Online Appendix for "Justified Communication Equilibrium"

Daniel Clark and Drew Fudenberg

July 24, 2020

OA.1 Formal Details Omitted from Section 2

OA.1.1 Strategy Mappings

The map $\sigma : \Delta(\mathcal{H}_1)^\Theta \times \Delta(\mathcal{H}_2) \to \Pi_1 \times \Pi_2$ taking the state in period t to the aggregate strategy profile has component mappings $\sigma_1 : \Delta(\mathcal{H}_1)^\Theta \to \Pi_1$ and $\sigma_2 : \Delta(\mathcal{H}_2) \to \Pi_2$ given by

$$\sigma_1(\mu_1)[s, m|\theta] = \sum_{h_1 : \mathbf{x}_\theta(h_1)=(s,m)} \mu_\theta[h_1] \text{ for all } s \in S, m \in M, \theta \in \Theta,$$
$$\sigma_2(\mu_2)[a|s, m] = \sum_{h_2 : \mathbf{y}(h_2)[s,m]=a} \mu_2[h_2] \text{ for all } a \in A, s \in S, m \in M.$$

OA.1.2 Update Rule

We define the various component mappings of the rule, $\mathbf{f} : \Delta(\mathcal{H}_1)^\Theta \times \Delta(\mathcal{H}_2) \to \Delta(\mathcal{H}_1)^\Theta \times \Delta(\mathcal{H}_2)$, taking the state in period t to the state in period $t+1$. The mapping $\mathbf{f}_\theta : \Delta(\mathcal{H}_1)^\Theta \times \Delta(\mathcal{H}_2) \to \Delta(\mathcal{H}_1)$ taking the state in period t to the distribution over type θ sender histories at period $t+1$ is given by

$$\mathbf{f}_\theta(\mu)[\emptyset] = 1 - \gamma,$$
$$\mathbf{f}_\theta(\mu)[(h_1, (s, m, a))] = \gamma \mu_\theta[h_1] \mathbb{1}_{\mathbf{x}_\theta^{-1}(s,m)}(h_1) \sigma_2(\mu)[a|s, m] \ \forall h_1 \in \mathcal{H}_1, s \in S, m \in M, a \in A,$$

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. Likewise, the mapping $\mathbf{f}_2 : \Delta(\mathcal{H}_1)^\Theta \times \Delta(\mathcal{H}_2) \to \Delta(\mathcal{H}_1)$ taking the state in period t to the distribution over receiver histories at period $t + 1$ is given by

$$\mathbf{f}_2(\mu)[\emptyset] = 1 - \gamma,$$
$$\mathbf{f}_2(\mu)[(h_2, (\theta, s, m))] = \gamma\mu_2[h_2]\lambda(\theta)\sigma_1(\mu)[s, m|\theta] \; \forall h_2 \in \mathcal{H}_2, \theta \in \Theta, s \in S, m \in M,$$

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 θ sender who plays (s, m).

OA.1.3 Aggregate Response Mappings

Here we define the aggregate response mappings, $\mathscr{R}_1^{\delta,\gamma_1,\gamma_2} : \Pi_2 \to \Pi_1$ and $\mathscr{R}_2^{\delta,\gamma_1,\gamma_2} : \Pi_1 \to \Pi_2$. To do so, we first define two mappings, $\mathscr{L}_1^{\delta,\gamma_1,\gamma_2} : \Pi_2 \to \Delta(\mathcal{H}_1)^\Theta$ and $\mathscr{L}_2^{\delta,\gamma_1,\gamma_2} : \Pi_1 \to \Delta(\mathcal{H}_2)$. $\mathscr{L}_1^{\delta,\gamma_1,\gamma_2}$ takes a fixed receiver aggregate behavior strategy and outputs the shares of the sender types with the various possible histories. For each $\theta \in \Theta$, given by

$$\mathscr{L}_\theta^{\delta,\gamma_1,\gamma_2}(\pi_2)[h_1] = \begin{cases} 1 - \gamma & \text{if } h_1 = \emptyset \\ (1-\gamma)\gamma^{|h_1|} \times_{t \le |h_1|} \mathbb{1}_{\mathbf{x}_\theta^{-1}(s,m)}((h_1(0), ..., h_1(t-1)))\pi_2[h_1(t)[A]|h_1(t)[X]] & \text{if } h_1 \neq \emptyset \end{cases},$$

where $|h_1|$ is the length of history $h_1 \in \mathcal{H}_1$, $h_1(t)$ denotes the t-th observation in history h_1, $h_1(t)[A]$ denotes the receiver action played in the t-th observation of history h_1, and $h_1(t)[X]$ denotes the sender signal-message pair played in the t-th observation of history h_1. $\mathscr{L}_2^{\delta,\gamma_1,\gamma_2}$ takes a fixed sender aggregate behavior strategy and gives the shares of receivers with the various possible histories according to

$$\mathscr{L}_2^{\delta,\gamma_1,\gamma_2}(\pi_1)[h_2] = \begin{cases} 1 - \gamma & \text{if } h_2 = \emptyset \\ (1-\gamma)\gamma^{|h_2|} \times_{t \le |h_2|} \lambda(h_2(t)[\Theta])\pi_1[h_2(t)[X]|h_2(t)[\Theta]] & \text{if } h_2 \neq \emptyset \end{cases},$$

where $|h_2|$ is the length of history $h_2 \in \mathcal{H}_2$, $h_2(t)$ denotes the t-th observation in history h_2, $h_2(t)[\Theta]$ denotes the sender type in the t-th observation of history h_2, and $h_2(t)[X]$ denotes the sender signal-message pair played in the t-th observation of history h_2.

$\mathscr{R}_1^{\delta,\gamma_1,\gamma_2}$ is then given by $\mathscr{R}_1^{\delta,\gamma_1,\gamma_2} = \sigma_1(\mathscr{L}_1^{\delta,\gamma_1,\gamma_2}(\pi_2))$, and $\mathscr{R}_2^{\delta,\gamma_1,\gamma_2}$ is given by $\mathscr{R}_2^{\delta,\gamma_1,\gamma_2} = \sigma_2(\mathscr{L}_2^{\delta,\gamma_1,\gamma_2}(\pi_2))$.

OA.2 Proof of Proposition 5

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

Proof. Because π is a uniformly justified JCE in a strictly monotonic signaling game, $\pi_2(\cdot|s,m) = \pi_2(\cdot|s,m')$ for all $s \in S$, $m, m' \in M$ such that $(s,m),(s,m') \in X^{on}$. Thus, for every $s \in S^{on}$, there is some $a_s \in A$ such that $\pi_2(a_s|s,m) = 1$ for all $(s,m) \in X^{on}$. For all $s \in S^{off}$, 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 each type $\theta \in \Theta$ plays s_θ 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 π.

Because π 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_θ paired with message m is strictly better for type θ than playing any other $s' \neq s_\theta$ paired with any other m'. Second, if $\pi_1(s_\theta, m|\theta) \geq 1-\varepsilon$ for every $\theta \in \Theta$, then, it is strictly optimal for the receiver to respond to (s,m) with a_s for every $s \in S^{on}$. Fix such an ε.

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 $z \geq 1$. ($\kappa(z) = \max\{\min\{z,1\},0\}$ is an example of such a κ.) 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 $s \in S$, $m \in M$. Note that ϕ 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 $\widetilde{\mathscr{R}}^{\delta,\gamma_1,\gamma_2}(\pi_1,\pi_2) = (\mathscr{R}^{\delta,\gamma_1,\gamma_2}(\pi_2), \phi(\pi_1, \mathscr{R}^{\delta,\gamma_1,\gamma_2}(\pi_1)))$. 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 π. 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 $s \in S$, $m \in M$. Lemma 5 thus requires that $\pi'_1(s_\theta|\theta) = 1$ for all $\theta \in \Theta$, which is a contradiction.

Next we show that $\pi'_2(a_s|s,m)$ for all $s \in S^{on}$ and $m \in M$ such that $\pi'_1(s,m|\theta) > 0$ for some $\theta \in \Theta$. Fix $s \in S^{on}$. 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 almost identical argument to the one which proves 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, α and α' must be optimal responses to s under the posterior distributions obtained by updating λ under Bayes' rule 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 5 implies that $\alpha = \alpha'$. Since $\pi'_2(\cdot|s,m)$ is the same for all $m \in M$ where there is some

$\theta \in \Theta$ such that $\pi_1(s, m|\theta) > 0$ and $\pi'_1(s_\theta|\theta) \geq 1 - \varepsilon$, 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 $\theta \in \Theta$.

We now show that, for all $\theta \in \Theta$, $\pi'_1(s|\theta) = 0$ for all $s \in S^{off}$. 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 5 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 5 requires that $u_1(\theta, s, \pi'_2(\cdot|s, m)) \leq u_1(\theta, \pi') = u_1(\theta, \pi)$ for all $\theta \in \Theta$, $s \in S$, $m \in M$. Now, suppose that there is some $s \in S^{off}$, $m \in M$ 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 θ 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 π 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 5, implies that $\pi'_1(s, m|\theta) = 0$ for all $\theta \in \Theta$, a contradiction.

We have shown that $\pi^{\delta,\gamma_1,\gamma_2}$ induces the same distribution over $\Theta \times S \times A$ as π 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 $\theta \in \Theta$, $\pi_2^{\delta,\gamma_1,\gamma_2} = \phi(\pi_1^{\delta,\gamma_1,\gamma_2}, \mathscr{R}^{\delta,\gamma_1,\gamma_2}(\pi_1^{\delta,\gamma_1,\gamma_2})) = \mathscr{R}^{\delta,\gamma_1,\gamma_2}(\pi_1^{\delta,\gamma_1,\gamma_2})$ 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 π' is a stable profile. ■

OA.3 Proof of Proposition 6

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

Proof. We specify that the marginal of the receiver prior g_2 on In 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 (θ_1, In, m) and 1 on (θ_2, In, m). This means that initial suggestibility is satisfied: When a receiver first encounters a sender who plays $(In, m_{In,\theta})$, the probability they place on the receiver having type θ is 2/3 so $BR(\theta_1, In)$ is optimal.

We claim first that if a receiver has encountered past play of (In, m) and all such plays have been by senders with the same type θ, then the receiver will respond to the next instance of (In, m) with $BR(\theta, In)$. We demonstrate this for the case $m = m_{\theta_1}$; analogous arguments handle the other cases. When $\theta = \theta_1$, the receiver's conditional distribution over the sender's type after (In, m_{θ_1}) must put probability at least 4/5 on θ_1, and $BR(p_1\theta_1 + (1 - p_1)\theta_2, In) = \{a_1\} = BR(\theta_1, In)$ if $p_1 \geq 4/5$. When $\theta = \theta_2$, the receiver's conditional distribution over the sender's type after (In, m_{θ_1}) must put probability at least 3/5 on θ_2, and $BR((1 - p_2)\theta_1 + p_2\theta_2, In) = \{a_2\} = BR(\theta_2, In)$ if $p_2 \geq 3/5$.

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

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

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 ρ is upper hemicontinuous and coincides with the identity correspondence whenever $\pi_2(a_3|In, m) \leq 1/4$ 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 $\widetilde{\mathscr{R}}^{\delta,\gamma_1,\gamma_2}(\pi_1, \pi_2) = \{(\pi_1', \pi_2') \in \Pi_1 \times \Pi_2 : \pi_1' = \mathscr{R}_1^{\delta,\gamma_1,\gamma_2}(\pi_2) \text{ and } \pi_2' \in \rho(\mathscr{R}_2^{\delta,\gamma_1,\gamma_2}(\pi_1))\}$. Since $\mathscr{R}$ is upper hemicontinuous, 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 5 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 $\lim_{\gamma \to 1} \mathscr{R}_2^{\delta,\gamma_1,\gamma_2}(\pi_1^{\delta,\gamma_1,\gamma_2})(a_3|In, m) = 0$ for all $m \in M$. Since $\rho(\pi_2) = \{\pi_2\}$ if $\pi_2(a_3|In, m) \leq 1/4$ for all m, $\pi_2^{\delta,\gamma_1,\gamma_2} = \rho(\mathscr{R}_2^{\delta,\gamma_1,\gamma_2})(\pi_1^{\delta,\gamma_1,\gamma_2}) = \mathscr{R}_2^{\delta,\gamma_1,\gamma_2}(\pi_1^{\delta,\gamma_1,\gamma_2})$ for fixed $\gamma_2 \in [0, 1)$ when γ_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 δ is sufficiently close to 1 and, given δ, γ_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 $\theta \in \Theta$, $m \in M$. 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 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 θ_1 sender playing (In, m) but never encounters a type θ_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 5 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.4 Proof of Lemma 8

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

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

Proof. Let $\pi = (\pi_1, \pi_2) \in \Pi_1 \times \Pi_2$ 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 π. 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 receiving their outcome under π. Both $\mathcal{A}_-$ and $\mathcal{A}_+$ are 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 $\mathcal{A}_- = \Delta(BR(\Theta, s))$ or $\Delta(BR(\Theta, s)) = \mathcal{A}_+$. $\Delta(BR(\Theta, s)) = \mathcal{A}_+$ is not possible when π is a JCE since then, for every $\alpha \in \Delta(BR(\overline{\Theta}(s, \pi), s))$, there is some θ such that $u_1(\theta, s, \alpha) > u_1(\theta, \pi)$. Therefore, $\Delta(BR(\Theta, s)) = \mathcal{A}_-$, which gives $u_1(\theta, s, a) < u_1(\theta, \pi)$ for all $a \in BR(\Theta, s)$. ■

OA.5 Analysis of Example 2

Proposition OA 1. *If π is a JCE in the game in Example 2:*

1. *$\pi_1(s = 0|\theta = 1) > 0$ and $\pi_2(a = 20|s = 0, m) = 1$ for all m such that $\pi_1(s = 0, m) > 0$,*

2. $\pi_1(s=20|\theta=2)>0$ *and* $\pi_2(a=40|s=20,m)=1$ *for all* $m\in M$ *such that* $\pi_1(s=20,m)>0$*, and*

3. $\pi_1(s=60|\theta=3)=1$ *and* $\pi_2(a=60|s=60,m)=1$ *for all* $m\in M$ *such that* $\pi_1(s=60,m)>0$*.*

Condition 1 says that the $\theta=1$ sender plays $s=0$ with positive probability, and that the receiver responds with $a=20$ to every on-path signal-message pair with $s=0$. Condition 2 says that the $\theta=2$ sender plays $s=20$ with positive probability, and that the receiver responds with $a=40$ to every on-path signal-message pair with $s=20$. Condition 3 says that the $\theta=3$ sender plays $s=60$ with positive probability, and that the receiver responds with $a=60$ to every on-path signal-message pair with $s=60$.

Proof. We first establish that in any JCE π, the receiver's response to each signal-message pair played by $\theta=3$ has expected value of at least $60-10/3$. Suppose otherwise that there is some signal-message pair (s,m) that $\theta=3$ plays which induces a receiver response with expected value $\widetilde{a}<60-10/3$. It must be that $s<100$ as 100 is a strictly dominated signal for types other than $\theta=3$ and the receiver's response to any signal-message pair played only by $\theta=3$ must be 60 since $BR(3,s)=\{60\}$ for all s. Thus, $s+10\in S$. Note that $u_1(3,\pi)=3\widetilde{a}-s$, while $u_1(\theta,\pi)\leq\theta\widetilde{a}-s$ for $\theta=1$ and $\theta=2$. Since $u_1(3,s+10,a)=3a-s-10$, we have that $u_1(3,s+10,a)\geq u_1(3,\pi)$ if and only if $a\geq\widetilde{a}+10/3$, with the inequality strict for all $a>\widetilde{a}+10/3$. Moreover, $u_1(\theta,s+10,a)\geq u_1(\theta,\pi)$ for $\theta=1$ or $\theta=2$ only if $u_1(\theta,s+10,a)=\theta a-s-10\geq\theta\widetilde{a}-s$, which requires $a\geq\widetilde{a}+5$. Thus, $\overline{\Theta}(s+10,\pi)=\{3\}$, so the only justified response to $s+10$ is 60. As this is strictly greater than $\widetilde{a}+10/3$ when $\widetilde{a}<60-10/3$, it follows that in every JCE, any signal-message pair played $\theta=3$ must induce a receiver response with expected value at least $60-10/3$.

An immediate implication of this is that there must be some signal-message pair that $\theta=2$ sends with positive probability that $\theta=3$ does not send. The reason is that the receiver's best responses to distributions where the relative weight on $\theta=2$ versus $\theta=3$ exceeds that under the prior are all no more than 50.

We now show that the receiver's response to each signal-message pair played by $\theta = 2$ but not by $\theta = 3$ must have an expected value between 35 and 40. Since every receiver action strictly higher than 40 is strictly dominated whenever the probability of $\theta = 3$ is 0, we need only show that the expected value of the receiver response to any signal-message pair played by $\theta = 2$ must exceed 35. Suppose otherwise that there is some signal-message pair (s, m) that $\theta = 2$ plays but $\theta = 3$ does not play which induces a receiver response with expected value $\widetilde{a} < 35$. Again, it must be that $s < 100$, so $s + 10 \in S$. Note that $u_1(2, \pi) = 2\widetilde{a} - s$, while $u_1(1, \pi) \leq \widetilde{a} - s$. Since $u_1(2, s + 10, a) = 2a - s - 10$, we have that $u_1(2, s + 10, a) \geq u_1(2, \pi)$ if and only if $a \geq \widetilde{a}+5$, with the inequality strict for all $a > \widetilde{a}+5$. Moreover, $u_1(1, s+10, a) \geq u_1(1, \pi)$ only if $u_1(1, s + 10, a) = a - s - 10 \geq \widetilde{a} - s$, which requires $a \geq \widetilde{a} + 10$. Thus, $1 \notin \overline{\Theta}(s + 10, \pi)$, so justified responses to $s + 10$ must weakly exceed 40. As this is strictly greater than $\widetilde{a}+5$ when $\widetilde{a} < 35$, it follows that in every JCE, any signal-message pair played by $\theta = 2$ must induce a receiver response with expected value at least 35.

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, m) played by both $\theta = 1$ and $\theta = 3$, the expected value of the receiver response $\widetilde{a}$ must be less than 50, because increasing differences in θ and a in the sender utility function implies that every signal-message pair played by $\theta = 2$ must induce a receiver response with this same expected value $\widetilde{a}$. This implies that $\widetilde{a} \leq 50$ since the receiver's best responses to distributions where the weight on $\theta = 2$ exceeds that under the prior are all no more than 50. However, this contradicts the fact that every signal-message played by $\theta = 3$ must induce a receiver response with expected value weakly greater than $60 - 10/3$. Additionally, $\theta = 1$ cannot only play signal-message pairs that are also played by $\theta = 2$. Otherwise, there would be some signal-message pair played by $\theta = 2$ that induces a receiver response with expected value weakly less than 30, which has already been shown to not be possible. This follows from the fact that the receiver's best responses to distributions where the weight on $\theta = 3$ is 0 and the relative weight of $\theta = 1$ to $\theta = 2$ exceeds that under the prior are all no more than 30.

Moreover, for every signal-message pair that only $\theta = 1$ plays, $s = 0$ and the receiver responds with $a = 20$. That the receiver responds with $a = 20$ to every signal-message pair that only $\theta = 1$ plays follows from the fact that $BR(1, s) = \{20\}$ for all s. So the payoff $\theta = 1$ obtains from a signal-message pair (s, m) that only $\theta = 1$ plays is $20 - s$, which is strictly less than 20 for all $s > 0$. However, $\theta = 1$ can secure a payoff of 20 by simply playing $s = 0$, since every $a < 20$ is strictly dominated for the receiver. This, $s = 0$ for every signal-message pair that only $\theta = 1$ plays.

We now argue that, for every signal-message pair played by $\theta = 2$ but not by $\theta = 3$, $s = 20$ and the receiver responds with $a = 40$. We have previously established that the expected value of the receiver's response, $\widetilde{a}$, to such a signal-message pair, (s, m), must be between 35 and 40. For $\widetilde{a} < 40$ to hold, it must be that $\theta = 1$ also plays this signal-message pair. This requires $u_1(1, s, \widetilde{a}) = \widetilde{a} - s = u_1(1, \pi)$. As previously established, $u_1(1, \pi) = 20$, so it must be that $s = \widetilde{a} - 20$. However, there is no $\widetilde{a} \in [35, 40)$ such that $\widetilde{a} - 20 \in S$. Therefore, $\widetilde{a} = 40$ so the receiver's response is necessarily 40 since the receiver never responds to any on-path signal-message pair with a mixture over non-adjacent actions. From $u_1(1, s, 40) = 40 - s \leq 20 = u_1(1, \pi)$, we obtain $s \geq 20$. All that remains is to rule out $s \geq 30$. If $s \geq 30$, $u_1(1, s - 10, a) = a - s + 10 \geq 20 = u_1(1, \pi)$ if and only if $a \geq 40$. On the other hand, $u_1(2, s - 10, a) = 2a - s + 10 \geq 80 - s = u_1(2, \pi)$ if and only if $a \geq 35$, with the inequality strict for all $a > 35$. Thus, $1 \notin \overline{\Theta}(s - 10, \pi)$, so justified responses to $s + 10$ must weakly exceed 40. It follows that $s = 20$.

Finally, we show that, for every signal-message pair played by $\theta = 3$, $s = 60$ and the receiver responds with $a = 60$. We have previously established that the expected value of the receiver's response, $\widetilde{a}$, to such a signal-message pair, (s, m), must be between $60 - 10/3$ and 60. For $\widetilde{a} < 60$ to hold, it must be that $\theta = 2$ also plays this signal-message pair. This requires $u_1(2, s, \widetilde{a}) = 2\widetilde{a} - s = u_1(2, \pi)$. As previously established, $u_1(2, \pi) = 60$, so it must be that $s = 2\widetilde{a} - 60$. However, there is no $\widetilde{a} \in [60 - 10/3, 60)$ such that $2\widetilde{a} - 60 \in S$. Therefore, $\widetilde{a} = 60$ so the receiver's response is necessarily 60. From $u_1(2, s, 60) = 120 - s \leq 60 = u_1(2, \pi)$, we obtain $s \geq 60$. All that remains is to rule out $s > 60$. If $s \geq 70$, $u_1(\theta, s-10, a) = \theta a - s + 10 \geq u_1(\theta, \pi)$ for either $\theta = 1$ or $\theta = 2$ requires that $a \geq 60$. On the other hand, $u_1(3, s-10, a) = 3a - s + 10 \geq 180 - s = u_1(3, \pi)$ if and only if $a \geq 170/3$, with the inequality strict for all $a > 170/3$. Thus, $\overline{\Theta}(s-10, \pi) = \{3\}$, so the only justified response to $s - 10$ is 60. It follows that $s = 60$. ■

OA.6 Omitted Examples

OA.6.1 Example Where Stability Does Not Imply the Intuitive Criterion without Initially Suggestible Receivers

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

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

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

Table 1: The payoffs for Example OA 1.

Out strictly dominates In for type θ_2, so θ_2 plays Out in every equilibrium of this game. However, there are equilibria in which θ_1 plays In and equilibria in which θ_1 plays Out. The equilibria where θ_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 θ_1, and θ_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 θ_1 plays Out.

Holding fixed an aggregate sender behavior strategy π_1 in the population, the receiver policy leads to a unique aggregate receiver behavior strategy. Likewise, holding fixed an aggregate receiver behavior strategy π_2 in the population, the profile of sender policies leads to a unique aggregate sender behavior strategy. There is a steady state where the aggregate strategy profile is $\pi = (\pi_1, \pi_2)$ precisely when the continuation probability parameters and policies are such that π_1 induces π_2 and π_2 induces π_1.

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 θ_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 θ_1 sender plays In approaches 0 as $\gamma_1 \to 1$.

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

$$\psi(\pi_2) = \left\{ \pi_2' \in \Pi_2 : \pi_2'(a_1|In, m) = \min\left\{ \pi_2(a_1|In, m), \frac{1}{3} \right\} \ \forall m \in M \right\}.$$

Note that ψ 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 $\widetilde{\mathscr{R}}^{\delta,\gamma_1,\gamma_2}(\pi_1, \pi_2) = (\mathscr{R}_1^{\delta,\gamma_1,\gamma_2}(\pi_2), \psi(\mathscr{R}_2^{\delta,\gamma_1,\gamma_2}(\pi_1)))$. 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 5 implies that $\lim_{\gamma_1 \to 1} \pi_1^{\delta,\gamma_1,\gamma_2}[In] = 0$ 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), $\lim_{\gamma_1 \to 1} \mathscr{R}_2^{\delta,\gamma_1,\gamma_2}(\pi_1^{\delta,\gamma_1,\gamma_2})(a_1|In, m) = 0$ for all m, $\gamma_2 \in [0, 1)$, so the $\pi_2(a_1|In, m) \leq 1/3$ constraint does not bind when δ is sufficiently close to 1 and, given δ, γ_1 is sufficiently close to 1. Formally, since $\pi_2^{\delta,\gamma_1,\gamma_2} \neq \mathscr{R}_2^{\delta,\gamma_1,\gamma_2}(\pi_1^{\delta,\gamma_1,\gamma_2})$ only if $\mathscr{R}_2^{\delta,\gamma_1,\gamma_2}(\pi_1^{\delta,\gamma_1,\gamma_2})(a_1|In, m) > 1/3$ for some m, we have that, for fixed $\gamma_2 \in [0, 1)$, $\pi_2^{\delta,\gamma_1,\gamma_2} = \mathscr{R}_2^{\delta,\gamma_1,\gamma_2}(\pi_1^{\delta,\gamma_1,\gamma_2})$ for δ sufficiently close to 1 and, given δ, γ_1 sufficiently close to 1. Combining this with the fact that $\pi_1^{\delta,\gamma_1,\gamma_2} = \mathscr{R}_1^{\delta,\gamma_1,\gamma_2}(\pi_2^{\delta,\gamma_1,\gamma_2})$ 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 δ sufficiently close to 1 and, given δ, γ_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. $\square$

OA.6.2 Example Where D1 Does Not Imply JCE

Example OA 2. The sender's type space is $\Theta = \{\theta_1, \theta_2, \theta_3\}$, and the receiver's prior is that all types are equally likely. The sender's signal space is $S = \{In, Out\}$, and the receiver's action space is $A = \{a_1, a_2, a_3\}$. The payoffs to the sender and receiver are given in Table 2 below.

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

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

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

Table 2: The payoffs for Example OA 2.

Every type playing Out is a D1 equilibrium outcome. To see this, let π denote a strategy profile in which every type plays Out and note that $\{\alpha \in MBR(\Theta, In) : \alpha[a_3] < 1/2\} \subset D_{\theta_3}(In, \pi)$, while $\{\nu a_2 + (1-\nu)a_3 : \nu \in [0,1]\} \cap D_{\theta_1}(In, \pi) = \emptyset$ and $\{\nu a_1 + (1-\nu)a_3 : \nu \in [0,1]\} \cap D_{\theta_2}(In, \pi) = \emptyset$. Thus, $D_{\theta_3}(In, \pi) \not\subseteq D_{\theta_1}(In, \pi)$ and $D_{\theta_3}(In, \pi) \not\subseteq D_{\theta_2}(In, \pi)$, so the receiver responding to In with a_3, which deters all types from playing In, is consistent with D1 since $BR(\theta_3, In) = \{a_3\}$.

However, JCE rules out profiles in which every sender type plays Out. Again, let π denote such a strategy profile. $\widetilde{D}^0_{\theta_3}(In, \pi) = \{\alpha \in \Delta(A) : \alpha[a_3] = 1/2\}$. Since

$$u_1(\theta_1, In, \alpha) + u_1(\theta_2, In, \alpha) = 3\alpha[a_1] + 3\alpha[a_2] - 2\alpha[a_3] = 3 - 5\alpha[a_3],$$

it follows that $\{\alpha \in \Delta(A) : \alpha[a_3] = 1/2\} \subset \widetilde{D}_{\theta_1}(In, \pi) \cup \widetilde{D}_{\theta_2}(In, \pi)$. Note that $BR(\{\theta_1, \theta_2\}, In) = \{a_1, a_2\}$, so no such profile can be a JCE since type θ_3 cannot be deterred from playing In by the receiver with either a_1 or a_2, as $u_1(\theta_3, In, a_1) = u_1(\theta_3, In, a_2) = 1 > 0$. □

OA.6.3 Example Where Stability Does Not Imply D1

Example OA 3. The sender's type space is $\Theta = \{\theta_1, \theta_2\}$, and the receiver's prior is that both types are equally likely. The sender's signal space is $S = \{In, Out\}$, and the receiver's action space is $A = \{a_1, a_2, a_3\}$. The payoffs to the sender and receiver are given in Table 3 below.

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

θ_2	a_1	a_2	a_3
In	$1, -4$	$-1, 5$	$0, .1$
Out	$0, 0$	$0, 0$	$0, 0$

Table 3: The payoffs for Example OA 3.

Every type playing Out is a JCE outcome. Since $u_1(\theta_2, In, (1/2)a_1 + (1/2)a_2) = 0$ and $u_1(\theta_1, In, (1/2)a_1 + (1/2)a_2) < 0$, θ_2 is a justified type, so a_2, which deters both types from playing In, is a justified response as is it the unique best response to θ_2 playing In.

However, every type playing Out is not a D1 equilibrium outcome. To see this, note that

$$MBR(\Theta, In) = \{\alpha \in \Delta(A) : \alpha[a_1] = 0 \text{ or } \alpha[a_2] = 0\}.$$

Let π denote a strategy profile in which every type plays Out. As $D_{\theta_2}(In, \pi) \cup D^0_{\theta_2}(In, \pi) = \{\alpha \in MBR(\Theta, In) : \alpha[a_2] = 0\}$ and $D_{\theta_1}(In, \pi) = \{\alpha \in MBR(\Theta, In) : \alpha[a_2] < 1/10\}$, we have that $D_{\theta_2}(In, \pi) \cup D^0_{\theta_2}(In, \pi) \subset D_{\theta_1}(In, \pi)$. Since $\{a_1\} = BR(\theta_1, In)$ and neither type is deterred from playing In by a_1, it follows that no such strategy profile is D1.

We now show that there are stable profiles in which all types play Out. We specify that the marginal of the receiver prior g_2 on In is a Dirichlet distribution with initial weight 1 on $(\theta_1, In, m_{In,\theta_1})$ and 1/4 on $(\theta_2, In, m_{In,\theta_1})$, and, for all other messages $m \neq m_{In,\theta_1}$, initial weight 1/4 on (θ_1, In, m) and 1 on (θ_2, In, m). Note that initial suggestibility is satisfied: When a receiver first encounters a sender who plays $(In, m_{In,\theta})$, the probability they place on the receiver having type θ is 4/5 so $BR(\theta, In)$ is optimal.

We observe that a_2 is the receiver's unique best response to In under any distribution that puts weakly more weight on θ_2 than the prior. Additionally, if a receiver has encountered past play of (In, m) and all such plays have been by senders with type θ_2, then the receiver will respond to the next instance of (In, m) with a_2. To see that this holds for the case $m = m_{In,\theta_1}$, note that the receiver's conditional distribution over the sender's type after (In, m_{In,θ_1}) must put probability at least 5/9 on θ_2. Analogous arguments handle the other cases.

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 1/10. Under such responses, whenever it is weakly optimal for θ_1 to play In, it must be strictly optimal for θ_2 to do so. To see this, note that

$$u_1(\theta_1, In, \alpha) = 3\alpha[a_1] + 2.1\alpha[a_3] - 2,$$

so $\alpha[a_1] \geq 2/3 - 7/10\alpha[a_3]$ whenever $u_1(\theta_1, In, \alpha) \geq 0$. Additionally,

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

which is strictly positive whenever $\alpha[a_1] \geq 2/3 - 7/10\alpha[a_3]$ and $\alpha[a_3] \leq 1/10$. We argue that such steady state profiles exist in the iterated limit where $\gamma_1 \to 1$ 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

$$\chi(\alpha) = \begin{cases} \{\alpha\} & \text{if } \alpha[a_3] \leq \frac{1}{10} \\ \{\alpha' \in \Delta(A) : \alpha'[a_3] = \frac{1}{10}\} & \text{if } \alpha[a_3] > \frac{1}{10} \end{cases},$$

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 ρ is upper hemicontinuous and coincides with the identity correspondence whenever $\pi_2(a_3|In, m) \leq 1/10$ 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 $\mathscr{R}^{\delta,\gamma_1,\gamma_2}(\pi_1, \pi_2) = \{(\pi_1', \pi_2') \in \Pi_1 \times \Pi_2 : \pi_1' = \mathscr{R}_1^{\delta,\gamma_1,\gamma_2}(\pi_2) \text{ and } \pi_2' \in \rho(\mathscr{R}_2^{\delta,\gamma_1,\gamma_2}(\pi_1))\}$. Since $\mathscr{R}$ is upper hemicontinuous, 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 $\lim_{\gamma_2\to 1} \lim_{\delta\to 1} \lim_{\gamma_1\to 1} \pi_1^{\delta,\gamma_1,\gamma_2}[In|\theta_1] = 0$. 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, (d) $\lim_{j\to\infty} \lim_{k\to\infty} \lim_{l\to\infty} \pi_1^{\delta_{j,k},\gamma_{1,j,k,l},\gamma_{2,j}}[In, m|\theta]$ exists for all $\theta \in \Theta$, $m \in M$, and (e) $\lim_{j\to\infty} \lim_{k\to\infty} \lim_{l\to\infty} \pi_1^{\delta_{j,k},\gamma_{1,j,k,l},\gamma_{2,j}}[In|\theta_1] > 0$. Then since $\pi_2^{\delta,\gamma_1,\gamma_2}(a_3|In, m) \leq 1/10$ for all $m \in M$, Lemma 5 implies that $\lim_{j\to\infty} \lim_{k\to\infty} \lim_{l\to\infty} \pi_1^{\delta_{j,k},\gamma_{1,j,k,l},\gamma_{2,j}}[In|\theta_2] = 1$. Therefore, there exists some $m \in M$ such that $\lim_{j\to\infty} \lim_{k\to\infty} \lim_{l\to\infty} \pi_1^{\gamma_{1,j,k,l},\gamma_{2,j}}[In|\theta_2] > 0$ and $\lim_{j\to\infty} \lim_{k\to\infty} \lim_{l\to\infty} \pi_1^{\delta_{j,k},\gamma_{1,j,k,l},\gamma_{2,j}}[In|\theta_2] \geq \lim_{j\to\infty} \lim_{k\to\infty} \lim_{l\to\infty} \pi_1^{\delta_{j,k},\gamma_{1,j,k,l},\gamma_{2,j}}[In|\theta_1]$. By Lemma 2, this implies that $\lim_{j\to\infty} \lim_{k\to\infty} \lim_{l\to\infty} \mathscr{R}_2^{\delta_{j,k},\gamma_{1,j,k,l},\gamma_{2,j}}(\pi_1^{\delta_{j,k},\gamma_{1,j,k,l},\gamma_{2,j}})(a_2|In, m) = 1$. Since $\chi(\pi_2(\cdot|In, m)) = \{\pi_2(\cdot|In, m)\}$ if $\pi_2(\cdot|In, m) \leq 1/10$, it follows that $\lim_{j\to\infty} \lim_{k\to\infty} \lim_{l\to\infty} \pi_2^{\delta_{j,k},\gamma_{1,j,k,l},\gamma_{2,j}}(\pi_1)(a_2|In, m) = 1$. However, by Lemma 5, this requires that $\lim_{j\to\infty} \lim_{k\to\infty} \lim_{l\to\infty} \pi_1^{\delta_{j,k},\gamma_{1,j,k,l},\gamma_{2,j}}[In, m] = 0$ must hold, a contradiction.

A very similar argument establishes that $\lim_{\gamma_2\to 1}\lim_{\delta\to 1}\lim_{\gamma_1\to 1}\pi_1^{\delta,\gamma_1,\gamma_2}[In|\theta_1]=0$. Thus, $\lim_{\gamma_2\to 1}\lim_{\delta\to 1}\lim_{\gamma_1\to 1}\pi_1^{\delta,\gamma_1,\gamma_2}[In]=0$. Since a receiver will only play a_3 in response to some (In,m) if they have previously encountered a sender playing (In,m), we have that, for all sufficiently high γ_2, $\mathscr{R}_2^{\delta_{j,k},\gamma_{1,j,k,l},\gamma_{2,j}}(\pi_1^{\gamma_{1,k,l},\gamma_{2,k}})(a_3|In,m)\le 1/10$ for all $m\in M$ when γ_1 is sufficiently high. Since $\rho(\pi_2)=\{\pi_2\}$ if $\pi_2(a_3|In,m)\le 1/10$ for all m, $\pi_2^{\delta,\gamma_1,\gamma_2}=\rho(\mathscr{R}_2^{\delta,\gamma_1,\gamma_2})(\pi_1^{\delta,\gamma_1,\gamma_2})=\mathscr{R}_2^{\delta,\gamma_1,\gamma_2}(\pi_1^{\delta,\gamma_1,\gamma_2})$ for fixed, sufficiently high $\gamma_2\in[0,1)$ when δ is sufficiently close to 1 and, given δ, γ_1 is sufficiently close to 1. Thus, for fixed, sufficiently high $\gamma_2\in[0,1)$, $(\pi_1^{\delta,\gamma_1,\gamma_2},\pi_2^{\delta,\gamma_1,\gamma_2})$ is a fixed point of $\mathscr{R}^{\delta,\gamma_1,\gamma_2}$ when when δ is sufficiently close to 1 and, given δ, γ_1 is sufficiently close to 1, and we conclude that there are stable profiles in which every type plays *Out*. □

OA.7 Stability Under Alternative Assumptions

OA.7.1 Weakening Initial Suggestibility

Here we discuss the refinement satisfied by all stable profiles under the alternative assumption to initial suggestibility introduced in Section 6. As before, any stable profile must be a PBE-H. Moreover, stability also imposes additional conditions for profiles π that are on-path strict for the receiver or are such that the sender types' payoffs would not be changed if the receiver deviated.[1] For such a profile to be stable, it must be that, for every signal s where $u_1(\theta,s,\alpha)<u_1(\theta,\pi)$ for all $\theta\notin\overline{\Theta}(s,\pi)$, $\alpha\in\Delta(BR(\overline{\Theta}(s,\pi),s))$, 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 $u_1(\theta,s,\alpha)<u_1(\theta,\pi)$ for all $\theta\notin\overline{\Theta}(s,\pi)$, $\alpha\in\Delta(BR(\overline{\Theta}(s,\pi),s))$, this requirement is the same as Condition 2 of Definition 4.

This refinement is weaker than JCE, strictly so in some games, including some co-monotonic games. Thus, it preserves the equilibria we focus on in Examples 1 and 3. This refinement does not make the same predictions as JCE does in Example 2;

[1]These restrictions on π guarantee that a typical receiver agent will learn the equilibrium payoffs of the sender types with high probability.

however, if the game were altered to have finer action spaces that sufficiently approximate a continuum, then this refinement, like JCE, would select only equilibria that are close to the least-cost separating equilibrium. The D1 equilibrium in Example OA.6.2 satisfies this refinement, but there are other games in which this refinement rules out D1 equilibria.

OA.7.2 Costly Messages

Suppose that we allow the sender's utility function $u_1 : \Theta \times S \times M \times A \to \mathbb{R}$ and the receiver's utility function $u_2 : \Theta \times S \times M \times A \to \mathbb{R}$ to also depend on the sender's message m. We can compute $\overline{\Theta}(s, m, \pi)$, the set of justified types for signal s and message m given profile π, in a way analogous to before. Then, under initial suggestibility, any stable profile π must satify the following requirement: $\pi_2(\cdot|s, m_{s,\widetilde{\Theta}}) \in \Delta(BR(\widetilde{\Theta}, s))$ for all $s \in S$ and $\widetilde{\Theta} \subseteq \Theta$ such that $\overline{\Theta}(s, m_{s,\widetilde{\Theta}}, \pi) \subseteq \widetilde{\Theta}$. Observe that, when $\overline{\Theta}(s, m, \pi) = \overline{\Theta}(s, \pi)$ is independent of $m \in M$, as in the case where the sender's message is payoff irrelevant, applying this requirement with $\widetilde{\Theta} = \overline{\Theta}(s, \pi)$ implies Condition 2 of Definition 4.

OA.7.3 Strengthening Initial Suggestibility

Suppose that we strengthen initial suggestibility 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 $(s, m_{s,\widetilde{\Theta}})$ belongs to $BR(\widetilde{\Theta} \cup \widetilde{\Theta}', s)$. 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]).

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

$$\widetilde{D}_{\theta}^{0,n}(s, \pi) = \{\alpha \in \Delta(BR(\overline{\Theta}^{n-1}(s, \pi), s)) : u_1(\theta, s, \alpha) = u_1(\theta, \pi)\},$$

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

$$\overline{\Theta}^{n}(s, \pi) = \begin{cases} \Theta^{\dagger,n}(s, \pi) & \text{if } \Theta^{\dagger,n}(s, \pi) \neq \emptyset \\ \overline{\Theta}^{n-1}(s, \pi) & \text{if } \Theta^{\dagger,n}(s, \pi) = \emptyset \end{cases}.$$

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

Every stable profile π must satisfy the following requirement: For every signal s, there is some $m \in M$ such that $\pi_2(\cdot|s, m) \in \Delta(BR(\overline{\Theta}^{\infty}(s, \pi), s))$.

References

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

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