The Canonical Reputation Model and Reputation Effects

George J. Mailath

University of Pennsylvania

Lecture at the Summer School in Economic Theory
Jerusalem, Israel
June 25, 2015

The Canonical Reputation Model

A long-lived player 1 faces a sequence of short-lived players, in the role of player 2 of the stage game.

- A_i, finite action set for each player.
- Y, finite set of public signals of player 1's actions, a_1.
- $\rho(y \mid a_1)$, prob of signal $y \in Y$, given $a_1 \in A_1$.
- Player 2's ex post stage game payoff is $u_2^*(a_1, a_2, y)$, and 2's ex ante payoff is

$$u_2(a_1, a_2) := \sum_{y \in Y} u_2^*(a_1, a_2, y)\rho(y|a_1).$$

- Each player 2 maximizes her stage game payoff u_2.

- The player 2's are uncertain about the characteristics of player 1: Player 1's characteristics are described by his type, $\xi \in \Xi$.
- All the player 2's have a common prior μ on Ξ.
- Type space is partitioned into two sets, $\Xi = \Xi_1 \cup \Xi_2$, where
  - Ξ_1 is the set of payoff types and
  - Ξ_2 is the set of behavioral (or commitment or action) types.
- For $\xi \in \Xi_1$, player 1's ex post stage game payoff is $u_1^*(a_1, a_2, y, \xi)$, and 1's ex ante payoff is

$$u_1(a_1, a_2, \xi) := \sum\nolimits_{y \in Y} u_1^*(a_1, a_2, y, \xi)\rho(y|a_1).$$

- Each type $\xi \in \Xi_1$ of player 1 maximizes

$$(1 - \delta) \sum\nolimits_{t \geq 0} \delta^t u_1(a_1, a_2, \xi).$$

- Player 1 knows his type and observes all past actions and signals, while each player 2 only the history of past signals.
- A strategy for player 1:

$$\sigma_1 : \cup_{t=0}^{\infty}(A_1 \times A_2 \times Y)^t \times \Xi \to \Delta(A_1).$$

  If $\hat{\xi} \in \Xi_2$ is a simple action type, then $\exists! \hat{\alpha}_1 \in \Delta(A_1)$ such that $\sigma_1(h_1^t, \hat{\xi}) = \hat{\alpha}_1$ for all h_1^t.
- A strategy for player 2:

$$\sigma_2 : \cup_{t=0}^{\infty} Y^t \to \Delta(A_2).$$

- Space of outcomes: $\Omega := \Xi \times (A_1 \times A_2 \times Y)^\infty$.
- A profile (σ_1, σ_2) with prior μ induces the unconditional distribution $\mathbf{P} \in \Delta(\Omega)$.
- For a fixed simple type $\hat{\xi} = \xi(\hat{\alpha}_1)$, the probability measure on Ω conditioning on $\hat{\xi}$ (and so induced by $\hat{\alpha}_1$ in every period and σ_2), is denoted $\widehat{\mathbf{P}} \in \Delta(\Omega)$.
- Denoting by $\widetilde{\mathbf{P}}$ the measure induced by (σ_1, σ_2) and conditioning on $\xi \neq \hat{\xi}$, we have

$$\mathbf{P} = \mu(\hat{\xi})\widehat{\mathbf{P}} + (1 - \mu(\hat{\xi}))\widetilde{\mathbf{P}}. \tag{1}$$

- Given a strategy profile σ, $U_1(\sigma, \xi)$ denotes the type-ξ long-lived player's payoff in the repeated game,

$$U_1(\sigma, \xi) := E_{\mathbf{P}}\left[(1-\delta) \sum_{t=0}^{\infty} \delta^t u_1(a^t, y^t, \xi) \middle| \xi \right].$$

Denote by $\Gamma(\mu, \delta)$ the game of incomplete information.

Definition

A strategy profile (σ_1', σ_2') is a Nash equilibrium of the game $\Gamma(\mu, \delta)$ if, for all $\xi \in \Xi_1$, σ_1' maximizes $U_1(\sigma_1, \sigma_2', \xi)$ over player 1's repeated game strategies, and if for all t and all $h_2^t \in \mathcal{H}_2$ that have positive probability under (σ_1', σ_2') and μ (i.e., $\mathbf{P}(h_2^t) > 0$),

$$E_{\mathbf{P}}\left[u_2(\sigma_1'(h_1^t, \xi), \sigma_2'(h_2^t)) \mid h_2^t\right] = \max_{a_2 \in A_2} E_{\mathbf{P}}\left[u_2(\sigma_1'(h_1^t, \xi), a_2) \mid h_2^t\right].$$

Denote by $\Gamma(\mu, \delta)$ the game of incomplete information.

Definition

A strategy profile (σ_1', σ_2') is a Nash equilibrium of the game $\Gamma(\mu, \delta)$ if, for all $\xi \in \Xi_1$, σ_1' maximizes $U_1(\sigma_1, \sigma_2', \xi)$ over player 1's repeated game strategies, and if for all t and all $h_2^t \in \mathcal{H}_2$ that have positive probability under (σ_1', σ_2') and μ (i.e., $\mathbf{P}(h_2^t) > 0$),

$$E_{\mathbf{P}}\left[u_2(\sigma_1'(h_1^t, \xi), \sigma_2'(h_2^t)) \mid h_2^t\right] = \max_{a_2 \in A_2} E_{\mathbf{P}}\left[u_2(\sigma_1'(h_1^t, \xi), a_2) \mid h_2^t\right].$$

Our goal: Reputation Bound (Fudenberg & Levine '89 '92)

Fix a payoff type, $\xi \in \Xi_1$. What is a "good" lower bound, uniform across Nash equilibria σ' and Ξ, for $U_1(\sigma', \xi)$?

Our tool (Gossner 2011): relative entropy.

Relative Entropy

- X a finite set of outcomes.
- The relative entropy or Kullback-Leibler distance between probability distributions p and q over X is

$$d(p\|q) := \sum_{x \in X} p(x) \log \frac{p(x)}{q(x)}.$$

  By convention, $0 \log \frac{0}{q} = 0$ for all $q \in [0,1]$ and $p \log \frac{p}{0} = \infty$ for all $p \in (0,1]$. In our applications of relative entropy, the support of q will always contain the support of p.
- Since relative entropy is not symmetric, often say $d(p\|q)$ is the relative entropy of q with respect to p.
- $d(p\|q) \geq 0$, and $d(p\|q) = 0 \iff p = q$.

Relative entropy is expected prediction error

$d(p\|q)$ measures observer's expected prediction error on $x \in X$ using q when true dsn is p:

- n i.i.d. draws from X under p has probability $\prod_x p(x)^{n_x}$, where n_x is the number of realization of x in sample.
- Observer assigns same sample probability $\prod_x q(x)^{n_x}$.
- Log likelihood ratio is

$$\mathcal{L}(x_1, \ldots, x_n) = \sum_x n_x \log \frac{p(x)}{q(x)},$$

and so

$$\frac{1}{n}\mathcal{L}(x_1, \ldots, x_n) \to d(p\|q).$$

The chain rule

Lemma

Suppose $P, Q \in \Delta(X \times Y)$, X and Y finite sets. Then

$$\begin{aligned} d(P\|Q) &= d(P_X\|Q_X) + \sum_x P_X(x) d(P_Y(\cdot|x)\|Q_Y(\cdot|x)) \\ &= d(P_X\|Q_X) + E_{P_X} d(P_Y(\cdot|x)\|Q_Y(\cdot|x)). \end{aligned}$$

The chain rule

Lemma

Suppose $P, Q \in \Delta(X \times Y)$, X and Y finite sets. Then

$$\begin{aligned} d(P\|Q) &= d(P_X\|Q_X) + \sum_x P_X(x) d(P_Y(\cdot|x)\|Q_Y(\cdot|x)) \\ &= d(P_X\|Q_X) + E_{P_X} d(P_Y(\cdot|x)\|Q_Y(\cdot|x)). \end{aligned}$$

Proof.

$$d(P\|Q) = \sum_{x,y} P(x,y) \log \frac{P_X(x)}{Q_X(x)} \frac{P(x,y)}{P_X(x)} \frac{Q_X(x)}{Q(x,y)}$$

□

The chain rule

Lemma

Suppose $P, Q \in \Delta(X \times Y)$, X and Y finite sets. Then

$$d(P\|Q) = d(P_X\|Q_X) + \sum_x P_X(x) d(P_Y(\cdot|x)\|Q_Y(\cdot|x))$$
$$= d(P_X\|Q_X) + E_{P_X} d(P_Y(\cdot|x)\|Q_Y(\cdot|x)).$$

Proof.

$$d(P\|Q) = \sum_{x,y} P(x,y) \log \frac{P_X(x)}{Q_X(x)} \frac{P(x,y)}{P_X(x)} \frac{Q_X(x)}{Q(x,y)}$$
$$= d(P_X\|Q_X) + \sum_{x,y} P(x,y) \log \frac{P_Y(y|x)}{Q_Y(y|x)}$$

□

The chain rule

Lemma

Suppose $P, Q \in \Delta(X \times Y)$, X *and* Y *finite sets. Then*

$$d(P\|Q) = d(P_X\|Q_X) + \sum_x P_X(x) d(P_Y(\cdot|x)\|Q_Y(\cdot|x))$$
$$= d(P_X\|Q_X) + E_{P_X} d(P_Y(\cdot|x)\|Q_Y(\cdot|x)).$$

Proof.

$$d(P\|Q) = \sum_{x,y} P(x,y) \log \frac{P_X(x)}{Q_X(x)} \frac{P(x,y)}{P_X(x)} \frac{Q_X(x)}{Q(x,y)}$$
$$= d(P_X\|Q_X) + \sum_{x,y} P(x,y) \log \frac{P_Y(y|x)}{Q_Y(y|x)}$$

□

The chain rule

Lemma

Suppose $P, Q \in \Delta(X \times Y)$, X and Y finite sets. Then

$$\begin{aligned} d(P\|Q) &= d(P_X\|Q_X) + \textstyle\sum_x P_X(x) d(P_Y(\cdot|x)\|Q_Y(\cdot|x)) \\ &= d(P_X\|Q_X) + E_{P_X} d(P_Y(\cdot|x)\|Q_Y(\cdot|x)). \end{aligned}$$

Proof.

$$\begin{aligned} d(P\|Q) &= \textstyle\sum_{x,y} P(x,y) \log \frac{P_X(x)}{Q_X(x)} \frac{P(x,y)}{P_X(x)} \frac{Q_X(x)}{Q(x,y)} \\ &= d(P_X\|Q_X) + \textstyle\sum_{x,y} P(x,y) \log \frac{P_Y(y|x)}{Q_Y(y|x)} \end{aligned}$$

□

The chain rule

Lemma

Suppose $P, Q \in \Delta(X \times Y)$, X and Y finite sets. Then

$$\begin{aligned} d(P \| Q) &= d(P_X \| Q_X) + \textstyle\sum_x P_X(x) d(P_Y(\cdot|x) \| Q_Y(\cdot|x)) \\ &= d(P_X \| Q_X) + E_{P_X} d(P_Y(\cdot|x) \| Q_Y(\cdot|x)). \end{aligned}$$

Proof.

$$\begin{aligned} d(P \| Q) &= \textstyle\sum_{x,y} P(x,y) \log \frac{P_X(x)}{Q_X(x)} \frac{P(x,y)}{P_X(x)} \frac{Q_X(x)}{Q(x,y)} \\ &= d(P_X \| Q_X) + \textstyle\sum_{x,y} P(x,y) \log \frac{P_Y(y|x)}{Q_Y(y|x)} \\ &= d(P_X \| Q_X) + \textstyle\sum_x P_X(x) \sum_y P_Y(y|x) \log \frac{P_Y(y|x)}{Q_Y(y|x)}. \end{aligned}$$

□

A grain of truth

Lemma

Let X be a finite set of outcomes. Suppose $p, p' \in \Delta(X)$ and $q = \varepsilon p + (1 - \varepsilon)p'$ for some $\varepsilon > 0$. Then,

$$d(p \| q) \leq -\log \varepsilon.$$

A grain of truth

Lemma

Let X be a finite set of outcomes. Suppose $p, p' \in \Delta(X)$ and $q = \varepsilon p + (1-\varepsilon)p'$ for some $\varepsilon > 0$. Then,

$$d(p\|q) \leq -\log \varepsilon.$$

Proof.

Since $q(x)/p(x) \geq \varepsilon$, we have

$$-d(p\|q) = \sum_x p(x) \log \frac{q(x)}{p(x)} \geq \sum_x p(x) \log \varepsilon = \log \varepsilon.$$

□

Back to reputations!

- Fix $\hat{\alpha}_1 \in \Delta(A_1)$ and suppose $\mu(\xi(\hat{\alpha}_1)) > 0$.
- In a Nash eq, at history h_2^t, $\sigma_2(h_2^t)$ is a best response to

$$\alpha_1(h_2^t) := E_{\mathbf{P}}[\sigma_1(h_1^t, \xi) \mid h_2^t] \in \Delta(A_1),$$

that is, $\sigma_2(h_2^t)$ maximizes

$$\sum_{a_1} \sum_{y} u_2^*(a_1, a_2, y)\rho(y|a_1)\alpha_1(a_1|h_2^t).$$

Back to reputations!

- Fix $\hat{\alpha}_1 \in \Delta(A_1)$ and suppose $\mu(\xi(\hat{\alpha}_1)) > 0$.
- In a Nash eq, at history h_2^t, $\sigma_2(h_2^t)$ is a best response to

$$\alpha_1(h_2^t) := E_{\mathbf{P}}[\sigma_1(h_1^t, \xi) \mid h_2^t] \in \Delta(A_1),$$

  that is, $\sigma_2(h_2^t)$ maximizes

$$\sum_{a_1} \sum_y u_2^*(a_1, a_2, y)\rho(y|a_1)\alpha_1(a_1|h_2^t).$$

- At h_2^t, 2's predicted dsn on the signal y^t is

$$p(h_2^t) := \rho(\cdot|\alpha_1(h_2^t)) = \sum_{a_1} \rho(\cdot|a_1)\alpha_1(a_1|h_2^t).$$

Back to reputations!

- Fix $\hat{\alpha}_1 \in \Delta(A_1)$ and suppose $\mu(\xi(\hat{\alpha}_1)) > 0$.
- In a Nash eq, at history h_2^t, $\sigma_2(h_2^t)$ is a best response to

$$\alpha_1(h_2^t) := E_{\mathbf{P}}[\sigma_1(h_1^t, \xi) \mid h_2^t] \in \Delta(A_1),$$

  that is, $\sigma_2(h_2^t)$ maximizes

$$\sum_{a_1} \sum_y u_2^*(a_1, a_2, y)\rho(y|a_1)\alpha_1(a_1|h_2^t).$$

- At h_2^t, 2's predicted dsn on the signal y^t is

$$p(h_2^t) := \rho(\cdot|\alpha_1(h_2^t)) = \sum_{a_1} \rho(\cdot|a_1)\alpha_1(a_1|h_2^t).$$

- If player 1 plays $\hat{\alpha}_1$, true dsn on y^t is

$$\hat{p} := \rho(\cdot|\hat{\alpha}_1) = \sum_{a_1} \rho(\cdot|a_1)\hat{\alpha}_1(a_1).$$

Back to reputations!

- Fix $\hat{\alpha}_1 \in \Delta(A_1)$ and suppose $\mu(\xi(\hat{\alpha}_1)) > 0$.
- In a Nash eq, at history h_2^t, $\sigma_2(h_2^t)$ is a best response to

$$\alpha_1(h_2^t) := E_{\mathbf{P}}[\sigma_1(h_1^t, \xi) \mid h_2^t] \in \Delta(A_1),$$

that is, $\sigma_2(h_2^t)$ maximizes

$$\sum_{a_1} \sum_y u_2^*(a_1, a_2, y)\rho(y|a_1)\alpha_1(a_1|h_2^t).$$

- At h_2^t, 2's predicted dsn on the signal y^t is

$$p(h_2^t) := \rho(\cdot|\alpha_1(h_2^t)) = \sum_{a_1} \rho(\cdot|a_1)\alpha_1(a_1|h_2^t).$$

- If player 1 plays $\hat{\alpha}_1$, true dsn on y^t is

$$\hat{p} := \rho(\cdot|\hat{\alpha}_1) = \sum_{a_1} \rho(\cdot|a_1)\hat{\alpha}_1(a_1).$$

- Player 2's one-step ahead prediction error is

$$d\left(\hat{p} \| p(h_2^t)\right).$$

Bounding prediction errors

- Player 2 is best responding to an action profile $\alpha_1(h_2^t)$ that is $d\big(\hat{p}\|p(h_2^t)\big)$-close to $\hat{\alpha}_1$ (as measured by the relative entropy of the induced signals).
- To bound player 1's payoff, it suffices to bound the number of periods in which $d\big(\hat{p}\|p(h_2^t)\big)$ is large.

Bounding prediction errors

- Player 2 is best responding to an action profile $\alpha_1(h_2^t)$ that is $d\big(\hat{p} \| p(h_2^t)\big)$-close to $\hat{\alpha}_1$ (as measured by the relative entropy of the induced signals).
- To bound player 1's payoff, it suffices to bound the number of periods in which $d\big(\hat{p} \| p(h_2^t)\big)$ is large.
- For any t, $\mathbf{P}_2^t$ is the marginal of $\mathbf{P}$ on Y^t. Then,

$$\mathbf{P}_2^t = \mu(\hat{\xi})\widehat{\mathbf{P}}_2^t + (1 - \mu(\hat{\xi}))\widetilde{\mathbf{P}}_2^t,$$

and so

$$d(\widehat{\mathbf{P}}_2^t \| \mathbf{P}_2^t) \leq -\log \mu(\hat{\xi}).$$

Applying the chain rule:

$$
\begin{aligned}
-\log \mu(\hat{\xi}) &\geq d(\widehat{\mathbf{P}}_2^t \| \mathbf{P}_2^t) \\
&= d(\widehat{\mathbf{P}}_2^{t-1} \| \mathbf{P}_2^{t-1}) + E_{\widehat{\mathbf{P}}} d(\hat{p} \| p(h_2^{t-1})) \\
&= \sum_{\tau=0}^{t-1} E_{\widehat{\mathbf{P}}} d(\hat{p} \| p(h_2^{\tau})).
\end{aligned}
$$

Applying the chain rule:

$$\begin{aligned} -\log \mu(\hat{\xi}) &\geq d(\widehat{\mathbf{P}}_2^t \| \mathbf{P}_2^t) \\ &= d(\widehat{\mathbf{P}}_2^{t-1} \| \mathbf{P}_2^{t-1}) + E_{\widehat{\mathbf{P}}} d(\hat{p} \| p(h_2^{t-1})) \\ &= \sum_{\tau=0}^{t-1} E_{\widehat{\mathbf{P}}} d(\hat{p} \| p(h_2^{\tau})). \end{aligned}$$

Since this holds for all t,

$$\sum_{\tau=0}^{\infty} E_{\widehat{\mathbf{P}}} d(\hat{p} \| p(h_2^{\tau})) \leq -\log \mu(\hat{\xi}).$$

From prediction bounds to payoff bounds

Definition

An action $\alpha_2 \in \Delta(A_2)$ is an ε-entropy confirming best response to $\alpha_1 \in \Delta(A_1)$ if there exists $\alpha_1' \in \Delta(A_1)$ such that

1. α_2 is a best response to α_1'; and
2. $d(\rho(\cdot|\alpha_1)\|\rho(\cdot|\alpha_1')) \leq \varepsilon$.

The set of ε-entropy confirming BR's to α_1 is denoted $B_\varepsilon^d(\alpha_1)$.

From prediction bounds to payoff bounds

Definition

An action $\alpha_2 \in \Delta(A_2)$ is an ε-entropy confirming best response to $\alpha_1 \in \Delta(A_1)$ if there exists $\alpha_1' \in \Delta(A_1)$ such that

1. α_2 is a best response to α_1'; and
2. $d(\rho(\cdot|\alpha_1)\|\rho(\cdot|\alpha_1')) \leq \varepsilon$.

The set of ε-entropy confirming BR's to α_1 is denoted $B^d_\varepsilon(\alpha_1)$.

In a Nash eq, at any on-the-eq-path history h_2^t, player 2's action is a $d(\hat{p}\|p(h_2^t))$-entropy confirming BR to $\hat{\alpha}_1$.

From prediction bounds to payoff bounds

Definition

An action $\alpha_2 \in \Delta(A_2)$ is an ε-entropy confirming best response to $\alpha_1 \in \Delta(A_1)$ if there exists $\alpha_1' \in \Delta(A_1)$ such that

1. α_2 is a best response to α_1'; and
2. $d(\rho(\cdot|\alpha_1) \| \rho(\cdot|\alpha_1')) \leq \varepsilon$.

The set of ε-entropy confirming BR's to α_1 is denoted $B_\varepsilon^d(\alpha_1)$.

Define, for all payoff types $\xi \in \Xi_1$,

$$\underline{v}_{\alpha_1}^{\xi}(\varepsilon) := \min_{\alpha_2 \in B_\varepsilon^d(\alpha_1)} u_1(\alpha_1, \alpha_2, \xi),$$

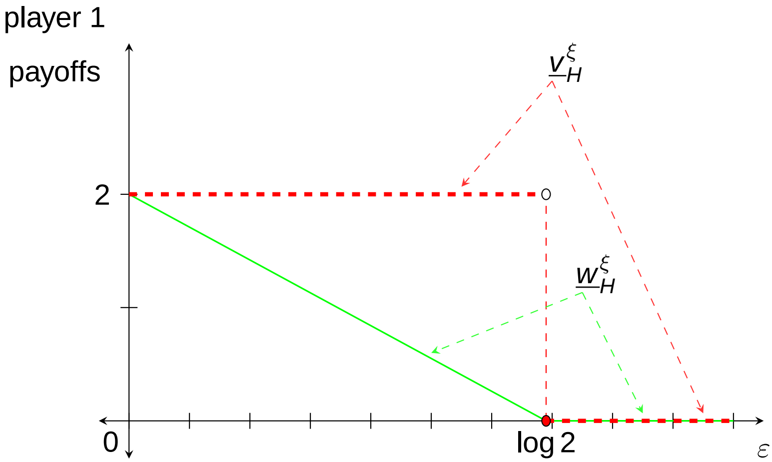
and denote by $\underline{w}_{\alpha_1}^{\xi}$ the largest convex function below $\underline{v}_{\alpha_1}^{\xi}$.

The product-choice game I

	c	s
H	$2, 3$	$0, 2$
L	$3, 0$	$1, 1$

- Suppose $\hat{\alpha}_1 = 1 \circ H$.
- c is unique BR to α_1 if $\alpha_1(H) > \frac{1}{2}$.
- s is also a BR to α_1 if $\alpha_1(H) = \frac{1}{2}$.
- $d(1 \circ H \| \frac{1}{2} \circ H + \frac{1}{2} \circ L) = \log \frac{1}{1/2} = \log 2 \approx 0.69$.
- $$\underline{v}_H^{\xi}(\varepsilon) = \begin{cases} 2, & \text{if } \varepsilon < \log 2, \\ 0, & \text{if } \varepsilon \geq \log 2. \end{cases}$$

A picture is worth a thousand words

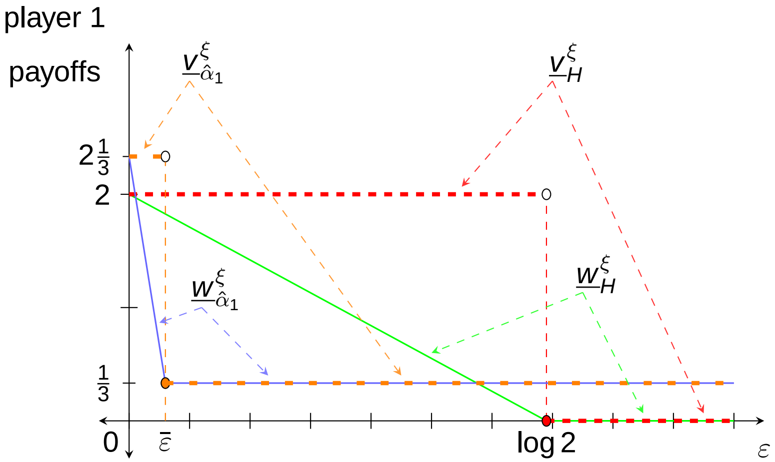


The product-choice game II

	c	s
H	$2,3$	$0,2$
L	$3,0$	$1,1$

- Suppose $\hat{\alpha}_1 = \frac{2}{3} \circ H + \frac{1}{3} \circ L$.
- c is unique BR to α_1 if $\alpha_1(H) > \frac{1}{2}$.
- s is also a BR to α_1 if $\alpha_1(H) = \frac{1}{2}$.
- $$\begin{aligned} d(\hat{\alpha}_1 \| \tfrac{1}{2} \circ H + \tfrac{1}{2} \circ L) &= \tfrac{2}{3} \log \tfrac{2/3}{1/2} + \tfrac{1}{3} \log \tfrac{1/3}{1/2} \\ &= \tfrac{5}{3} \log 2 - \log 3 \\ &=: \bar{\varepsilon} \approx 0.06. \end{aligned}$$

Two thousand?



The reputation bound

Proposition

Suppose the action type $\hat{\xi} = \xi(\hat{\alpha}_1)$ *has positive prior probability,* $\mu(\hat{\xi}) > 0$*, for some potentially mixed action* $\hat{\alpha}_1 \in \Delta(A_1)$*. Then, player* 1 *type* ξ*'s payoff in any Nash equilibrium of the game* $\Gamma(\mu, \delta)$ *is greater than or equal to* $\underline{w}^{\xi}_{\hat{\alpha}_1}(\hat{\varepsilon})$*, where*

$$\hat{\varepsilon} := -(1 - \delta) \log \mu(\hat{\xi}).$$

The only aspect of the set of types and the prior that plays a role in the proposition is the probability assigned to $\hat{\xi}$.
The set of types may be very large, and other quite crazy types may receive significant probability under the prior μ.

The proof

Since in any Nash equilibrium (σ_1', σ_2'), each payoff type ξ has the option of playing $\hat{\alpha}_1$ in every period, we have

$$U_1(\sigma', \xi) = (1-\delta)\sum_{t=0}^{\infty} \delta^t E_{\mathbf{P}}[u_1(\sigma_1'(h_1^t), \sigma_2'(h_2^t), \xi) \mid \xi]$$
$$\geq (1-\delta)\sum_{t=0}^{\infty} \delta^t E_{\widehat{\mathbf{P}}} u_1(\hat{\alpha}_1, \sigma_2'(h_2^t), \xi)$$

The proof

Since in any Nash equilibrium (σ_1', σ_2'), each payoff type ξ has the option of playing $\hat{\alpha}_1$ in every period, we have

$$
\begin{aligned}
U_1(\sigma', \xi) &= (1-\delta) \sum_{t=0}^{\infty} \delta^t E_{\mathbf{P}}[u_1(\sigma_1'(h_1^t), \sigma_2'(h_2^t), \xi) \mid \xi] \\
&\geq (1-\delta) \sum_{t=0}^{\infty} \delta^t E_{\widehat{\mathbf{P}}} u_1(\hat{\alpha}_1, \sigma_2'(h_2^t), \xi) \\
&\geq (1-\delta) \sum_{t=0}^{\infty} \delta^t E_{\widehat{\mathbf{P}}} \underline{v}_{\hat{\alpha}_1}^{\xi}(d(\hat{p} \| p(h_2^t)))
\end{aligned}
$$

The proof

Since in any Nash equilibrium (σ'_1, σ'_2), each payoff type ξ has the option of playing $\hat{\alpha}_1$ in every period, we have

$$\begin{aligned} U_1(\sigma', \xi) &= (1-\delta)\sum\nolimits_{t=0}^{\infty} \delta^t E_{\mathbf{P}}[u_1(\sigma'_1(h_1^t), \sigma'_2(h_2^t), \xi) \mid \xi] \\ &\geq (1-\delta)\sum\nolimits_{t=0}^{\infty} \delta^t E_{\widehat{\mathbf{P}}} u_1(\hat{\alpha}_1, \sigma'_2(h_2^t), \xi) \\ &\geq (1-\delta)\sum\nolimits_{t=0}^{\infty} \delta^t E_{\widehat{\mathbf{P}}} \underline{v}_{\hat{\alpha}_1}^{\xi}(d(\hat{p} \| p(h_2^t))) \\ &\geq (1-\delta)\sum\nolimits_{t=0}^{\infty} \delta^t E_{\widehat{\mathbf{P}}} \underline{w}_{\hat{\alpha}_1}^{\xi}(d(\hat{p} \| p(h_2^t))) \end{aligned}$$

The proof

Since in any Nash equilibrium (σ_1', σ_2'), each payoff type ξ has the option of playing $\hat{\alpha}_1$ in every period, we have

$$
\begin{aligned}
U_1(\sigma', \xi) &= (1-\delta)\sum_{t=0}^{\infty} \delta^t E_{\mathbf{P}}[u_1(\sigma_1'(h_1^t), \sigma_2'(h_2^t), \xi) \mid \xi] \\
&\geq (1-\delta)\sum_{t=0}^{\infty} \delta^t E_{\widehat{\mathbf{P}}} u_1(\hat{\alpha}_1, \sigma_2'(h_2^t), \xi) \\
&\geq (1-\delta)\sum_{t=0}^{\infty} \delta^t E_{\widehat{\mathbf{P}}} \underline{v}_{\hat{\alpha}_1}^{\xi}(d(\hat{p} \| p(h_2^t))) \\
&\geq (1-\delta)\sum_{t=0}^{\infty} \delta^t E_{\widehat{\mathbf{P}}} \underline{w}_{\hat{\alpha}_1}^{\xi}(d(\hat{p} \| p(h_2^t))) \\
&\geq \underline{w}_{\hat{\alpha}_1}^{\xi}\left((1-\delta)\sum_{t=0}^{\infty} \delta^t E_{\widehat{\mathbf{P}}} d(\hat{p} \| p(h_2^t))\right)
\end{aligned}
$$

The proof

Since in any Nash equilibrium (σ_1', σ_2'), each payoff type ξ has the option of playing $\hat{\alpha}_1$ in every period, we have

$$
\begin{aligned}
U_1(\sigma', \xi) &= (1-\delta)\sum\nolimits_{t=0}^{\infty} \delta^t E_{\mathbf{P}}[u_1(\sigma_1'(h_1^t), \sigma_2'(h_2^t), \xi) \mid \xi] \\
&\geq (1-\delta)\sum\nolimits_{t=0}^{\infty} \delta^t E_{\widehat{\mathbf{P}}} u_1(\hat{\alpha}_1, \sigma_2'(h_2^t), \xi) \\
&\geq (1-\delta)\sum\nolimits_{t=0}^{\infty} \delta^t E_{\widehat{\mathbf{P}}} \underline{v}_{\hat{\alpha}_1}^{\xi}(d(\hat{p} \| p(h_2^t))) \\
&\geq (1-\delta)\sum\nolimits_{t=0}^{\infty} \delta^t E_{\widehat{\mathbf{P}}} \underline{w}_{\hat{\alpha}_1}^{\xi}(d(\hat{p} \| p(h_2^t))) \\
&\geq \underline{w}_{\hat{\alpha}_1}^{\xi}\left((1-\delta)\sum\nolimits_{t=0}^{\infty} \delta^t E_{\widehat{\mathbf{P}}} d(\hat{p} \| p(h_2^t))\right) \\
&\geq \underline{w}_{\hat{\alpha}_1}^{\xi}\left(-(1-\delta)\log \mu(\hat{\xi})\right) = \underline{w}_{\hat{\alpha}_1}^{\xi}(\hat{\varepsilon}).
\end{aligned}
$$

Patient player 1

Corollary

Suppose the action type $\hat{\xi} = \xi(\hat{\alpha}_1)$ *has positive prior probability,* $\mu(\hat{\xi}) > 0$*, for some potentially mixed action* $\hat{\alpha}_1 \in \Delta(A_1)$*. Then, for all* $\xi \in \Xi_1$ *and* $\eta > 0$*, there exists a* $\bar{\delta} < 1$ *such that, for all* $\delta \in (\bar{\delta}, 1)$*, player* 1 *type* ξ*'s payoff in any Nash equilibrium of the game* $\Gamma(\mu, \delta)$ *is greater than or equal to*

$$\underline{v}^{\xi}_{\hat{\alpha}_1}(0) - \eta.$$

When does $B_0^d(\alpha_1) = BR(\alpha_1)$?

- Suppose $\rho(\cdot|a_1) \neq \rho(\cdot|a_1')$ for all $a_1 \neq a_1'$. Then pure action Stackelberg payoff is a reputation lower bound provided the simple Stackelberg type has positive prob.
- Suppose $\rho(\cdot|\alpha_1) \neq \rho(\cdot|\alpha_1')$ for all $\alpha_1 \neq \alpha_1'$. Then mixed action Stackelberg payoff is a reputation lower bound provided the prior includes in its support a dense subet of $\Delta(A_1)$.

How general is the result?

- The same argument (with slightly worse notation) works if the monitoring distribution depends on both players actions (though statistical identifiability is a more demanding requirement, particularly for extensive form stage games).
- The same argument (with slightly worse notation) also works if the game has private monitoring. Indeed, notice that player 1 observing the signal played no role in the proof.