21\alpha[e_H] + 6\alpha[e_M] - 5,$$

so $\alpha[e_H] \geq 5/21 - 6/21\alpha[e_M]$ whenever $u_1(\theta_H, Hire, \alpha) \geq 0$, and

$$u_1(\theta_M, Hire, \alpha) = 12\alpha[e_H] + 10\alpha[e_M] - 4,$$

so $\alpha[e_H] \geq 1/3 - 5/6\alpha[e_M]$ whenever $u_1(\theta_M, Hire, \alpha) \geq 0$. Additionally,

$$u_1(\theta_L, Hire, \alpha) = 5\alpha[e_H] + 2\alpha[e_M] - 1,$$

which is strictly positive whenever $\alpha[e_H] \geq \min\{5/21 - 6/21\alpha[e_M], 1/3 - 5/6\alpha[e_M]\}$ and $\alpha[e_M] \leq 1/4$. We argue that such steady state profiles exist in the iterated limit where $\gamma_1 \to 1$, then $\delta \to 1$, and then $\gamma_2 \to 1$, and that the corresponding aggregate probability that any type plays $Hire$ converges to 0.

Let $\chi : \Delta(A) \rightrightarrows \Delta(A)$ be the correspondence given by

$$\chi(\alpha) = \begin{cases} \{\alpha\} & \text{if } \alpha[e_M] < \frac{1}{4} \\ \{\alpha' \in \Delta(A) : \alpha'[e_M] = \frac{1}{4}\} & \text{if } \alpha[e_M] \geq \frac{1}{4} \end{cases},$$

and let $\rho : \Pi_2 \rightrightarrows \Pi_2$ be the correspondence given by

$$\rho(\pi_2) = \{\pi_2' \in \Pi_2 : \pi_2'(\cdot|Hire, m) \in \chi(\pi_2(\cdot|Hire, m)) \ \forall m \in M\}.$$

Note that $\rho$ is upper hemicontinuous, convex-valued, and coincides with the identity correspondence whenever $\pi_2(e_M|In, m) < 1/4$ for all $m$. Let $\upsilon : \Pi_1 \rightrightarrows \Pi_1$ be the correspondence given by

$$\begin{aligned} \upsilon(\pi_1) = \Bigg\{\pi_1' \in \Pi_1 : &(1)\ \pi_1'[Hire, m|\theta] = \min\left\{\pi_1[Hire, m|\theta], \frac{\lambda(\theta_L)}{2\lambda(\theta)}\right\} \ \forall m \in M,\ \theta \in \{\theta_H, \theta_M\}, \\ &(2)\ \pi_1'[Pass, m|\theta] = \pi_1[Hire, m|\theta] \ \forall m \neq m_{Pass,\theta_H},\ \theta \in \{\theta_H, \theta_M\}, \\ &(3)\ \pi_1'[s, m|\theta_L] = \pi_1[s, m|\theta_L] \ \forall s \in \{Hire, Pass\},\ m \in M, \Bigg\}. \end{aligned}$$

Note that $\upsilon$ is upper hemicontinuous, convex-valued, and coincides with the identity correspondence whenever $\pi_1(Hire, m|\theta) < \lambda(\theta_L)/(2\lambda(\theta))$ for all $m \in M$ and $\theta \in \{\theta_H, \theta_M\}$.

Consider the correspondence $\widetilde{\mathscr{R}}^{\delta,\gamma_1,\gamma_2} : \Pi_1 \times \Pi_2 \rightrightarrows \Pi_1 \times \Pi_2$ given by $\widetilde{\mathscr{R}}^{\delta,\gamma_1,\gamma_2}(\pi_1, \pi_2) = \{(\pi_1', \pi_2') \in \Pi_1 \times \Pi_2 : \pi_1' = \upsilon(\mathscr{R}_1^{\delta,\gamma_1}(\pi_2))$ and $\pi_2' \in \rho(\mathscr{R}_2^{\gamma_2}(\pi_1))\}$. Since $\widetilde{\mathscr{R}}^{\delta,\gamma_1,\gamma_2}$ is upper hemicontinuous and convex-valued, Kakutani's fixed point theorem guarantees the existence of a fixed point $(\pi_1^{\delta,\gamma_1,\gamma_2}, \pi_2^{\delta,\gamma_1,\gamma_2})$.

We establish that $\lim_{\gamma_2\to 1}\lim_{\delta\to 1}\lim_{\gamma_1\to 1}\pi_1^{\delta,\gamma_1,\gamma_2}[Hire|\theta] = 0$ for $\theta \in \{\theta_H, \theta_M\}$. Suppose towards a contradiction that there is a sequence of worker continuation probabilities $\{\gamma_{2,j}\}_{j\in\mathbb{N}}$, a collection of sequences of firm discount factors $\{\delta_{j,k}\}_{j,k\in\mathbb{N}}$, and a collection of sequences of firm continuation probabilities $\{\gamma_{1,j,k,l}\}_{j,k,l\in\mathbb{N}}$ such that (a) $\lim_{j\to\infty}\gamma_{2,j} = 1$, (b) $\lim_{k\to\infty}\delta_{j,k} = 1$ for all $j$, (c) $\lim_{l\to\infty}\gamma_{1,j,k,l} = 1$ for all $j, k$, (d) $\lim_{j\to\infty}\lim_{k\to\infty}\lim_{l\to\infty}\pi_1^{\delta_{j,k},\gamma_{1,j,k,l},\gamma_{2,j}}[Hire, m|\theta]$ exists for all $\theta \in \Theta$ and $m \in M$, and (e) $\lim_{j\to\infty}\lim_{k\to\infty}\lim_{l\to\infty}\pi_1^{\delta_{j,k},\gamma_{1,j,k,l},\gamma_{2,j}}[Hire|\theta] > 0$ for either $\theta = \theta_H$ or $\theta = \theta_M$. Then since $\pi_2^{\delta,\gamma_1,\gamma_2}(e_M|Hire, m) \leq 1/4$ for all $m \in M$, Lemma B1 implies that $\lim_{j\to\infty}\lim_{k\to\infty}\lim_{l\to\infty}\pi_1^{\delta_{j,k},\gamma_{1,j,k,l},\gamma_{2,j}}[Hire|\theta_L] = 1$. Therefore, there exists some $m \in M$ such that $\lim_{j\to\infty}\lim_{k\to\infty}\lim_{l\to\infty}\pi_1^{\gamma_{1,j,k,l},\gamma_{2,j}}[Hire, m|\theta_L] > 0$ and

$$\lim_{j\to\infty}\lim_{k\to\infty}\lim_{l\to\infty}\pi_1^{\delta_{j,k},\gamma_{1,j,k,l},\gamma_{2,j}}[Hire|\theta_L] \geq \frac{\lambda(\theta_L)}{2\lambda(\theta)}\lim_{j\to\infty}\lim_{k\to\infty}\lim_{l\to\infty}\pi_1^{\delta_{j,k},\gamma_{1,j,k,l},\gamma_{2,j}}[Hire|\theta]$$

for both $\theta \in \{\theta_H, \theta_M\}$. By Lemma 2 and the fact that the unique worker best response to $Hire$ is $e_L$ when the probability the type is $\theta_L$ is at least 1/2, this implies that $\lim_{j\to\infty}\lim_{k\to\infty}\lim_{l\to\infty}\mathscr{R}_2^{\gamma_{2,j}}(\pi_1^{\delta_{j,k},\gamma_{1,j,k,l},\gamma_{2,j}})(e_L|Hire, m) = 1$. Since $\chi(\pi_2(\cdot|Hire, m)) = \{\pi_2(\cdot|Hire, m)\}$ if $\pi_2(e_M|Hire, m) < 1/4$, it follows that $\lim_{j\to\infty}\lim_{k\to\infty}\lim_{l\to\infty}\pi_2^{\delta_{j,k},\gamma_{1,j,k,l},\gamma_{2,j}}(\pi_1)(e_L|Hire, m) = 1$. However, by Lemma B1, this requires that $\lim_{j\to\infty}\lim_{k\to\infty}\lim_{l\to\infty}\pi_1^{\delta_{j,k},\gamma_{1,j,k,l},\gamma_{2,j}}[Hire, m] = 0$ must hold, a contradiction.

A similar argument establishes that $\lim_{\gamma_2\to 1}\lim_{\delta\to 1}\lim_{\gamma_1\to 1}\pi_1^{\delta,\gamma_1,\gamma_2}[Hire|\theta_L] = 0$, so $\lim_{\gamma_2\to 1}\lim_{\delta\to 1}\lim_{\gamma_1\to 1}\pi_1^{\delta,\gamma_1,\gamma_2}[Hire] = 0$. Since a worker will only play $e_M$ in response to some $(Hire, m)$ if they have previously encountered a firm playing $(Hire, m)$, we have that $\mathscr{R}_2^{\gamma_{2,j}}(\pi_1^{\gamma_{1,k,l},\gamma_{2,k}})(e_M|Hire,m) < 1/4$ for all $m \in M$ in the iterated limit. Since $\rho(\pi_2) = \{\pi_2\}$ if $\pi_2(e_M|Hire,m) < 1/4$ for all $m$, $\pi_2^{\delta,\gamma_1,\gamma_2} = \rho(\mathscr{R}_2^{\gamma_2}(\pi_1^{\delta,\gamma_1,\gamma_2})) = \mathscr{R}_2^{\gamma_2}(\pi_1^{\delta,\gamma_1,\gamma_2})$ for fixed, sufficiently high $\gamma_2 \in [0,1)$ when $\delta$ is sufficiently close to 1 and, given $\delta$, $\gamma_1$ is sufficiently close to 1. For similar reasons, $\pi_1^{\delta,\gamma_1,\gamma_2} = \upsilon(\mathscr{R}_1^{\delta,\gamma_1}(\pi_2^{\delta,\gamma_1,\gamma_2})) = \mathscr{R}_2^{\gamma_2}(\pi_1^{\delta,\gamma_1,\gamma_2})$ also holds in the iterated limit. Thus, $(\pi_1^{\delta,\gamma_1,\gamma_2}, \pi_2^{\delta,\gamma_1,\gamma_2})$ is a fixed point of $\mathscr{R}^{\delta,\gamma_1,\gamma_2}$ for fixed, sufficiently high $\gamma_2 \in [0,1)$, when $\delta$ is sufficiently close to 1 and, given $\delta$, $\gamma_1$ is sufficiently close to 1. We conclude that there are stable profiles in which every type plays $Pass$. ■

### OA.6.2 Analysis of Example 3

**Proposition OA 5.** *The game in Example 3 has stable profiles where both types play Out with probability* 1.

*Proof.* We specify that the receiver prior $g_2$ is a Dirichlet distribution with initial weight 1 on $(\theta_1, In, m_{In,\theta_1})$ and 1/2 on $(\theta_2, In, m_{In,\theta_1})$, and, for all other messages $m \neq m_{In,\theta_1}$, initial weight 1/2 on $(\theta_1, In, m)$ and 1 on $(\theta_2, In, m)$. This means that initial trust is satisfied: When a receiver first encounters a sender who plays $(In, m_{In,\theta})$, the probability they place on the receiver having type $\theta$ is 2/3 so $BR(\theta, In)$ is optimal.

We claim first that if a receiver has encountered past plays 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 $BR(\theta, In)$. We demonstrate this for the case $m = m_{\theta_1}$; analogous arguments handle the other case. If this message has only ever been sent by $\theta_1$, the receiver's belief about the sender's type after $(In, m_{\theta_1})$ must put probability at least $(1+1)/(1+1+.5) = 4/5$ on $\theta_1$, which makes $a_1$ the unique receiver best response. When $\theta = \theta_2$, the receiver's conditional distribution over the sender's type after $(In, m_{\theta_1})$ must put probability at least $(1+.5)/(1+1+.5) = 3/5$ on $\theta_2$, which makes $a_2$ the unique receiver best response.

We focus on steady state profiles in which, for every $m \in M$, the aggregate probability that a receiver responds to $(In, m)$ with $a_3$ is less than 1/4. Under such responses, it can never be weakly optimal for both types to play $In$ with the same message. To see this, note that

$$u_1(\theta_1, In, \alpha) + u_2(\theta_2, In, \alpha) = -\alpha[a_1] - \alpha[a_2] + 2\alpha[a_3] = -1 + 3\alpha[a_3],$$

which is strictly negative whenever $\alpha[a_3] \leq 1/4$. We argue that such steady state profiles exist in the iterated limit where $\gamma_1 \to 1$ then $\delta \to 1$ then $\gamma_2 \to 1$ and that the corresponding aggregate probability that either sender type plays $In$ converges to 0.

Let $\chi : \Delta(A) \rightrightarrows \Delta(A)$ be the correspondence given by

$$\chi(\alpha) = \begin{cases} \{\alpha\} & \text{if } \alpha[a_3] < \frac{1}{4} \\ \{\alpha' \in \Delta(A) : \alpha'[a_3] = \frac{1}{4}\} & \text{if } \alpha[a_3] \geq \frac{1}{4} \end{cases},$$

and let $\rho : \Pi_2 \rightrightarrows \Pi_2$ be the correspondence given by

$$\rho(\pi_2) = \{\pi_2' \in \Pi_2 : \pi_2'(\cdot|In, m) \in \chi(\pi_2(\cdot|In, m))\ \forall m \in M\}.$$

Note that $\rho$ is upper hemicontinuous, convex-valued, and coincides with the identity correspondence whenever $\pi_2(a_3|In, m) < 1/4$ for all $m$.

Consider the correspondence $\widetilde{\mathscr{R}}^{\delta,\gamma_1,\gamma_2} : \Pi_1 \times \Pi_2 \rightrightarrows \Pi_1 \times \Pi_2$ given by $\widetilde{\mathscr{R}}^{\delta,\gamma_1,\gamma_2}(\pi_1, \pi_2) = \{(\pi_1', \pi_2') \in \Pi_1 \times \Pi_2 : \pi_1' = \mathscr{R}_1^{\delta,\gamma_1}(\pi_2) \text{ and } \pi_2' \in \rho(\mathscr{R}_2^{\gamma_2}(\pi_1))\}$. Since $\mathscr{R}$ is upper hemicontinuous and convex-valued, Kakutani's fixed point theorem guarantees the existence of a fixed point $(\pi_1^{\delta,\gamma_1,\gamma_2}, \pi_2^{\delta,\gamma_1,\gamma_2})$. As $\pi_2^{\delta,\gamma_1,\gamma_2}(a_3|s, m) \leq 1/4$ for all $(s, m)$ by construction, Lemma B1 implies that, for all $\gamma_2 \in [0, 1)$ and $(s, m)$, either $\lim_{\gamma_1 \to 1} \pi_1^{\delta,\gamma_1,\gamma_2}[In, m|\theta_1] = 0$ or $\lim_{\gamma_1 \to 1} \pi_1^{\delta,\gamma_1,\gamma_2}[In, m|\theta_2] = 0$. This means that, as $\gamma_1 \to 1$ then $\delta \to 1$, the probability that a receiver encounters senders with both types that pair $In$ with the same message $m$ approaches 0. Since a receiver would only ever play $a_3$ in response to $(In, m)$ if they have previously encountered senders of both types play $(In, m)$, this means that $\lim_{\delta \to 1} \lim_{\gamma_1 \to 1} \mathscr{R}_2^{\gamma_2}(\pi_1^{\delta,\gamma_1,\gamma_2})(a_3|In, m) = 0$ for all $m \in M$. Since $\rho(\pi_2) = \{\pi_2\}$ if $\pi_2(a_3|In, m) < 1/4$ for all $m$, $\pi_2^{\delta,\gamma_1,\gamma_2} = \rho(\mathscr{R}_2^{\gamma_2})(\pi_1^{\delta,\gamma_1,\gamma_2}) = \mathscr{R}_2^{\gamma_2}(\pi_1^{\delta,\gamma_1,\gamma_2})$ for fixed

$\gamma_2 \in [0,1)$ when $\delta$ is sufficiently close to 1 and, given $\delta$, $\gamma_1$ is sufficiently close to 1. Thus, for fixed $\gamma_2 \in [0,1)$, $(\pi_1^{\delta,\gamma_1,\gamma_2}, \pi_2^{\delta,\gamma_1,\gamma_2})$ is a fixed point of $\mathscr{R}^{\delta,\gamma_1,\gamma_2}$ when $\delta$ is sufficiently close to 1 and, given $\delta$, $\gamma_1$ sufficiently close to 1.

To show that $\lim_{\gamma_2 \to 1} \lim_{\delta \to 1} \lim_{\gamma_1 \to 1} \pi_1^{\delta,\gamma_1,\gamma_2}[In] = 0$, suppose towards a contradiction that there is a sequence of receiver continuation probabilities $\{\gamma_{2,j}\}_{j\in\mathbb{N}}$, a collection of sequences of sender discount factors $\{\delta_{j,k}\}_{j,k\in\mathbb{N}}$, and a collection of sequences of sender continuation probabilities $\{\gamma_{1,j,k,l}\}_{j,k,l\in\mathbb{N}}$ such that (a) $\lim_{j\to\infty} \gamma_{2,j} = 1$, (b) $\lim_{k\to\infty} \delta_{j,k} = 1$ for all $j$, (c) $\lim_{l\to\infty} \gamma_{1,j,k,l} = 1$ for all $j,k$, and (d) $\lim_{j\to\infty} \lim_{k\to\infty} \lim_{l\to\infty} \pi_1^{\delta_{j,k},\gamma_{1,j,k,l},\gamma_{2,j}}[In, m|\theta] > 0$ for some $\theta \in \Theta$ and $m \in M$. Without loss of generality, take $\theta = \theta_1$. By what we have shown, it must be that $\lim_{k\to\infty} \lim_{l\to\infty} \pi_1^{\delta_{j,k},\gamma_{1,j,k,l},\gamma_{2,j}}[In, m|\theta_2] = 0$ for all sufficiently large $j$. Combining this with $\lim_{j\to\infty} \lim_{k\to\infty} \lim_{l\to\infty} \pi_1^{\delta_{j,k},\gamma_{1,j,k,l},\gamma_{2,j}}[In, m|\theta_1] > 0$ and $\lim_{j\to\infty} \gamma_{2,j} = 1$ gives $\lim_{j\to\infty} \lim_{k\to\infty} \lim_{l\to\infty} \pi_2^{\delta_{j,k},\gamma_{1,j,k,l},\gamma_{2,j}}(a_1|s,m) = 1$, because with probability 1 every receiver encounters a type $\theta_1$ sender playing $(In, m)$ but never encounters a type $\theta_2$ sender playing $(In, m)$. However, since $u_1(\theta_1, In, a_1) < 0$, $\lim_{j\to\infty} \lim_{k\to\infty} \lim_{l\to\infty} \pi_2^{\delta_{j,k},\gamma_{1,j,k,l},\gamma_{2,j}}(a_1|s,m) = 1$ combined with Lemma B1 requires $\lim_{j\to\infty} \lim_{k\to\infty} \lim_{l\to\infty} \pi_1^{\delta_{j,k},\gamma_{1,j,k,l},\gamma_{2,j}}[In, m|\theta_1] = 0$, a contradiction. ■

### OA.6.3 Analysis of Example 4

**Proposition OA 6.** *The least-cost separating equilibrium of the game in Example 4 has $\theta = 1$ play $(s_1^*(1), s_2^*(1)) = (1/2, 0)$, to which the receiver responds $a^*(1) = 10$, $\theta = 2$ play $(s_1^*(2), s_2^*(2)) = (1/2, 5)$, to which the receiver responds $a^*(2) = 20$, and $\theta = 3$ play $(s_1^*(3), s_2^*(3)) = (1/2, 15)$, to which the receiver responds $a^*(3) = 30$.*

*Proof.* We first establish that this play is consistent with a separating PBE. Given an arbitrary $(s_1, s_2)$ and a belief $\tilde{\lambda}$ about the sender's type, the receiver's best responses are the closest actions to $20 s_1 \mathbb{E}_{\tilde{\lambda}}[\theta]$, as can be readily verified using the receiver's utility function. For $s_1 = 1/2$ and the belief that the type is $\theta$, the receiver's best response is $10\theta$, so the prescribed receiver play following the on-path sender play is indeed optimal.

Fix the receiver's response to any off-path signal-message pair $(s_1, s_2, m)$ to be $20s_1$, i.e. the best response under a belief putting probability 1 on $\theta = 1$. All that remains is to check that the incentives of the sender types are satisfied. We verify this for the $\theta = 3$ sender type. (Similar arguments handle the other two types.) Under the prescribed play, the payoff of the $\theta = 3$ sender type is $u_1(3, 1/2, 15, 30) = 30$. If the $\theta = 3$ sender were instead to mimic $\theta = 1$ or $\theta = 3$, their payoff would be 15 or 25, respectively. Moreover, if the $\theta = 3$ sender were to deviate to some off-path signal-message pair $(s_1, s_2, m)$, their payoff would be $60(1 - s_1)s_1 - s_2$, which is strictly lower than 30 for all $s_1 \in [0, 1]$ and $s_2 \geq 0$.

We now show that every other separating equilibrium results in (weakly) lower payoffs to each of the sender types. The payoff of the $\theta = 1$ sender from $(s_1, s_2)$ when the receiver responds with $20s_1$ is $20(1 - s_1)s_1 - s_2$, which attains its maximum value of 5 at $(s_1^*(1), s_2^*(1))$. The maximum possible payoff of the $\theta = 2$ sender from playing some $(s_1, s_2)$ when the receiver responds with $40s_1$, subject to the constraint that $\theta = 1$ would obtain a lower payoff than 5 by imitating $\theta = 2$ is

$$\max_{(s_1,s_2)\in S} 80(1 - s_1)s_1 - s_2 \text{ s.t. } 40(1 - s_1)s_1 - s_2 \leq 5.$$

The solution to this problem is $(s_1^*(2), s_2^*(2))$, and the resulting payoff to $\theta = 2$ is 15. Finally, the maximum possible payoff of the $\theta = 3$ sender from playing some $(s_1, s_2)$ when the receiver responds with $60s_1$, subject to the constraint that $\theta = 2$ would obtain a lower payoff than 15 by imitating $\theta = 3$ is

$$\max_{(s_1,s_2)\in S} 120(1 - s_1)s_1 - s_2 \text{ s.t. } 80(1 - s_1)s_1 - s_2 \leq 15.$$

The solution to this problem is $(s_1^*(3), s_2^*(3))$. ■

**Proposition OA 7.** *If $\pi$ is a JCE in the game in Example 4, then each $\theta$ plays $(s_1^*(\theta), s_2^*(\theta))$ with strictly positive probability, and the receiver responds to all on-path $(s_1^*(\theta), s_2^*(\theta), m)$ with $a^*(\theta)$ as in the least-cost separating equilibrium.*

*Proof.* We first establish that in a JCE $\pi$, for each signal-message pair $(s_1, s_2, m)$ played by $\theta = 3$, the product of $(1 - s_1)$ and the receiver's response has expected value at least $44/3$. Suppose otherwise that there is some signal-message pair $(s_1, s_2, m)$ that $\theta = 3$ plays which induces a receiver response with expected value $\tilde{a}$ such that $(1 - s_1)\tilde{a} < 44/3$. It must be that $s_2 < 44$, as otherwise $\theta = 3$ would obtain a strictly negative payoff. Thus, $s_2' = \lceil s_2 + 30 - 2(1 - s_1)\tilde{a} \rceil \in S$. Note that $u_1(3, \pi) = 3(1 - s_1)\tilde{a} - s_2$, while $u_1(\theta, \pi) \leq \theta(1 - s_1)\tilde{a} - s_2$ for $\theta \in \{1, 2\}$. Since $u_1(3, 1/2, s_2', a) = 3a/2 - s_2'$, we have that $u_1(3, 1/2, s_2', a) \geq u_1(3, \pi)$ if and only if $a \geq 2(1 - s_1)\tilde{a} + 2(s_2' - s_2)/3$, with the inequality strict for all $a > 2(1 - s_1)\tilde{a} + 2(s_2' - s_2)/3$. Moreover, $u_1(\theta, 1/2, s_2', a) \geq u_1(\theta, \pi)$ for $\theta = 1$ or $\theta = 2$ only if $u_1(\theta, 1/2, s_2', a) = \theta a/2 - s_2' \geq \theta(1 - s_1)\tilde{a} - s_2$, which requires $a \geq 2(1 - s_1)\tilde{a} + s_2' - s_2$. Since $s_2' > s_2$, $2(1 - s_1)\tilde{a} + s_2' - s_2 > 2(1 - s_1)\tilde{a} + 2(s_2' - s_2)/3$ which means that $\overline{\Theta}(1/2, s_2', \pi) = \{3\}$ and the only justified response to $(1/2, s_2')$ is 30. As this is strictly greater than $2(1 - s_1)\tilde{a} + 2s_2'/3 - 2s_2/3$ when $(1 - s_1)\tilde{a} < 44/3$, the claim follows.

An immediate implication is that there must be some signal-message pair that $\theta = 2$ sends with positive probability that $\theta = 3$ does not send, because $(1 - s_1)a \leq 25/2$ for any signal $(s_1, s_2)$ and receiver best response $a$ to a belief where the relative weight on $\theta = 2$ versus $\theta = 3$ is at least that of the prior.

We now show that, for each signal-message pair $(s_1, s_2, m)$ played by $\theta = 2$ but not by $\theta = 3$, the product of $1 - s_1$ and the receiver's response must have an expected value between $19/2$ and $10$. Whenever the probability of $\theta = 3$ is 0, the product of $(1 - s_1)$ and any undominated receiver response is no more than 10, so we need only show that the expected value of the product must exceed $19/2$. Suppose otherwise that there is some signal-message pair $(s_1, s_2, m)$ that $\theta = 2$ plays but $\theta = 3$ does not play for which the expected value of the receiver response $\tilde{a}$ satisfies $(1 - s_1)\tilde{a} < 19/2$. It must be that $s_2 < 19$, so $s_2' = \lceil s_2 + 10 - (1 - s_1)a \rceil \in S$. Note that $u_1(2, \pi) = 2(1 - s_1)\tilde{a} - s_2$, while $u_1(1, \pi) \leq (1 - s_1)\tilde{a} - s$. Since $u_1(2, 1/2, s_2', a) = a - s_2'$, we have that $u_1(2, 1/2, s_2', a) \geq u_1(2, \pi)$ if and only if $a \geq 2(1 - s_1)\tilde{a} + s_2' - s_2$, with the inequality strict for all $a > 2(1 - s_1)\tilde{a} + s_2' - s_2$. Moreover, $u_1(1, 1/2, s_2', a) \geq u_1(1, \pi)$

only if $a/2 - s_2' \geq (1-s_1)\tilde{a} - s$, which requires $a \geq 2(1-s_1)\tilde{a} + 2(s_2' - s_2)$. Since $s_2' > s_2$, $2(1-s_1)\tilde{a} + 2(s_2' - s_2) > 2(1-s_1)\tilde{a} + s_2' - s_2$, which means that $\overline{\Theta}(s+10, \pi) \subseteq \{2,3\}$ so justified responses to $(1/2, s_2')$ must weakly exceed 20. As this is strictly greater than $2(1-s_1)\tilde{a} + s_2' - s_2$ when $(1-s_1)\tilde{a} < 19/2$, the claim follows.

There must be some signal-message pair that only $\theta = 1$ plays. To see this, first observe that there can be no signal-message pair played by both $\theta = 1$ and $\theta = 3$. If there were some signal-message pair $(s_1, s_2, m)$ played by both $\theta = 1$ and $\theta = 3$, the product of $1 - s_1$ and the expected value of the receiver response $\tilde{a}$ must be less than $25/2$, because increasing differences in $\theta$ and $(1-s_1)a$ in the sender utility function implies that every signal-message pair played by $\theta = 2$ must induce the same expected value $(1-s_1)\tilde{a}$. This contradicts the fact that, for every signal-message pair played by $\theta = 3$, the product of $1 - s_1$ and the expected value of the receiver response must be weakly greater than $44/3$. Additionally, $\theta = 1$ cannot only play signal-message pairs that are also played by $\theta = 2$. Otherwise, there would be some signal-message pair $(s_1, s_2, m)$ played by $\theta = 2$, for which the product of $1 - s_1$ and the receiver response would have expected value weakly less than $15/2$ since $(1-s_1)a \leq 15/2$ for any receiver best response $a$ to a belief where the weight on $\theta = 3$ is 0 and the weight on $\theta = 1$ is at least that of the prior.

For every signal-message pair that only $\theta = 1$ plays, $s_1 = 1/2$, $s_2 = 0$, and the receiver responds with $a = 10$. The reason is the payoff $\theta = 1$ obtains from a signal-message pair $(s_1, s_2, m)$ that only $\theta = 1$ plays is $20(1-s_1)s_1 - s_2$, which is strictly less than 5 if $s_1 \neq 1/2$ or $s_2 > 0$. However, $\theta = 1$ can secure a payoff of 5 by simply playing $(s_1, s_2) = (1/2, 0)$, since every $a < 10$ is a strictly dominated response for the receiver.

We now argue that, for every signal-message pair played by $\theta = 2$ but not by $\theta = 3$, $s_1 = 1/2$, $s_2 = 5$, and the receiver responds with $a = 20$. We have previously established that the product of $1 - s_1$ and the expected value of the receiver's response $\tilde{a}$ must be between $19/2$ and 10. For $(1-s_1)\tilde{a} < 10$ to hold, it must be that $\theta = 1$ also plays this signal-message pair. This requires $u_1(1, s, \tilde{a}) = (1-s_1)\tilde{a} - s_2 = u_1(1, \pi)$. As previously established, $u_1(1, \pi) = 5$, so it must be that $s_2 = (1-s_1)\tilde{a} - 5$. However, there is no $\tilde{a} \in$

$[19/2, 10)$ such that $\tilde{a} - 5 \in S$. Therefore, $(1 - s_1)\tilde{a} = 10$. Since $(1 - s_1)\tilde{a} \leq 40(1 - s_1)s_1$ and $40(1 - s_1)s_1 < 10$ for all $s_1 \neq 1/2$, it follows that $s_1 = 1/2$ and thus $\tilde{a} = 20$. From $u_1(1, 1/2, s_2, 20) = 10 - s_2 \leq 5 = u_1(1, \pi)$, we obtain $s_2 \geq 5$. All that remains is to rule out $s_2 > 5$. If $s_2 > 5$, $u_1(1, 1/2, s_2 - 1, a) = a/2 - s_2 + 1 \geq 5 = u_1(1, \pi)$ only if $a \geq 20$. On the other hand, $u_1(2, 1/2, s_2 - 1, a) = a - s_2 + 1 \geq 20 - s_2 = u_1(2, \pi)$ if and only if $a \geq 19$, with the inequality strict for all $a > 19$. Thus, $\overline{\Theta}(1/2, s_2 - 1, \pi) \subseteq \{2, 3\}$, so justified responses to $(1/2, s_2 - 1)$ must weakly exceed 20. It follows that $s_2 = 5$.

Finally, we show that, for every signal-message pair played by $\theta = 3$, $s_1 = 1/2$, $s_2 = 15$, and the receiver responds with $a = 40$. We have previously established that the product of $1 - s_1$ and the expected value of the receiver's response $\tilde{a}$ must be between $44/3$ and 15. For $(1 - s_1)\tilde{a} < 15$ to hold, it must be that $\theta = 2$ also plays this signal-message pair. This requires $u_1(2, s_1, s_2, \tilde{a}) = 2(1 - s_1)\tilde{a} - s_2 = u_1(2, \pi)$. As previously established, $u_1(2, \pi) = 15$, so it must be that $s_2 = 2(1 - s_1)\tilde{a} - 15$. However, there is no $(1 - s_1)\tilde{a} \in [44/3, 15)$ such that $2(1 - s_1)\tilde{a} - 15 \in S$. Therefore, $(1 - s_1)\tilde{a} = 15$. Since $(1 - s_1)\tilde{a} \leq 60(1 - s_1)s_1$ and $60(1 - s_1)s_1 < 15$ for all $s_1 \neq 1/2$, it follows that $s_1 = 1/2$ and thus $\tilde{a} = 30$. From $u_1(2, 1/2, s_2, 30) = 30 - s_2 \leq 15 = u_1(2, \pi)$, we obtain $s_2 \geq 15$. All that remains is to rule out $s > 15$. If $s > 15$, $u_1(\theta, 1/2, s_2 - 1, a) = \theta a/2 - s + 1 \geq u_1(\theta, \pi)$ for either $\theta = 1$ or $\theta = 2$ requires that $a \geq 40$. On the other hand, $u_1(3, 1/2, s_2 - 1, a) = 3a/2 - s_2 + 1 \geq 45 - s_2 = u_1(3, \pi)$ if and only if $a \geq 29/3$, with the inequality strict for all $a > 29/3$. Thus, $\overline{\Theta}(1/2, s_2 - 1, \pi) = \{3\}$, so the only justified response to $(1/2, s_2 - 1)$ is 30. It follows that $s_2 = 15$. ■

# OA.7 Other Examples

## OA.7.1 Stability without Initially Trusting Receivers

*Example OA 1.* The sender's type space is $\Theta = \{\theta_1, \theta_2\}$, signal space is $S = \{In, Out\}$, and the receiver's action space is $A = \{a_1, a_2\}$. The payoffs to the sender and receiver are given below.

| $\theta_1$ | $a_1$ | $a_2$ |
|---|---|---|
| $In$ | $1,1$ | $-1,-1$ |
| $Out$ | $0,0$ | $0,0$ |

| $\theta_2$ | $a_1$ | $a_2$ |
|---|---|---|
| $In$ | $-1,-1$ | $-1,1$ |
| $Out$ | $0,0$ | $0,0$ |

$Out$ strictly dominates $In$ for type $\theta_2$, so $\theta_2$ plays $Out$ in every equilibrium of this game. However, there are equilibria in which $\theta_1$ plays $In$ and equilibria in which $\theta_1$ plays $Out$. The equilibria where $\theta_1$ plays $Out$ do not survive the Intuitive Criterion since $a_1$ is the receiver's unique best response to $In$ when the sender's type is $\theta_1$, and $\theta_1$ obtains a strictly higher payoff from $(In_1, a_1)$ than from playing $Out$.

We show that, when $g_2$ is such that a receiver plays $a_2$ when they first encounter a sender playing $(In, m)$ for every message $m \in M$, there are stable profiles in which $\theta_1$ plays $Out$.

We focus on steady state profiles in which the aggregate probability that a receiver responds to $(In, m)$ with $a_1$ is less than $1/3$ for every message $m \in M$, which makes it strictly optimal for type $\theta_1$ senders to play $Out$. We show that, for fixed $\gamma_2 \in [0, 1)$, such steady state profiles exist, and, moreover, that the corresponding aggregate probability that a type $\theta_1$ sender plays $In$ approaches 0 as $\gamma_1 \to 1$ and then $\delta \to 1$.

Let $\psi : \Pi_2 \to \Pi_2$ be the mapping given by

$$\psi(\pi_2)(a_1|In, m) = \min\left\{\pi_2(a_1|In, m), \frac{1}{3}\right\} \ \forall m \in M.$$

Note that $\psi$ is continuous and coincides with the identity mapping whenever $\pi_2(a_1|In, m) \leq 1/3$ for all $m$.

Consider the mapping $\widetilde{\mathscr{R}}^{\delta,\gamma_1,\gamma_2} : \Pi_1 \times \Pi_2 \to \Pi_1 \times \Pi_2$ given by $\widetilde{\mathscr{R}}^{\delta,\gamma_1,\gamma_2}(\pi_1, \pi_2) = (\mathscr{R}_1^{\delta,\gamma_1}(\pi_2), \psi(\mathscr{R}_2^{\gamma_2}(\pi_1)))$. Since $\widetilde{\mathscr{R}}^{\delta,\gamma_1,\gamma_2}$ is continuous, Brouwer's fixed point theorem guarantees the existence of a fixed point $(\pi_1^{\delta,\gamma_1,\gamma_2}, \pi_2^{\delta,\gamma_1,\gamma_2})$. As $\pi_2^{\delta,\gamma_1,\gamma_2}(a_1|In, m) \leq 1/3$ for all $m$ by construction, Lemma B1 implies that $\lim_{\delta\to 1} \lim_{\gamma_1 \to 1} \pi_1^{\delta,\gamma_1,\gamma_2}[In] = 0$ for all $\gamma_2 \in [0, 1)$. Furthermore, because $g_2$ is such that every receiver would play $a_2$ at a first encounter with a sender playing $(In, m)$, $\lim_{\delta\to 1}\lim_{\gamma_1\to 1}\mathscr{R}_2^{\gamma_2}(\pi_1^{\delta,\gamma_1,\gamma_2})(a_1|In, m) = 0$ for all $m$, $\gamma_2 \in [0,1)$, so the $\pi_2(a_1|In, m) \le 1/3$ constraint does not bind when $\delta$ is sufficiently close to 1 and, given $\delta$, $\gamma_1$ is sufficiently close to 1. Formally, since $\pi_2^{\delta,\gamma_1,\gamma_2} \neq \mathscr{R}_2^{\gamma_2}(\pi_1^{\delta,\gamma_1,\gamma_2})$ only if $\mathscr{R}_2^{\gamma_2}(\pi_1^{\delta,\gamma_1,\gamma_2})(a_1|In, m) > 1/3$ for some $m$, we have that, for fixed $\gamma_2 \in [0,1)$, $\pi_2^{\delta,\gamma_1,\gamma_2} = \mathscr{R}_2^{\gamma_2}(\pi_1^{\delta,\gamma_1,\gamma_2})$ for $\delta$ sufficiently close to 1 and, given $\delta$, $\gamma_1$ sufficiently close to 1. Combining this with the fact that $\pi_1^{\delta,\gamma_1,\gamma_2} = \mathscr{R}_1^{\delta,\gamma_1}(\pi_2^{\delta,\gamma_1,\gamma_2})$ for all $\gamma_1, \gamma_2 \in [0,1)$, it follows that, for fixed $\gamma_2 \in [0,1)$, $(\pi_1^{\delta,\gamma_1,\gamma_2}, \pi_2^{\delta,\gamma_1,\gamma_2})$ is a fixed point of $\mathscr{R}^{\delta,\gamma_1,\gamma_2}$ for $\delta$ sufficiently close to 1 and, given $\delta$, $\gamma_1$ sufficiently close to 1. Since $\lim_{\gamma_2\to 1}\lim_{\delta\to 1}\lim_{\gamma_1\to 1}\pi_1^{\delta,\gamma_1,\gamma_2}[In] = 0$, we conclude that there are stable profiles in which both types plays *Out*. $\square$

In this example, $In$ is strictly dominated for type $\theta_2$. If the priors of the receiver agents put 0 probability on sender types for whom a given signal is strictly dominated after an observation of that signal, the receivers would respond to $In$ with $a_1$, which would preclude the "All *Out*" equilibria. Depending on the context, such belief restrictions might be plausible, though they do rely on the receivers knowing the sender payoff function. However, even with such restrictions, stability can still allow implausible outcomes when initial trust is not satisfied. For example, we could modify the payoffs above so that $In$ is no longer strictly dominated for $\theta_2$, but rather conditionally dominated when the receiver response to *Out* uses a particular action, say $a_2$, with high probability. When the receiver priors are non-degenerate, we could choose the receiver payoffs so that both types playing *Out* is stable.[1]

### OA.7.2 Alternate Example Where D1 Does Not Imply JCE

*Example OA 2.* Here we analyze a simple example that is related to the idea of corporate culture as a way of telling workers what to do in unforeseen contingencies (see e.g. Camerer and Vepsalainen (1988) and Kreps (1990)). The sender is a firm, and the

[1]We could further restrict the receiver priors to assign probability 0 to sender types for whom a given signal is equilibrium dominated, but such restrictions are not consistent with a learning foundation for equilibrium, since they require that the receivers know the equilibrium being played in the population.

receiver is a recently hired worker. The firm's signal $s \in \{Creative, Standard\}$ is its choice of job assignment for the worker: The firm can either assign the worker to one of its "standard" jobs or to a "creative" job. Standard jobs carry out the firm's operation as currently designed, and let the firm effectively control the actions of workers through a combination of monitoring and provision of incentives. Creative jobs are intended to lead to innovations which the firm can then incorporate into its main operations, and the firm has relatively little direct control over the work these workers carry out. The worker's choice of action $a \in \{a_1, a_2, a_3\}$ represents the focus and intensity of their costly effort when assigned a creative job: $a_1$ and $a_2$ both represent intense effort directed at productive innovation but with focuses in different sectors, while $a_3$ represents a lack of productive effort.

The firm has three possible types, $\Theta = \{\theta_1, \theta_2, \theta_3\}$. Type $\theta_1$ and $\theta_2$ firms obtain higher payoffs than the relatively unproductive type $\theta_3$ firms. Moreover, type $\theta_1$ firms are particularly well suited to exploit innovations that workers with creative jobs choosing action $a_1$ may create, and type $\theta_2$ firms have an advantage with innovations from $a_2$. Due to their high payoffs from standard jobs, type $\theta_1$ gains relatively less from a worker with a creative job working on $a_2$ than type $\theta_3$ does. (Likewise for type $\theta_2$ and $a_1$.) A worker with a creative job is incentivized by rewards that come from successful innovation, so such a worker would like to take action $a_1$ if the firm has type $\theta_1$, $a_2$ if the firm has type $a_2$, and $a_3$ if the firm has type $\theta_3$.

The payoffs are given below.[2]

[2]The table indicates the worker can take any action in $\{a_1, a_2, a_3\}$ when assigned a standard job. However, we think of the firm as controlling the actual effort of a worker with a standard job, which is why the payoffs are independent of the formal action of a worker assigned a standard job.

| $\theta_1$ | $a_1$ | $a_2$ | $a_3$ |
|---|---|---|---|
| *Creative* | $4, 1$ | $2, 0$ | $0, -1$ |
| *Standard* | $2, 0$ | $2, 0$ | $2, 0$ |

| $\theta_2$ | $a_1$ | $a_2$ | $a_3$ |
|---|---|---|---|
| *Creative* | $2, 0$ | $4, 1$ | $0, -1$ |
| *Standard* | $2, 0$ | $2, 0$ | $2, 0$ |

| $\theta_3$ | $a_1$ | $a_2$ | $a_3$ |
|---|---|---|---|
| *Creative* | $1, 0$ | $1, 0$ | $-1, 1$ |
| *Standard* | $0, 0$ | $0, 0$ | $0, 0$ |

In every JCE, there is a positive probability of the worker being assigned a creative job. The reason is that the worker must, with positive probability, respond to *Creative* with $a_3$ in order to deter the firm from playing *Creative*, but there is no justified response to *Creative* that uses $a_3$, because $a_3$ is an optimal response to *Creative* only when the worker assigns a positive probability to the firm being type $\theta_3$. However, either $\theta_1$ or $\theta_2$ strictly prefers to play *Creative* whenever $\theta_3$ weakly prefers *Creative*, so $\theta_3$ is not a justified type for *Creative*.

Every stable profile has a positive probability of the worker being assigned a creative job because, for every firm type to learn that *Standard* is weakly optimal, the aggregate worker response must use $a_3$ with positive probability whenever *Creative* is played. Since responding to *Creative* with $a_3$ is optimal only for beliefs with positive probability on $\theta_3$, Initial Trust implies that some $\theta_3$ firms must be learning to play *Creative* while claiming to be either type $\theta_1$ or $\theta_2$. But if $\theta_3$ firms learn that it is weakly optimal to play *Creative*, then either the $\theta_1$ or $\theta_2$ firms learn that it is strictly optimal to do so.

Unlike JCE, many existing refinements allow equilibria in which all types play *Standard*. We discuss why this is the case for D1, which is typically thought of as a strong refinement. D1 allows the worker to respond to *Creative* with $a_3$, because there is no type which strictly prefers to play *Creative* whenever $\theta_3$ weakly prefers to do so. In particular, $\theta_3$ strictly prefers to play *Creative* whenever the worker plays either $a_1$ or $a_2$ with probability 1. For the other two types, there are some mixtures over $a_1$

and $a_2$ at which *Creative* is strictly preferred to *Standard* and others where *Standard* is strictly preferred to *Creative*. In contrast, $\theta_3$ is not a justified type for *Creative* because whenever $\theta_3$ weakly prefers to play *Creative*, there is some type that strictly prefers to do so, but this type need not be the same across worker responses. $\square$

# OA.8 Stability Under Alternative Assumptions

## OA.8.1 Weakening Initial Trust

Here we discuss a refinement satisfied by all stable profiles under an alternative assumption to initial trust. Suppose that receivers know the payoff functions of the senders, as in Fudenberg and He (2020). Then receivers who are long-lived may feel that they have acquired a good sense of each sender type's equilibrium payoff. Suppose that such a receiver encounters a sender playing a pair $(s, m_{s,\widetilde{\Theta}})$ that the receiver has not previously seen types outside of $\widetilde{\Theta}$ play. If the receiver believes that only types in $\widetilde{\Theta}$ could improve their outcome by deviating to $s$ when the receiver's response is contained in $BR(s, \widetilde{\Theta})$, we assume the receiver finds such a message credible and respond accordingly.[3]

As before, any stable profile must be a PBE-H. Moreover, stability also imposes additional conditions for profiles $\pi$ that are on-path strict for the receiver or are such that the sender types' payoffs would not be changed if the receiver deviated.[4] For such a profile to be stable, it must be that, for every signal $s$ where $u_1(\theta, s, a) < u_1(\theta, \pi)$ for all $a \in BR(\overline{\Theta}(s, \pi), s)$ and $\theta \notin \overline{\Theta}(s, \pi)$, there is some $m \in M$ such that $\pi_2(\cdot|s, m) \in \Delta(BR(\overline{\Theta}(s, \pi), s))$. Aside from the qualifying condition $u_1(\theta, s, a) < u_1(\theta, \pi)$ for all $a \in BR(\overline{\Theta}(s, \pi), s)$ and $\theta \notin \overline{\Theta}(s, \pi)$, this requirement is the same as Condition 2 of Definition 3. Combined, these conditions are weaker than JCE, so they are satisfied

[3]The receiver responding to "credible" statements in this way is similar to the motivation underlying "credible robust neologisms" in Clark (2020).

[4]These restrictions on $\pi$ guarantee that a typical receiver agent will learn the equilibrium payoffs of the sender types with high probability.

by the equilibria we focus on in Examples 3 and 2. The conditions coincide with JCE in Example 4 provided that the game is altered to have sufficiently fine action spaces. Unlike JCE, the conditions are satisfied by the D1 equilibrium in Example 1, but there are other games in which the conditions rule out D1 equilibria.

### OA.8.2 Strengthening Initial Trust

Suppose that we strengthen initial trust to require that for any $s \in S$ and $\widetilde{\Theta}, \widetilde{\Theta}' \subseteq \Theta$, if the receiver has never seen a type outside of $\widetilde{\Theta} \cup \widetilde{\Theta}'$ play $(s, m_{s,\widetilde{\Theta}})$, then their response to a first instance of $(s, m_{s,\widetilde{\Theta}})$ will belong to $BR(\widetilde{\Theta} \cup \widetilde{\Theta}', s)$. This means that a receiver who has only observed types in $\widetilde{\Theta}'$ deceitfully play $(s, m_{s,\widetilde{\Theta}})$ puts high probability on the sender type being in either $\widetilde{\Theta}$ or $\widetilde{\Theta}'$ after observing this signal-message pair. This seems plausible; however, we focus on initial trust because of JCE is simpler and easier to apply than its iterated version.

The stable profiles then satisfy an iterated version of JCE, which itself is stronger than the *Iterated Intuitive Criterion* (Cho and Kreps, 1987) and *co-divinity* (Sobel, Stole and Zapater, 1990). Moreover, it is not nested with NWBR, but it is weaker than the refinement obtained by iteratively applying NWBR.

Fix $s \in S$ and $\pi \in \Pi_1 \times \Pi_2$. Consider the following iterated version of the JCE procedure for computing the set of justified types. Initialize $\overline{\Theta}^0(s,\pi) = \overline{\Theta}(s,\pi)$. For $n \in \{1, 2, 3, ...\}$, let

$$\begin{aligned}
\widetilde{D}^n_\theta(s,\pi) &= \{\alpha \in \Delta(BR(\overline{\Theta}^{n-1}(s,\pi), s)) : u_1(\theta, s, \alpha) > u_1(\theta, \pi)\}, \\
\widetilde{D}^{0,n}_\theta(s,\pi) &= \{\alpha \in \Delta(BR(\overline{\Theta}^{n-1}(s,\pi), s)) : u_1(\theta, s, \alpha) = u_1(\theta, \pi)\}, \\
\Theta^{\dagger,n}(s,\pi) &= \{\theta \in \Theta : \widetilde{D}^n_\theta(s,\pi) \cup \widetilde{D}^{0,n}_\theta(s,\pi) \not\subseteq \cup_{\theta' \neq \theta} \widetilde{D}_{\theta'}(s,\pi)\}, \\
\overline{\Theta}^n(s,\pi) &= \begin{cases} \Theta^{\dagger,n}(s,\pi) & \text{if } \Theta^{\dagger,n}(s,\pi) \neq \emptyset \\ \overline{\Theta}^{n-1}(s,\pi) & \text{if } \Theta^{\dagger,n}(s,\pi) = \emptyset \end{cases}.
\end{aligned}$$

Set $\overline{\Theta}^\infty(s,\pi) = \cap_{n \in \mathbb{N}} \overline{\Theta}^n(s,\pi)$. Note that $\overline{\Theta}^{n+1}(s,\pi) \subseteq \overline{\Theta}^n(s,\pi)$ for all $n$ and that

$\overline{\Theta}^{\infty}(s,\pi) \subseteq \overline{\Theta}^{0}(s,\pi) = \overline{\Theta}(s,\pi)$.

Under this strengthening of initial trust, every stable profile $\pi$ must satisfy the following requirement: For every signal $s$, there is some $m \in M$ such that $\pi_2(\cdot|s,m) \in \Delta(BR(\overline{\Theta}^{\infty}(s,\pi),s))$. We refer to PBE-H that satisfy this requirement as *strongly justified communication equilibria*.

The proof proceeds by using similar arguments to the proof of Theorem 1 to inductively establish that $\pi_2(\cdot|s,m_{s,\overline{\Theta}^{\infty}(s,\pi)}) \in \Delta(BR(\overline{\Theta}^{n}(s,\pi),s))$ for all $n \in \mathbb{N}$.

### OA.8.3 Costs of Lying

Suppose that we allow the sender's utility function $u_1 : \Theta \times S \times M \times A \to \mathbb{R}$ to depend on the sender's message $m$ in the following way: For all $\theta \in \Theta$ and $\Theta', \Theta'' \subseteq \Theta$ such that $\theta \in \Theta' \cap \Theta''$, and $\Theta''' \subseteq \Theta$ such that $\theta \notin \Theta'''$, $u_1(\theta,s,m_{s,\Theta'},a) = u_1(\theta,s,m_{s,\Theta''},a) \geq u_1(\theta,s,m_{s,\Theta'''},a)$ for all $s \in S$ and $a \in A$. Here lying is weakly costly for the sender in that, for a given $s$ and $a$, the sender gets a lower payoff from a message that represents a set of types to which they do not belong. For simplicity, we assume that all messages that represent a set containing the true type give the sender the same payoff.

For each signal $s$, message $m$, and profile $\pi$, we will define a set of types $\overline{\Theta}(s,m,\pi)$ that is analogous to the set of justified types in our main setting where $m$ does not impact payoffs. To do this, first set

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

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

and

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

Then let

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

Under initial trust, a similar proof to that of Theorem 1 shows that any stable profile $\pi$ must satisfy the following requirement: $\pi_2(\cdot|s,m_{s,\overline{\Theta}(s,\Theta,\pi)})\in\Delta(BR(\overline{\Theta}(s,\Theta,\pi),s))$ for all $s\in S$. When the sender's message is payoff irrelevant, $\overline{\Theta}(s,m,\pi)=\overline{\Theta}(s,\pi)$, so this requirement implies Condition 2 of Definition 3. While lying costs make it less appealing for a non-justified type to falsely represent themself as justified, they can change the set of equilibria, so it is hard to give a precise summary of their effect in general games.

## OA.9 Stability Under a More General Limit

In this section, we study steady state aggregate play in the more general limit where first $\gamma_1$ tends to 1, and then $\delta$ and $\gamma_2$ tend to 1, without any restrictions on the relative speed with which $\delta$ and $\gamma_2$ converge. Formally, we consider $\lim_{(\delta,\gamma_2)\to(1,1)}\lim_{\gamma_1\to 1}\Pi^*(g,\delta,\gamma_1,\gamma_2)$. We will call these the *stable* profiles*.

**Definition OA 3.** *Strategy profile* $\pi$ *is* **stable*** *if there is a sequence* $\{\delta_j\}_{j\in\mathbb{N}}\to 1$*, sequence* $\{\gamma_{2,j}\}_{j\in\mathbb{N}}\to 1$*, and sequences* $\{\gamma_{1,j,k}\}_{j,k\in\mathbb{N}}$ *with* $\lim_{k\to\infty}\gamma_{1,j,k}=1$ *for all* $j$*, such that* $\pi=\lim_{j\to\infty}\lim_{k\to\infty}\pi_{j,k}$ *for some sequence* $\pi_{j,k}\in\Pi^*(g,\delta_{1,j},\gamma_{1,j,k},\gamma_{2,j})$.

Since every stable profile is also stable*, it follows that stable* profiles exist.

**Corollary OA 2.** *Stable* strategy profiles exist.*

As with stability, there is a strong relationship between the stable* profiles and the set of JCE.

**Definition OA 4.** *Strategy profile* $\pi$ *has* **strong incentives** *if, for every off-path* $s$ *and* $\theta\notin\overline{\Theta}(s,\pi)$*, there is some on-path* $(s',m')$ *such that* $u_1(\theta,s',a)>u_1(\theta,s,\pi_2(\cdot|s,m_{s,\overline{\Theta}(s,\pi)}))$

*for all $a \in BR(p_{(s',m')}, s')$, where $p_{(s',m')}$ is the posterior belief given $(s', m')$ obtained from $\pi_1$ and Bayes' rule.*

A strategy profile has strong incentives if for every off-path $s$, every type would obtain a strictly lower payoff from playing $(s, m_{s,\overline{\Theta}(s,\pi)})$ than they would from playing some on-path signal-message pair when the receiver responds with any best response to the corresponding posterior.

**Theorem OA 1.** *Suppose that the density of the prior of the sender agents is everywhere positive. If $\pi$ is stable* and has strong incentives, then it is a JCE.*

Theorem OA 1 says that a profile with strong incentives can be stable* only if it is a JCE. The assumption of strong incentives is vacuous if all signals are played with positive probability in $\pi$. Also, note that $u_1(\theta, \pi) > u_1(\theta, s, m_{s,\overline{\theta}(s,\pi)})$ for an arbitrary signal $s$ and profile $\pi$ whenever $\theta \notin \overline{\Theta}(s, \pi)$. Thus, every profile that is on-path strict for the receiver has strong incentives.[5]

The remainder of this section is devoted to the proof of Theorem OA 1. The argument that every stable* profile is a PBE-H proceeds very similarly to that for the stable profiles. The following lemma affirms the optimality of the aggregate sender play given the aggregate receiver play.

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

**Lemma OA 2.** *Suppose that $\pi$ is stable*. Then for any sender signal-message pair $(s, m) \in S \times M$ 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 $\{\pi_1(s, m|\theta)\}_{\theta \in \Theta}$.*

[5]Another sufficient condition is that no sender type would be hurt if the receiver were to change their response to some on-path signal-message pair, as is the case when all types choose an "exit" option.

We omit the proofs of Lemma OA 1 and Lemma OA 2, which are quite similar to the proofs of Lemma 1 and Lemma 2, respectively.

Lemma OA 3 below shows that when $\pi$ is a stable profile that has strong incentives, the aggregate receiver response to any $(s, m_{s,\overline{\Theta}(s,\pi)})$ must be supported on $BR(\overline{\Theta}(s,\pi), s)$.

**Lemma OA 3.** *Suppose that $\pi$ is stable* and has strong incentives. Then $\pi_2(\cdot|s, m_{s,\overline{\Theta}(s,\pi)}) \in \Delta(BR(\overline{\Theta}(s,\pi), s))$ for all $s \in S$.*

We prove Lemma OA 3 in the following subsection, but first we use Lemmas OA 1, OA 2, and OA 3 to prove Theorem OA 1.

*Proof of Theorem OA 1.* Lemma OA 1 implies Condition 1 of the definition of PBE-H, and Lemma OA 2 implies Condition 2. As before, Condition 3 of Definition 1 follows from the fact that the receivers in our model myopically optimize. Finally, the additional condition in Definition 3 follows from Lemma OA 3 and the assumption that $\pi$ has strong incentives. ■

### OA.9.1 Proof of Lemma OA 3

The following lemma relates the receiver's continuation parameter to the probability the aggregate receiver response to any on-path signal-message pair places on the corresponding receiver best responses.

**Lemma OA 4.** *Fix a strategy profile $\pi$. Let $X^{on}$ be the set of sender signal-message pairs that are on-path under $\pi_1$, and let $p_{(s,m)}$ be the posterior belief given $(s,m) \in X^{on}$ that is obtained from $\pi_1$ and Bayes' rule. There are $\nu, \eta > 0$ such that, for every $\pi_1' \in \Pi_1$ satisfying $\max_{(\theta,s,m)\in\Theta\times S\times M} |\pi_1'(s,m|\theta) - \pi_1(s,m|\theta)| < \nu$ and all $\delta, \gamma_1, \gamma_2 \in [0,1)$,*

$$\mathscr{R}_2^{\gamma_2}(\pi_1')(BR(p_{(s,m)}, s)|(s,m)) \geq 1 - \eta(1-\gamma_2)$$

*for all $(s,m) \in X^{on}$.*

*Proof.* Let $q(\theta, s, m) = \lambda(\theta)\pi_1(s, m|\theta)$ be the distribution over sender types, signals, and messages induced by $\lambda$ and $\pi_1$. For $\varepsilon > 0$, let $Q_\varepsilon = \{q' \in \Delta(\Theta \times S \times M) : \max_{(\theta,s,m)\in\Theta\times S\times M} |q'(\theta, s, m) - q(\theta, s, m)| \le \varepsilon\}$. Because best responses are upper hemi-continous, there exists $\varepsilon > 0$ such that every receiver whose belief $\widetilde{g}_2 \in \Delta(\Delta(\Theta\times S\times M))$ puts probability at least $1 - \varepsilon$ on $Q_\varepsilon$ will respond to every $(s, m) \in X^{\text{on}}$ with some action belonging to $BR(p_{(s,m)}, s)$.

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 $\varepsilon/3$ distance (in the sup-norm metric) of the empirical distribution they have observed.

We provide a lower bound on the share of receivers who have lived more than $T$ periods and who have observed an empirical distribution within $\varepsilon/3$ distance of the true distribution $q' \in \Delta(\Theta \times S \times M)$. By Hoeffding's inequality, the probability that the fraction of $(\theta, s, m)$ observations is outside of $[q'(\theta, s, m) - \varepsilon/3, q'(\theta, s, m) + \varepsilon/3]$ for a receiver with $t$ observations is less than $2e^{-\frac{2\varepsilon^2}{9}t}$, so the probability that the empirical distribution of a receiver with $t$ observations is greater than $\varepsilon/3$ distance from $q'$ is no more than $2|S||M|e^{-\frac{2\varepsilon^2}{9}t}$. Thus, the share of receivers who have lived longer than $T$ periods and who have observed an empirical distribution within $\varepsilon/3$ distance of $q'$ is at least

$$\begin{aligned}
\sum_{t=T}^{\infty}(1-\gamma_2)\gamma_2^t\left(1 - 2|S||M|e^{-\frac{2\varepsilon^2}{9}t}\right) &= \gamma_2^T - \frac{2|S||M|(1-\gamma_2)\gamma_2^T e^{-\frac{2\varepsilon^2}{9}T}}{1-\gamma_2 e^{-\frac{2\varepsilon^2}{9}}}, \\
&= 1 - \left(\frac{1-\gamma_2^T}{1-\gamma_2} + \frac{2|S||M|\gamma_2^T e^{-\frac{2\varepsilon^2}{9}T}}{1-\gamma_2 e^{-\frac{2\varepsilon^2}{9}}}\right)(1-\gamma_2), \\
&\ge 1 - \left(T + \frac{2|S||M|}{1-e^{-\frac{2\varepsilon^2}{9}}}\right)(1-\gamma_2),
\end{aligned}$$

where the inequality follows from the facts that $(1-\gamma_2^T)/(1-\gamma_2) < T$ and $\gamma_2^T e^{-\frac{2\varepsilon^2 T}{9}}/(1-\gamma_2 e^{-\frac{2\varepsilon^2}{9}}) < 1/(1-e^{-\frac{2\varepsilon^2}{9}})$ for all $\gamma_2 \in [0, 1)$.

Let $\eta = T + 2|S||M|/(1-e^{-\frac{2\varepsilon^2}{9}})$, and let $\nu > 0$ be such that, for every $\pi_1' \in \Pi_1$ satisfying $\max_{(\theta,s,m)\in\Theta\times S\times M}|\pi_1'(s,m|\theta)-\pi_1(s,m|\theta)|<\nu$, the corresponding distribution over sender types, signals, and messages belongs to $Q_{\varepsilon/3}$. It follows from the arguments above that, for all $\pi_1'$ within $\nu$ distance (in the sup-norm metric) of $\pi_1$, the steady-state share of receivers who respond to each $(s,m)\in X^{\text{on}}$ with some element of $BR(p_{(s,m)},s)$ is at least $1-\eta(1-\gamma_2)$. ■

The next lemma builds on Lemma OA 4 to show that, in a sequence of steady states converging to a stable* profile with strong incentives, the ratio of the aggregate probability of a non-justified type playing $(s, m_{s,\overline{\Theta}(s,\pi)})$ to the expected lifetime of a receiver agent approaches 0.

**Lemma OA 5.** *Fix a stable* strategy profile $\pi$ with strong incentives. Let $\{\pi_{j,k}\in\Pi^*(g,\delta_j,\gamma_{1,j,k},\gamma_{2,j})\}_{j,k\in\mathbb{N}}$ be a sequence of steady state profiles such that $\lim_{j\to\infty}\lim_{k\to\infty}\pi_{j,k}=\pi$, where $\lim_{j\to\infty}\delta_j=1$, $\lim_{j\to\infty}\gamma_{2,j}=1$, and $\lim_{k\to\infty}\delta_{j,k}=1$ for all $j$. For every $\varepsilon>0$, there exists some $J\in\mathbb{N}$ and function $K:\mathbb{N}\to\mathbb{N}$ such that*

$$\pi_{1,j,k}(s,m_{s,\overline{\Theta}(s,\pi)}|\theta)\le\varepsilon(1-\gamma_{2,j})$$

*for all $s$, $\theta\notin\overline{\Theta}(s,\pi)$, $j>J$, and $k>K(j)$.*

*Proof.* By Lemma OA 4 and the fact that $\lim_{j\to\infty}\lim_{k\to\infty}\pi_{j,k}=\pi$, there exists some $\eta>0$, $J'\in\mathbb{N}$, and function $K':\mathbb{N}\to\mathbb{N}$ such that

$$\pi_{2,j,k}(BR(p_{(s,m)},s)|(s,m))\ge 1-\eta(1-\gamma_{2,j})\tag{1}$$

for all $(s,m)$ on-path under $\pi_1$, $j>J'$, and $k>K'(j)$.

Fix a signal $s$ and type $\theta$ such that $\theta\notin\overline{\Theta}(s,\pi)$. Since $\pi$ has strong incentives, there is some $(s',m')$ that is on-path under $\pi_1$ such that $u_1(\theta,s',a)>u_1(\theta,s,\pi_2(\cdot|s,m_{s,\overline{\Theta}(s,\pi)}))$ for all $a\in BR(p_{(s',m')},s')$. For any $\alpha\in\Delta(A)$ and $z>0$, let $\mathcal{A}_{(\alpha,z)}=\{\alpha'\in\Delta(A):\max_{a\in A}|\alpha'[a]-\alpha[a]|\le z\}$ be the set of mixtures over $A$ that are no greater than $z$

away from $\alpha$ in the sup-norm metric. Let $\nu > 0$ be such that

$$(1-\nu)u_1(\theta, s', a) + \nu \min_{a' \in A} u_1(\theta, s', a') > u_1(\theta, s, \alpha) + \nu \tag{2}$$

for all $a \in BR(p_{(s',m')}, s')$ and $\alpha \in \mathcal{A}_{(\pi_2(\cdot|s, m_{s,\overline{\Theta}(s,\pi)}),\nu)}$.

Suppose that a sender has played $(s', m')$ at least $N > 0$ times. Combining Equation 1 with Lemma A.1 of Fudenberg and Levine (2006) implies that the probability that the fraction of times the sender observed a receiver play something outside of $BR(p_{(s',m')}, s')$ in response to $(s', m')$ exceeds $\nu/2$ is no more than $2^{11}\eta(1-\gamma_{2,j})/(3\nu^4 N)$. For a fixed $\varepsilon > 0$, let $N_{(s',m')}$ be such that $2^{11}\eta/(3\nu^4 N_{(s',m')}) < \varepsilon/4$. For such an $N_{(s',m')}$, it follows that $2^{11}\eta(1-\gamma_{2,j})/(3\nu^4 N_{(s',m')}) < \varepsilon(1-\gamma_{2,j})/4$.

By the assumption that the sender's prior has a density $g_1(\pi_2)$ that is everywhere positive and continuous in $\pi_2 \in \Pi_2$, we can find a lower bound on the probability that certain senders put on the receiver aggregate response to $(s', m')$ playing an element of $BR(p_{(s',m')}, s')$ with probability at least $1-\nu$. In particular, we will show there is a lower bound $\zeta > 0$ on the probability that the aggregate receiver response to $(s', m')$ puts probability at least $1-\nu$ on $BR(p_{(s',m')}, s')$ as determined by two classes of sender agents: (1) a sender agent who has played $(s', m')$ fewer than $N_{(s',m')}$ times, and (2) a sender agent who has played $(s', m')$ more than $N_{(s',m')}$ times and observed a response in $BR(p_{(s',m')}, s')$ greater than a fraction $1-\nu/2$ of the times. From the preceding paragraph, the share of sender agents who fall into either of these two classes exceeds $1-\varepsilon(1-\gamma_{2,j})/4$.

Consider a sender who, for each $a \in A$, has $n_a$ observations of a receiver responding to $(s', m')$ with $a$. Then such a sender puts probability at least

$$\frac{\min_{\pi_2 \in \Pi_2} g_1(\pi_2) \int_{\{\alpha \in \Delta(A): \alpha[BR(p_{(s',m')}, s')] \geq 1-\nu\}} \Pi_{a \in A} \alpha[a]^{n_a}}{\max_{\pi_2 \in \Pi_2} g_1(\pi_2) \int_{\Delta(A)} \Pi_{a \in A} \alpha[a]^{n_a}}$$

on the set of aggregate receiver responses to $(s', m')$ that have probability weakly greater than $1-\nu$ on $BR(p_{(s',m')}, s')$. This expression is uniformly bounded away from

0 when there are fewer than $N_{(s',m')}$ observations. Moreover, Theorem 4.2 of Diaconis and Freedman (1990) implies that this expression is uniformly bounded away from 0 when there are more than $N_{(s',m')}$ observations and the fraction of these observations where the receiver responding with some element of $BR(p_{(s',m')}, s')$ exceeds $1 - \nu/2$.

By similar arguments, there is some $N'_s \in \mathbb{N}$ such that, for a sender who has played $(s, m_{s,\overline{\Theta}(s,\pi)})$ at least $N'_s$ times, the sender's expectation of the aggregate receiver response to $(s, m_{s,\overline{\Theta}(s,\pi)})$ is within $\nu/3$ (in the sup-norm metric) of the empirical response the sender has observed. Moreover, by the law of large numbers, for any $j \in \mathbb{N}$, we can choose some $N'_{s,j} > N'_s$ to be such that there is a probability no greater than $\varepsilon(1 - \gamma_{2,j})/4$ that the empirical response to $(s, m_{s,\overline{\Theta}(s,\pi)})$ observed by a sender who has played $(s, m_{s,\overline{\Theta}(s,\pi)})$ at least $N'_{s,j}$ times is more than $\nu/3$ away from the aggregate receiver response $\pi_{2,j,k}(\cdot|s, m_{s,\overline{\Theta}(s,\pi)})$. Let $J'' \in \mathbb{N}$ and $K'' : \mathbb{N} \to \mathbb{N}$ be such that $\max_{a\in A} |\pi_{2,j,k}(a|s, m_{s,\overline{\Theta}(s,\pi)}) - \pi_2(a|s, m_{s,\overline{\Theta}(s,\pi)})| < \nu/3$ for all $j > J''$ and $k > K''(j)$. It follows that, for all such $j$ and $k$, the probability that $\mathcal{A}_{(\pi_2(\cdot|s,m_{s,\overline{\Theta}(s,\pi)}),\nu)}$ contains the expectation of the aggregate receiver response to $(s, m_{s,\overline{\Theta}(s,\pi)})$, as evaluated by a sender who has played $(s, m_{s,\overline{\Theta}(s,\pi)})$ at least $N'_{s,j}$ times, exceeds $1 - \varepsilon(1 - \gamma_{2,j})/4$.

Consider a sender belief $\widetilde{g}_1 \in \Delta(\Pi_2)$ that satisfies

$$\begin{aligned} \widetilde{g}_1(\pi_2(BR(p_{(s',m')}, s')|s', m') \geq 1 - \nu) &\geq \zeta, \\ \widetilde{g}_1(\pi_2(\cdot|s, m_{s,\overline{\theta}(s,\pi)}) \in \mathcal{A}_{(\pi_2(\cdot|s,m_{s,\overline{\Theta}(s,\pi)}),\nu)}) &\geq 1 - \frac{1}{2}\zeta. \end{aligned} \tag{3}$$

The first inequality says that the belief puts probability at least $\zeta$ on aggregate receiver responses to $(s', m')$ that play an element of $BR(p_{(s',m')}, s')$ with probability weakly greater than $1 - \nu$. The second inequality says that the belief puts probability at least $1-\zeta/2$ on the aggregate receiver response to $(s, m_{s,\overline{\Theta}(s,\pi)})$ belonging to $\mathcal{A}_{(\pi_2(\cdot|s,m_{s,\overline{\Theta}(s,\pi)}),\nu)}$. By Equation 2, all beliefs satisfying the conditions in (3) must put probability at least $\zeta/2$ on aggregate receiver behavior strategies where playing $(s', m')$ gives a type $\theta$ sender an expected payoff at least $\nu$ greater than that from playing $(s, m_{s,\overline{\Theta}(s,\pi)})$.

For a type $\theta$ sender with any belief that satisfies (3), the expected total lifetime payoff from the optimal policy exceeds the expected total lifetime payoff from only playing $(s, m_{s,\overline{\Theta}(s,\pi)})$ by an amount bounded away from 0 when $\delta$ and $\gamma_1$ are sufficiently high. In particular, for $\delta$ and $\gamma_1$ sufficiently close to 1, the difference in the expected payoff from the optimal policy and that from repeatedly playing $(s, m_{s,\overline{\Theta}(s,\pi)})$ exceeds $c = \zeta\nu/4 > 0$. Let $J''' \in \mathbb{N}$ and $K''' : \mathbb{N} \to \mathbb{N}$ be such that, whenever $j > J'''$ and $k > K'''(j)$, $\delta_j$ and $\gamma_{1,j,k}$ are sufficiently close to 1 so that this gap in the expected payoffs holds. Then, the version of Corollary 5.5 of Fudenberg and Levine (1993) presented in Fudenberg and He (2018) implies that, for every $j > J'''$, there is some $N''_{s,j}$ such that the share of type $\theta$ sender agents who have a belief satisfying the conditions in (3), have played $(s, m_{s,\overline{\Theta}(s,\pi)})$ more than $N''_{s,j}$ times, and are set to play $(s, m_{s,\overline{\Theta}(s,\pi)})$ in the current period is less than $\varepsilon(1 - \gamma_{2,j})/4$ for all $k > K'''(j)$.

Let $J = \max\{J', J'', J'''\}$, $K(j) = \max\{K'(j), K''(j), K'''(j)\}$ for all $j > J$, and $N_{s,j} = \max\{N'_{s,j}, N''_{s,j}\}$ for all $j > J$. Combining the preceding results shows that, when $j > J$ and $k > K(j)$, the share of type $\theta$ sender agents who have played $(s, m_{s,\overline{\Theta}(s,\pi)})$ more than $N_{s,j}$ times and are set to play $(s, m_{s,\overline{\Theta}(s,\pi)})$ in the current period is no more than $3\varepsilon(1 - \gamma_{2,j})/4$. Additionally, using the version of Lemma 5.7 of Fudenberg and Levine (1993) presented in Fudenberg and He (2018), it follows that, for all $j > J$, $K(j)$ can also be chosen so that $\pi_{1,j,k}(s, m_{s,\overline{\Theta}(s,\pi)}|\theta)$ exceeds the share of type $\theta$ sender agents who have played $(s, m_{s,\overline{\Theta}(s,\pi)})$ more than $N''_{s,j}$ times and are set to play $(s, m_{s,\overline{\Theta}(s,\pi)})$ in the current period by no more than $\varepsilon(1 - \gamma_{2,j})/4$ when $k > K(j)$. Thus, we conclude that $\pi_{1,j,k}(s, m_{s,\overline{\Theta}(s,\pi)}|\theta) \leq \varepsilon(1 - \gamma_2)$ for all $j > J$ and $k > K(j)$. ■

The proof of Lemma OA 3 uses Lemma OA 5 to show that, in a sequence of steady states converging to a stable* profile with strong incentives, the probability that a receiver encounters a non-justified sender type playing some $(s, m_{s,\overline{\Theta}(s,\pi)})$ over the course of their lifetime converges to 0. Initial trust then ensures that the aggregate receiver response to each $(s, m_{s,\overline{\Theta}(s,\pi)})$ is justified.

*Proof of Lemma OA 3.* Let $\{\pi_{j,k} \in \Pi^*(g, \delta_j, \gamma_{1,j,k}, \gamma_{2,j})\}_{j,k\in\mathbb{N}}$ be a sequence of steady state profiles such that $\lim_{j\to\infty} \lim_{k\to\infty} \pi_{j,k} = \pi$, where $\lim_{j\to\infty} \delta_j = 1$, $\lim_{j\to\infty} \gamma_{2,j} = 1$, and $\lim_{k\to\infty} \delta_{j,k} = 1$ for all $j$. By Lemma OA 5, for any $\varepsilon > 0$, there exists some $J \in \mathbb{N}$ and some function $K : \mathbb{N} \to \mathbb{N}$ such that $\pi_{1,j,k}(s, m_{s,\overline{\Theta}(s,\pi)}|\theta) \leq \varepsilon(1-\gamma_2)/\lambda(\theta)$ for all $\theta \notin \overline{\Theta}(s,\pi)$, $j > J$, and $k > K(j)$. Thus, when $j > J$ and $k > K(j)$, the probability that a receiver agent in a given period encounters a sender type outside of $\overline{\Theta}(s,\pi)$ playing $(s, m_{s,\overline{\Theta}(s,\pi)})$ is no greater than $\varepsilon(1-\gamma_{2,j})$. It follows that, when $j > J$ and $k > K(j)$, the probability that a receiver agent never encounters a sender type outside of $\overline{\Theta}(s,\pi)$ playing $(s, m_{s,\overline{\Theta}(s,\pi)})$ over the course of their lifetime is at least

$$\sum_{t=0}^{\infty}(1-\gamma_{2,j})\gamma_{2,j}^t(1-\varepsilon(1-\gamma_{2,j}))^t = \frac{1}{1+\gamma_{2,j}\varepsilon}.$$

Receivers who have never observed the signal-message pair $(s, m_{s,\overline{\Theta}(s,\pi)})$ played by a type outside of $\overline{\Theta}(s,\pi)$ would respond to this pair with an action belonging to $BR(\overline{\Theta}(s,\pi), s)$. Thus,

$$\pi_2(BR(\overline{\Theta}(s,\pi),s)|s, m_{s,\overline{\Theta}(s,\pi)}) = \lim_{j\to\infty}\lim_{k\to\infty}\pi_{2,j,k}(BR(\overline{\Theta}(s,\pi),s)|s, m_{s,\overline{\Theta}(s,\pi)}) \geq 1/(1+\varepsilon).$$

Since this holds for all $\varepsilon > 0$, we have that $\pi_2(BR(\overline{\Theta}(s,\pi),s)|s, m_{s,\overline{\Theta}(s,\pi)}) = 1$. ■

## OA.10 Details of Alternate Model

Consider a steady-state population of receivers who have geometric lifetimes with continuation probability $\gamma$, and are matched with a sender each period with i.i.d. probability $p$. We show that, when the receivers have expected lifespan $T = 1/(1-\gamma)$ and are expected to have $N_2 = pT$ matches over the course of their lifetime, the distribution of match experience in the receiver population is geometric with hit probability $\tilde{\gamma}_2 = (1-1/T)N_2/(1+(1-1/T)N_2)$. Because the aggregate play of receivers only depends on their experience, it follows that for every steady state in our main learning model given parameters $\gamma_1$, $\delta$, and $\gamma_2$, there is a steady state in this alternate model given parameters $\gamma = \gamma_1$, $\delta$, and $\tilde{\gamma}_2$ with the same aggregate strategy profile.

**Lemma OA 6.** *If receivers have geometric lifetimes with expected lifespan $T$ and are expected to have $N_2$ matches over the course of their lifetime, then the steady-state share of receivers who have previously been matched $n \in \mathbb{N}$ times is $(1-\tilde{\gamma}_2)\tilde{\gamma}_2^n$, where*

$$\tilde{\gamma}_2 = \frac{\left(1-\frac{1}{T}\right) N_2}{1+\left(1-\frac{1}{T}\right) N_2}.$$

*Proof.* Denote the steady-state share of receivers who have previously had $n$ matches by $\tilde{\mu}_2[n]$. We first derive $\tilde{\mu}_2[0]$. Since $1-\gamma$ is the share of newborn receivers and $\gamma(1-p)\tilde{\mu}_2[0]$ is the share of non-newborn receivers who have never been matched, it follows that $\tilde{\mu}_2[0] = (1-\gamma) + \gamma(1-p)\tilde{\mu}_2[0]$. Solving this gives

$$\tilde{\mu}_2[0] = \frac{1-\gamma}{1-\gamma+\gamma p}. \tag{OA 1}$$

Now we derive a recursive expression relating $\tilde{\mu}_2[n]$ to $\tilde{\mu}_2[n-1]$ for $n > 0$. Since $\gamma p \tilde{\mu}_2[n-1]$ is the share of receivers who in the previous period were matched for the $n$th time and $\gamma(1-p)\tilde{\mu}_2[0]$ is the share of receivers who have been matched $n$ times but were unmatched in the previous period, it follows that $\tilde{\mu}_2[n] = \gamma p \tilde{\mu}_2[n-1] + \gamma(1-p)\tilde{\mu}_2[n]$. Solving this gives

$$\tilde{\mu}_2[n] = \frac{\gamma p}{1-\gamma+\gamma p}\tilde{\mu}_2[n-1]. \tag{OA 2}$$

Combining Equations OA 1 and OA 2 gives

$$\tilde{\mu}_2[n] = \left(1-\frac{\gamma p}{1-\gamma+\gamma p}\right)\left(\frac{\gamma p}{1-\gamma+\gamma p}\right)^n.$$

Substituting $\gamma = 1 - 1/T$ and $p = N/T$ renders

$$\tilde{\mu}_2[n] = \left(1 - \frac{\left(1-\frac{1}{T}\right)N_2}{1+\left(1-\frac{1}{T}\right)N_2}\right)\left(\frac{\left(1-\frac{1}{T}\right)N_2}{1+\left(1-\frac{1}{T}\right)N_2}\right)^n$$

as desired. ■

# References


**Camerer, Colin, and Ari Vepsalainen.** 1988. "The Economic Efficiency of Corporate Culture." *Strategic Management Journal*, 9: 115–126.

**Cho, In-Koo, and David M. Kreps.** 1987. "Signaling Games and Stable Equilibria." *Quarterly Journal of Economics*, 102(2): 179–221.

**Clark, Daniel.** 2020. "Robust Neologism Proofness in Signaling Games." http://economics.mit.edu/files/16758.

**Diaconis, Persi, and David Freedman.** 1990. "On the Uniform Consistency of Bayes Estimates for Multinomial Probabilities." *The Annals of Statistics*, 18(3): 1317–1327.

**Fudenberg, Drew, and David K. Levine.** 1993. "Steady State Learning and Nash Equilibrium." *Econometrica*, 61(3): 547–573.

**Fudenberg, Drew, and David K. Levine.** 2006. "Superstition and Rational Learning." *American Economic Review*, 96(3): 630–651.

**Fudenberg, Drew, and Kevin He.** 2018. "Learning and Type Compatibility in Signaling Games." *Econometrica*, 86(4): 1215–1255.

**Fudenberg, Drew, and Kevin He.** 2020. "Payoff Information and Learning in Signaling Games." *Games and Economic Behavior*, 120: 96–120.

**Kreps, David M.** 1990. "Perspectives on Positive Political Economy." Chapter Corporate Culture and Economic Theory, 90–142. Cambridge University Press.

**Sobel, Joel, Lars Stole, and Iñigo Zapater.** 1990. "Fixed-Equilibrium Rationalizability in Signaling Games." *Journal of Economic Theory*, 52(2): 304–331.