# Repeated Games and Reputations: The Basic Structure

George J. Mailath

University of Pennsylvania
and
Australian National University

CEMMAP Lectures
November 17-18, 2016

The slides and associated bibliography
are on my webpage
http://economics.sas.upenn.edu/∼gmailath

# Introduction

- The theory of repeated games provides a central underpinning for our understanding of social, political, and economic institutions, both formal and informal.
- A key ingredient in understanding institutions and other long run relationships is the role of
  - shared expectations about behavioral norms (cultural beliefs), and
  - sanctions in ensuring that people follow the "rules."
- Repeated games allow for a clean description of both the myopic incentives that agents have to not follow the rules and, via appropriate specifications of future behavior (and so rewards and punishments), the incentives that deter such opportunistic behavior.

# Examples of Long-Run Relationships and Opportunistic Behavior

- Buyer-seller.
  The seller selling an inferior good.
- Employer and employees.
  Employees shirking on the job, employer reneging on implicit terms of employment.
- A government and its citizens.
  Government expropriates (taxes) all profits from investments.
- World Trade Organization
  Imposing tariffs to protect a domestic industry.
- Cartels
  A firm exceeding its share of the monopolistic output.

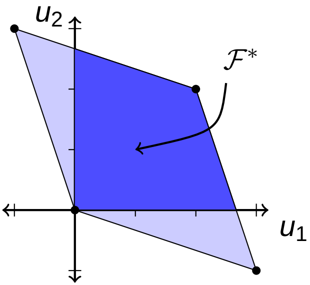
# Two particularly interesting examples

1. Dispute Resolution.
   Ellickson (1991) presents evidence that neighbors in Shasta County, CA, resolve disputes arising from the damage created by escaped cattle in ways that both ignore legal liability and are supported by intertemporal incentives.
2. Traders selling goods on consignment.
   Grief (1994) documents how the Maghribi and Genoese merchants deterred their agents from misreporting that goods were damaged in transport, and so were worth less. These two communities of merchants did this differently, and in ways consistent with the different cultural characteristics of the communities and repeated game analysis.

# The Leading Example

## The prisoners' dilemma as a partnership

| | $E$ | $S$ |
|---|---|---|
| $E$ | $2, 2$ | $-1, 3$ |
| $S$ | $3, -1$ | $0, 0$ |

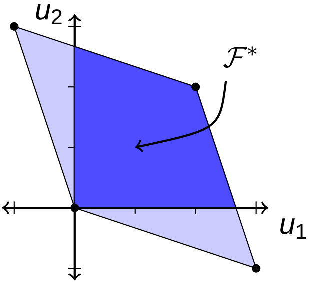



- Each player can guarantee herself a payoff of 0.
  A payoff vector is individually rational if it gives each player at least their guarantee.
- $\mathcal{F}^*$ is the set of feasible and individually rational payoffs.

# The Leading Example

## The prisoners' dilemma as a partnership

| | $E$ | $S$ |
|---|---|---|
| $E$ | $2, 2$ | $-1, 3$ |
| $S$ | $3, -1$ | $0, 0$ |




- Each player can guarantee herself a payoff of 0.
  A payoff vector is individually rational if it gives each player at least their guarantee.
- $\mathcal{F}^*$ is the set of feasible and individually rational payoffs.
- In the static (one shot ) play, each player will play $S$, resulting in the inefficient $SS$ outcome.

# Intertemporal Incentives

- Suppose the game is repeated (once), and payoffs are added.
- We "know" *SS* will be played in last period, so no intertemporal incentives.
- Infinite horizon—relationship never ends.
  The infinite stream of payoffs $(u_i^0, u_i^1, u_i^2, \ldots)$ is evaluated as the (average) discounted sum

$$\sum\nolimits_{t \geq 0} (1-\delta)\delta^t u_i^t.$$

- Individual $i$ is indifferent between $0, 1, 0, \ldots$ and $\delta, 0, 0 \ldots$.
- The normalization $(1-\delta)$ implies that repeated game payoffs are comparable to stage game payoffs.
  The infinite constant stream of 1 utils has a value of 1.

- A strategy $\sigma_i$ for individual $i$ describes how that individual behaves (at each point of time and after any possible history).
- A strategy profile $\sigma = (\sigma_1, \ldots, \sigma_n)$ describes how everyone behaves (at each point of...).

**Definition**

The profile $\sigma^*$ is a Nash equilibrium if for all individuals $i$, when everyone else is behaving according to $\sigma^*_{-i}$, then $i$ is also willing to behave as described by $\sigma_i^*$.
The profile $\sigma^*$ is a subgame perfect equilibrium if for all histories of play, the behavior described (induced) by the profile is a Nash equilibrium.

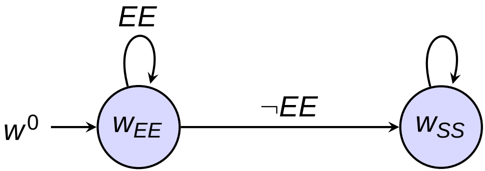
- Useful to think of social norms as equilibria: shared expectations over behavior that provide appropriate sanctions to deter deviations.

# Characterizing Equilibria

- Difficult problem: many possible deviations after many different histories.
- But repeated games are recursive, and the one shot deviation principle (from dynamic programming) holds.
- Simple penal codes (Abreu, 1988): use $i$'s worst eq to punish any (and all) deviation by $i$.

# Prisoners' Dilemma

## Grim Trigger




This is an equilibrium if

$$(1-\delta)\times 2+\delta\times 2 = 2 \geq (1-\delta)\times 3+\delta\times 0$$
$$\Rightarrow \delta \geq \tfrac{1}{3}.$$

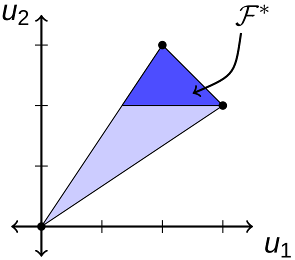
Grim trigger is subgame perfect: always $S$ is a Nash eq (because $SS$ is an eq of the stage game and in $w_{SS}$ behavior is history independent).

# The need for credibility of punishments

## The Purchase Game

A buyer and seller:

| | Buy | Don't buy |
|---|---|---|
| High effort | 2,3 | 0, 0 |
| Low effort | 3,2 | 0, 0 |



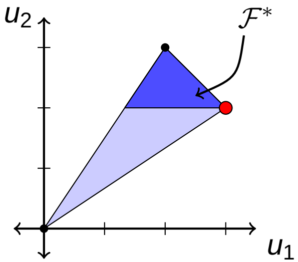

- The seller can guarantee himself 0, while the buyer can guarantee herself 2.

# The need for credibility of punishments

## The Purchase Game

A buyer and seller:

| | Buy | Don't buy |
|---|---|---|
| High effort | 2,3 | 0, 0 |
| Low effort | 3,2 | 0, 0 |



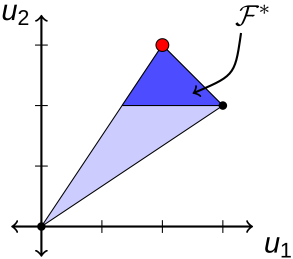

- The seller can guarantee himself 0, while the buyer can guarantee herself 2.
- There is an equilibrium in which the seller always chooses low effort and the buyer always buys.

# The need for credibility of punishments

## The Purchase Game

A buyer and seller:

| | Buy | Don't buy |
|---|---|---|
| High effort | 2,3 | 0, 0 |
| Low effort | 3,2 | 0, 0 |

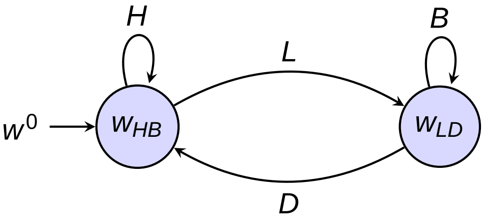



- The seller can guarantee himself 0, while the buyer can guarantee herself 2.
- There is an equilibrium in which the seller always chooses low effort and the buyer always buys.
- Is there a social norm in which the buyer threatens not to buy unless the seller chooses high effort?
  - Need to provide incentives for the buyer to do so.

# Why the buyer is willing to punish

Suppose, after the seller "cheats" the buyer by choosing low effort, the buyer expects the seller to continue to choose low effort until the buyer punishes him by not buying.

| | $B$ | $D$ |
|---|---|---|
| $H$ | $2, 3$ | $0, 0$ |
| $L$ | $3, 2$ | $0, 0$ |

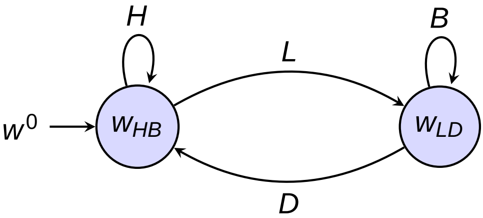



- The seller chooses high effort as long as $\delta \geq \frac{1}{2}$.
- The buyer is willing to punish as long as $\delta \geq \frac{2}{3}$.

# Why the buyer is willing to punish

Suppose, after the seller "cheats" the buyer by choosing low effort, the buyer expects the seller to continue to choose low effort until the buyer punishes him by not buying.

| | $B$ | $D$ |
|---|---|---|
| $H$ | $2, 3$ | $0, 0$ |
| $L$ | $3, 2$ | $0, 0$ |




- The seller chooses high effort as long as $\delta \geq \frac{1}{2}$.
- The buyer is willing to punish as long as $\delta \geq \frac{2}{3}$.
- This is a carrot and stick punishment (Abreu, 1986).

# The Game with Perfect Monitoring

- Action space for $i$ is $A_i$, with typical action $a_i \in A_i$.
- An action profile is denoted $a = (a_1, \ldots, a_n)$, with associated flow payoffs $u_i(a)$.
- Infinite stream of payoffs $(u_i^0, u_i^1, u_i^2, \ldots)$ is evaluated as the (average) discounted sum

$$\sum\nolimits_{t \geq 0} (1-\delta)\delta^t u_i^t,$$

  where $\delta \in [0,1)$ is the discount factor.
- Perfect monitoring: At the end of each period, all players observe the action profile $a$ chosen.
- History to date $t$: $h^t \equiv (a^0, \ldots, a^{t-1}) \in A^t \equiv H^t$; $H^0 \equiv \{\varnothing\}$.
- Set of all possible histories: $H \equiv \cup_{t=0}^{\infty} H^t$.
- Strategy for player $i$ is denoted $s_i : H \to A_i$.
- Set of all strategies for player $i$ is $S_i$.

# Automaton Representation of Behavior

An automaton is the tuple $(\mathcal{W}, w^0, f, \tau)$, where

- $\mathcal{W}$ is set of states,
- $w^0$ is initial state,
- $f : \mathcal{W} \to A$ is output function (decision rule), and
- $\tau : \mathcal{W} \times A \to \mathcal{W}$ is transition function.

Any automaton $(\mathcal{W}, w^0, f, \tau)$ induces a strategy profile. Define

$$\tau(w, h^t) := \tau(\tau(w, h^{t-1}), a^{t-1}).$$

The induced strategy $s$ is given by $s(\varnothing) = f(w^0)$ and

$$s(h^t) = f(\tau(w^0, h^t)), \quad \forall h^t \in H \backslash \{\varnothing\}.$$

Every profile can be represented by an automaton (set $\mathcal{W} = H$).

# Nash Equilibrium

**Definition**

An automaton is a Nash equilibrium if the strategy profile $s$ represented by the automaton is a Nash equilibrium.

# Subgames and Continuation Play

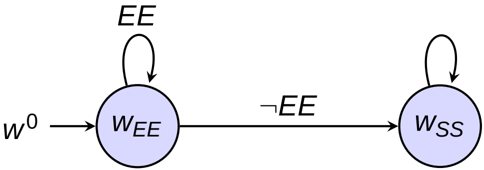
- Each history $h^t$ reaches ("indexes") a distinct subgame.
- Suppose $s$ is represented by $(\mathcal{W}, w^0, f, \tau)$. Recall that

$$\tau(w^0, h^t) := \tau(\tau(w^0, h^{t-1}), a^{t-1}).$$

- The continuation strategy profile after a history $h^t$, $s|_{h^t}$ is represented by the automaton $(\mathcal{W}, w^t, f, \tau)$, where

$$w^t := \tau(w^0, h^t).$$

- Grim Trigger after any $h^t = (EE)^t$:



# Subgames and Continuation Play

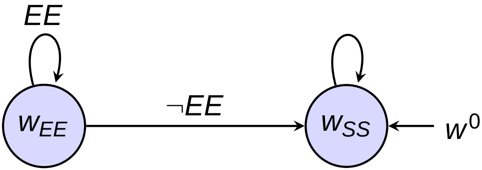
- Each history $h^t$ reaches ("indexes") a distinct subgame.
- Suppose $s$ is represented by $(\mathcal{W}, w^0, f, \tau)$. Recall that

$$\tau(w^0, h^t) := \tau(\tau(w^0, h^{t-1}), a^{t-1}).$$

- The continuation strategy profile after a history $h^t$, $s|_{h^t}$ is represented by the automaton $(\mathcal{W}, w^t, f, \tau)$, where

$$w^t := \tau(w^0, h^t).$$

- Grim Trigger after $h^t$ with an $S$ (equivalent to always $SS$):



# Subgame Perfection

## Definition

The state $w \in \mathcal{W}$ of an automaton $(\mathcal{W}, w^0, f, \tau)$ is reachable from $w^0$ if $w = \tau(w^0, h^t)$ for some history $h^t \in H$. Denote the set of states reachable from $w^0$ by $\mathcal{W}(w^0)$.

## Definition

The automaton $(\mathcal{W}, w^0, f, \tau)$ is a subgame perfect equilibrium if for all states $w \in \mathcal{W}(w^0)$, the automaton $(\mathcal{W}, w, f, \tau)$ is a Nash equilibrium.

The automaton $(\mathcal{W}, w, f, \tau)$ induces the sequences

$$
\begin{aligned}
\hat{w}^0 &:= w, & a^0 &:= f(\hat{w}^0) \\
\hat{w}^1 &:= \tau(\hat{w}^0, a^0), & a^1 &:= f(\hat{w}^1), \\
\hat{w}^2 &:= \tau(\hat{w}^1, a^1), & a^2 &:= f(\hat{w}^2), \\
&\vdots & &\vdots
\end{aligned}
$$

Given an automaton $(\mathcal{W}, w^0, f, \tau)$, let $V_i(w)$ be $i$'s value from being in the state $w \in \mathcal{W}$, i.e.,

$$
\begin{aligned}
V_i(w) &= (1-\delta)u_i(f(\hat{w}^0)) + \delta V_i(\tau(\hat{w}^0, f(\hat{w}^0))) \\
&= (1-\delta)u_i(a^0) + \delta\{(1-\delta)u_i(a^1) + \delta V_i(\hat{w}^2)\} \\
&\;\;\vdots \\
&= (1-\delta)\sum_t \delta^t u_i(a^t).
\end{aligned}
$$

# Principle of No One-Shot Deviation

## Definition

Player $i$ has a profitable one-shot deviation from $(\mathcal{W}, w^0, f, \tau)$, if there is a state $w \in \mathcal{W}(w^0)$ and some action $a_i \in A_i$ such that

$$V_i(w) < (1-\delta)u_i(a_i, f_{-i}(w)) + \delta V_i(\tau(w, (a_i, f_{-i}(w)))).$$

# Principle of No One-Shot Deviation

## Definition

Player $i$ has a profitable one-shot deviation from $(\mathcal{W}, w^0, f, \tau)$, if there is a state $w \in \mathcal{W}(w^0)$ and some action $a_i \in A_i$ such that

$$V_i(w) < (1-\delta)u_i(a_i, f_{-i}(w)) + \delta V_i(\tau(w, (a_i, f_{-i}(w)))).$$

## Theorem

*An automaton is subgame perfect iff there are no profitable one-shot deviations.*

# Principle of No One-Shot Deviation

## Definition

Player $i$ has a profitable one-shot deviation from $(\mathcal{W}, w^0, f, \tau)$, if there is a state $w \in \mathcal{W}(w^0)$ and some action $a_i \in A_i$ such that

$$V_i(w) < (1-\delta)u_i(a_i, f_{-i}(w)) + \delta V_i(\tau(w, (a_i, f_{-i}(w)))).$$

## Theorem

*An automaton is subgame perfect iff there are no profitable one-shot deviations.*

## Corollary

*The automaton $(\mathcal{W}, w^0, f, \tau)$ is subgame perfect iff, for all $w \in \mathcal{W}(w^0)$, $f(w)$ is a Nash eq of the normal form game with payoff function $g^w : A \to \mathbb{R}^n$, where*

$$g_i^w(a) = (1-\delta)u_i(a) + \delta V_i(\tau(w, a)).$$

# SGP if No Profitable One-Shot Deviations

## Proof I

- Let $\widetilde{V}_i(w)$ be player $i$'s payoff from the best response to $(\mathcal{W}, w, f_{-i}, \tau)$ (i.e., the strategy profile for the other players specified by the automaton with initial state $w$). Then

$$\widetilde{V}_i(w) = \max_{a_i \in A_i} \left\{ (1-\delta)u_i(a_i, f_{-i}(w)) + \delta \widetilde{V}_i(\tau(w, (a_i, f_{-i}(w)))) \right\}.$$

- Note that $\widetilde{V}_i(w) \geq V_i(w)$ for all $w$. Denote by $\bar{w}_i$, the state that maximizes $\widetilde{V}_i(w) - V_i(w)$ (if there is more than one, choose one arbitrarily).
- If $(\mathcal{W}, w^0, f, \tau))$ is not SGP, then for some player $i$,

$$\widetilde{V}_i(\bar{w}_i) - V_i(\bar{w}_i) > 0.$$

# SGP iff No Profitable One-Shot Deviations

## Proof II

Then, for all $w$,

$$\widetilde{V}_i(\bar{w}_i) - V_i(\bar{w}_i) > \delta[\widetilde{V}_i(w) - V_i(w)],$$

and so (where $a_i^{\bar{w}_i}$ yields $\widetilde{V}_i(\bar{w}_i)$)

$$\begin{aligned}\widetilde{V}_i(\bar{w}_i) - V_i(\bar{w}_i) \\ > \delta[\widetilde{V}_i(\tau(\bar{w}_i, (a_i^{\bar{w}_i}, f_{-i}(\bar{w}_i)))) - V_i(\tau(\bar{w}_i, (a_i^{\bar{w}_i}, f_{-i}(\bar{w}_i))))]\end{aligned}$$

# SGP iff No Profitable One-Shot Deviations

## Proof II

Then, for all $w$,

$$\widetilde{V}_i(\bar{w}_i) - V_i(\bar{w}_i) > \delta[\widetilde{V}_i(w) - V_i(w)],$$

and so (where $a_i^{\bar{w}_i}$ yields $\widetilde{V}_i(\bar{w}_i)$)

$$\begin{aligned}
&\widetilde{V}_i(\bar{w}_i) - V_i(\bar{w}_i) \\
&\qquad > \delta[\widetilde{V}_i(\tau(\bar{w}_i, (a_i^{\bar{w}_i}, f_{-i}(\bar{w}_i)))) - V_i(\tau(\bar{w}_i, (a_i^{\bar{w}_i}, f_{-i}(\bar{w}_i))))] \\
&\qquad\quad + [(1-\delta)u_i(a_i^{\bar{w}_i}, f_{-i}(\bar{w}_i)) - (1-\delta)u_i(a_i^{\bar{w}_i}, f_{-i}(\bar{w}_i))]
\end{aligned}$$

# SGP iff No Profitable One-Shot Deviations

## Proof II

Then, for all $w$,

$$\widetilde{V}_i(\bar{w}_i) - V_i(\bar{w}_i) > \delta[\widetilde{V}_i(w) - V_i(w)],$$

and so (where $a_i^{\bar{w}_i}$ yields $\widetilde{V}_i(\bar{w}_i)$)

$$\begin{aligned}
&\widetilde{V}_i(\bar{w}_i) - V_i(\bar{w}_i) \\
&\quad > \delta[\widetilde{V}_i(\tau(\bar{w}_i, (a_i^{\bar{w}_i}, f_{-i}(\bar{w}_i)))) - V_i(\tau(\bar{w}_i, (a_i^{\bar{w}_i}, f_{-i}(\bar{w}_i))))] \\
&\quad\quad + [(1-\delta)u_i(a_i^{\bar{w}_i}, f_{-i}(\bar{w}_i)) - (1-\delta)u_i(a_i^{\bar{w}_i}, f_{-i}(\bar{w}_i))] \\
&\quad = \widetilde{V}_i(\bar{w}_i) \\
&\quad\quad - \left\{(1-\delta)u_i(a_i^{\bar{w}_i}, f_{-i}(\bar{w}_i)) + \delta V_i(\tau(\bar{w}_i, (a_i^{\bar{w}_i}, f_{-i}(\bar{w}_i))))\right\}.
\end{aligned}$$

# SGP iff No Profitable One-Shot Deviations

## Proof II

Then, for all $w$,

$$\widetilde{V}_i(\bar{w}_i) - V_i(\bar{w}_i) > \delta[\widetilde{V}_i(w) - V_i(w)],$$

and so (where $a_i^{\bar{w}_i}$ yields $\widetilde{V}_i(\bar{w}_i)$)

$$\begin{aligned}
\widetilde{V}_i(\bar{w}_i) - V_i(\bar{w}_i) & \\
>&\delta[\widetilde{V}_i(\tau(\bar{w}_i, (a_i^{\bar{w}_i}, f_{-i}(\bar{w}_i)))) - V_i(\tau(\bar{w}_i, (a_i^{\bar{w}_i}, f_{-i}(\bar{w}_i))))] \\
&+ [(1-\delta)u_i(a_i^{\bar{w}_i}, f_{-i}(\bar{w}_i)) - (1-\delta)u_i(a_i^{\bar{w}_i}, f_{-i}(\bar{w}_i))] \\
=&\widetilde{V}_i(\bar{w}_i) \\
&- \left\{(1-\delta)u_i(a_i^{\bar{w}_i}, f_{-i}(\bar{w}_i)) + \delta V_i(\tau(\bar{w}_i, (a_i^{\bar{w}_i}, f_{-i}(\bar{w}_i))))\right\}.
\end{aligned}$$

Thus,

$$(1-\delta)u_i(a_i^{\bar{w}_i}, f_{-i}(\bar{w}_i)) + \delta V_i(\tau(\bar{w}_i, (a_i^{\bar{w}_i}, f_{-i}(\bar{w}_i)))) > V_i(w_i),$$

that is, player $i$ has a profitable one-shot deviation at $\bar{w}_i$.

# Enforceability and Decomposability

## Definition

An action profile $a' \in A$ is enforced by the continuation promises $\gamma : A \to \mathbb{R}^n$ if $a'$ is a Nash eq of the normal form game with payoff function $g^{\gamma} : A \to \mathbb{R}^n$, where

$$g_i^{\gamma}(a) = (1 - \delta)u_i(a) + \delta\gamma_i(a).$$

# Enforceability and Decomposability

## Definition

An action profile $a' \in A$ is enforced by the continuation promises $\gamma : A \to \mathbb{R}^n$ if $a'$ is a Nash eq of the normal form game with payoff function $g^\gamma : A \to \mathbb{R}^n$, where

$$g_i^\gamma(a) = (1-\delta)u_i(a) + \delta\gamma_i(a).$$

## Definition

A payoff $v$ is decomposable on a set of payoffs $\mathcal{V}$ if there exists an action profile $a'$ enforced by some continuation promises $\gamma : A \to \mathcal{V}$ satisfying, for all $i$,

$$v_i = (1-\delta)u_i(a') + \delta\gamma_i(a').$$

The payoff $v$ is decomposed by $a'$ on $\mathcal{V}$.

# The Purchase Game 1

| | Buy | Don't buy |
|---|---|---|
| High effort | 2,3 | $0, 0$ |
| Low effort | 3,2 | $0, 0$ |

- Only *LB* can be enforced by constant continuation promises, and so
- only $(3, 2)$ can be decomposed on a singleton set, and that set is $\{(3, 2)\}$.

# The Purchase Game 2

| | Buy | Don't buy |
|---|---|---|
| High effort | 2,3 | $0, 0$ |
| Low effort | 3,2 | $0, 0$ |

Suppose $\mathcal{V} = \{(2\delta, 3\delta), (2, 3)\}$, and $\delta > \frac{2}{3}$.

- $(2, 3)$ is decomposed on $\mathcal{V}$ by *HB* and promises

$$\gamma(a) = \begin{cases} (2, 3), & \text{if } a_1 = H, \\ (2\delta, 3\delta), & \text{if } a_1 = L. \end{cases}$$

- $(2\delta, 3\delta)$ is decomposed on $\mathcal{V}$ by *LD* and promises

$$\gamma(a) = \begin{cases} (2, 3), & \text{if } a_2 = D, \\ (2\delta, 3\delta), & \text{if } a_2 = B. \end{cases}$$

- No one-shot deviation principle $\Longrightarrow$
  every payoff in $\mathcal{V}$ is a subgame perfect eq payoff.

# The Purchase Game 3

| | Buy | Don't buy |
|---|---|---|
| High effort | 2,3 | $0, 0$ |
| Low effort | 3,2 | $0, 0$ |

Suppose $\mathcal{V} = \{(2\delta, 3\delta), (2, 3)\}$, and $\delta > \frac{2}{3}$.

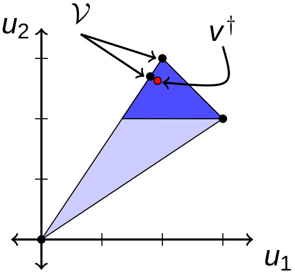
- $(3 - 3\delta + 2\delta^2, 2 - 2\delta + 3\delta^2) =: v^\dagger$ is decomposed on $\mathcal{V}$ by *LB* and the constant promises

$$\gamma(a) = (2\delta, 3\delta).$$

- So, payoffs outside $\mathcal{V}$ can also be decomposed on $\mathcal{V}$.
- No one-shot deviation principle $\implies$ $v^\dagger$ is a subgame perfect eq payoff.

# The Purchase Game 4

| | Buy | Don't buy |
|---|---|---|
| High effort | 2,3 | 0, 0 |
| Low effort | 3,2 | 0, 0 |



# Subgame Perfection redux

Let $\mathcal{E}^p(\delta) \subset \mathcal{F}^{p*}$ be the set of pure strategy subgame perfect equilibrium payoffs.

**Theorem**

*A payoff $v \in \mathbb{R}^n$ is decomposable on $\mathcal{E}^p(\delta)$ if, and only if, $v \in \mathcal{E}^p(\delta)$.*

# Subgame Perfection redux

Let $\mathcal{E}^p(\delta) \subset \mathcal{F}^{p*}$ be the set of pure strategy subgame perfect equilibrium payoffs.

**Theorem**

*A payoff $v \in \mathbb{R}^n$ is decomposable on $\mathcal{E}^p(\delta)$ if, and only if, $v \in \mathcal{E}^p(\delta)$.*

**Theorem**

*Suppose every payoff $v$ in some bounded set $\mathcal{V} \subset \mathbb{R}^n$ is decomposable with respect to $\mathcal{V}$. Then, $\mathcal{V} \subset \mathcal{E}^p(\delta)$.*

Any set of payoffs with the property described above is said to be self-generating.

# A Folk Theorem

- Intertemporal incentives allow for efficient outcomes, but also for inefficient outcomes, as well as crazy outcomes.
- This is illustrated by the "Folk" Theorem, so called because results of this type have been part of game theory folklore since at least the late sixties.

**The Discounted Folk Theorem (Fudenberg&Maskin 1986)**

Suppose $v$ is a feasible and strictly individually rational vector of payoffs. If the individuals are sufficiently patient (there exists $\underline{\delta} \in (0,1)$ such that for all $\delta \in (\underline{\delta}, 1)$), then there is a subgame perfect equilibrium with payoff $v$.

# Interpretation

- While efficient payoffs are consistent with equilibrium, so are many other payoffs, and associated behaviors. (Consistent with experimental evidence.)
- Moreover, multiple equilibria are consistent with the same payoff.
- The theorem does not justify restricting attention to efficient payoffs.

# Interpretation

- While efficient payoffs are consistent with equilibrium, so are many other payoffs, and associated behaviors. (Consistent with experimental evidence.)
- Moreover, multiple equilibria are consistent with the same payoff.
- The theorem does not justify restricting attention to efficient payoffs.

Nonetheless:

- In many situations, understanding the potential scope of equilibrium incentives helps us to understand possible plausible behaviors.
- Understanding what it takes to achieve efficiency gives us important insights into the nature of equilibrium incentives.
- It is sometimes argued that the punishments imposed are too severe. But this does simplify the analysis.

# What we learn from perfect monitoring

- Multiplicity of equilibria is to be expected.
  - This is necessary for repeated games to serve as a building block for any theory of institutions.
  - Selection of equilibrium can (should) be part of modelling.
- In general, efficiency requires being able to reward and punish individuals independently (this is the role of the full dimensionality assumption).
- Histories coordinate behavior to provide intertemporal incentives by punishing deviations. This requires monitoring (communication networks) and a future.
  - Intertemporal incentives require that individuals have something at stake: "Freedom's just another word for nothin' left to lose."

# Repeated Games and Reputations: Imperfect Public Monitoring

George J. Mailath

University of Pennsylvania
and
Australian National University

CEMMAP Lectures
November 17-18, 2016

The slides and associated bibliography
are on my webpage
http://economics.sas.upenn.edu/∼gmailath

# What we learned from perfect monitoring

- Multiplicity of equilibria is to be expected.
- In general, efficiency requires being able to reward and punish individuals independently.
- Histories coordinate behavior to provide intertemporal incentives by punishing deviations. This requires monitoring (communication networks) and a future.

# What we learned from perfect monitoring

- Multiplicity of equilibria is to be expected.
- In general, efficiency requires being able to reward and punish individuals independently.
- Histories coordinate behavior to provide intertemporal incentives by punishing deviations. This requires monitoring (communication networks) and a future.

But suppose deviations are not observed?
Suppose instead actions are only imperfectly observed.

# Collusion in Oligopoly

## Perfect Monitoring

- In each period, firms $i = 1, \ldots, n$ simultaneously choose quantities $q_i$.
- Firm $i$ profits

$$\pi_i(q_1, \ldots, q_n) = pq_i - c(q_i),$$

  where $p$ is market clearing price, and $c(q_i)$ is the cost of $q_i$.
- Suppose $p = P(\sum_i q_i)$ and $P$ is a strictly decreasing function of $Q := \sum_i q_i$.
- If firms are patient, there is a subgame perfect equilibrium in which the each firm sells $Q^m/n$, where $Q^m$ is monopoly output, supported by the threat that any deviation results in perpetual Cournot (static Nash) competition.

# Collusion in Oligopoly

## Imperfect Monitoring

- In each period, firms $i = 1, \ldots, n$ simultaneously choose quantities $q_i$.
- Firm $i$ profits

$$\pi_i(q_1, \ldots, q_n) = pq_i - c(q_i),$$

where $p$ is market clearing price, and $c(q_i)$ is the cost of $q_i$.

- Suppose $p = P(\sum_i q_i)$ and $P$ is a strictly decreasing function of $Q := \sum_i q_i$.
- Suppose now $q_1, \ldots, q_n$ are not public, but the market clearing price $p$ still is (so each firm knows its profit). Nothing changes! A deviation is still necessarily detected, since the market clearing price changes.

# Collusion in Oligopoly

## Noisy Imperfect Monitoring–Green and Porter (1984)

- In each period, firms $i = 1, \ldots, n$ simultaneously choose quantities $q_i$.
- Firm $i$ profits

$$\pi_i(q_1, \ldots, q_n) = pq_i - c(q_i),$$

where $p$ is market clearing price, and $c(q_i)$ is the cost of $q_i$.

- But suppose demand is random, so that the market clearing price $p$ is a function of$Q$ and a demand shock $\eta$. Moreover, suppose $p$ has full support for all $Q$.

$\implies$ no deviation is detected.

# Repeated Games with Noisy Imperfect Monitoring

- In a setting with noisy imperfect monitoring where it is impossible to detect deviations, are there still intertemporal incentives?
- If so, what is their nature?
- And, how effective are these intemporal incentives?

# Repeated Games with Noisy Imperfect Monitoring

- In a setting with noisy imperfect monitoring where it is impossible to detect deviations, are there still intertemporal incentives? Yes
- If so, what is their nature?
- And, how effective are these intemporal incentives?

# Repeated Games with Noisy Imperfect Monitoring

- In a setting with noisy imperfect monitoring where it is impossible to detect deviations, are there still intertemporal incentives?
- If so, what is their nature?
- And, how effective are these intemporal incentives? Surprisingly strong!

# Repeated Games with Imperfect Public Monitoring

## Structure 1

- Action space for $i$ is $A_i$, with typical action $a_i \in A_i$.
- Profile $a$ is not observed.
- All players observe a public signal $y \in Y$, $|Y| < \infty$, with

$$\Pr\{y \mid (a_1, \ldots, a_n)\} =: \rho(y \mid a).$$

- Since $y$ is a possibly noisy signal of the action profile $a$ in that period, the actions are imperfectly monitored.
- Since the signal is public (observed by all players), the game is said to have public monitoring.
- Assume $Y$ is finite.
- $u_i^* : A_i \times Y \to \mathbb{R}$, $i$'s ex post or realized payoff.
- Stage game (ex ante) payoffs:

$$u_i(a) \equiv \sum_{y \in Y} u_i^*(a_i, y)\rho(y \mid a).$$

# Ex post payoffs

## Oligopoly with imperfect monitoring

- Ex post payoffs are given by realized profits,

$$u_i^*(q_i, p) = pq_i - c(q_i),$$

where $p$ is the public signal.

- Ex ante payoffs are given by expected profits,

$$\begin{aligned} u_i(q_1, \ldots, q_n) &= E[pq_i - c(q_i) \mid q_1, \ldots q_n] \\ &= E[p \mid q_1, \ldots q_n]q_i - c(q_i). \end{aligned}$$

# Ex post payoffs II

## Prisoners' Dilemma with Noisy Monitoring

- There is a noisy signal of actions (output), $y \in \{\underline{y}, \bar{y}\} =: Y$,

$$\Pr(\bar{y} \mid a) := \rho(\bar{y} \mid a) = \begin{cases} p, & \text{if } a = EE, \\ q, & \text{if } a = SE \text{ or } SE, \text{ and} \\ r, & \text{if } a = SS. \end{cases}$$

- Player $i$'s ex post payoffs

| | $\bar{y}$ | $\underline{y}$ |
|---|---|---|
| $E$ | $\frac{(3-p-2q)}{(p-q)}$ | $-\frac{(p+2q)}{(p-q)}$ |
| $S$ | $\frac{3(1-r)}{(q-r)}$ | $-\frac{3r}{(q-r)}$ |

ex ante payoffs

| | $E$ | $S$ |
|---|---|---|
| $E$ | $2, 2$ | $-1, 3$ |
| $S$ | $3, -1$ | $0, 0$ |

# Repeated Games with Imperfect Public Monitoring

## Structure 2

- Public histories:

$$H \equiv \cup_{t=0}^{\infty} Y^t,$$

  with $h^t \equiv (y^0, \ldots, y^{t-1})$ being a $t$ period history of public signals ($Y^0 \equiv \{\varnothing\}$).

- Public strategies:

$$s_i : H \to A_i.$$

# Automaton Representation of Public Strategies

An automaton is the tuple $(\mathcal{W}, w^0, f, \tau)$, where

- $\mathcal{W}$ is set of states,
- $w^0$ is initial state,
- $f : \mathcal{W} \to A$ is output function (decision rule), and
- $\tau : \mathcal{W} \times Y \to \mathcal{W}$ is transition function.

The automaton is strongly symmetric if $f_i(w) = f_j(w) \quad \forall i, j, w$.

Any automaton $(\mathcal{W}, w^0, f, \tau)$ induces a strategy profile. Define

$$\tau(w, h^t) := \tau(\tau(w, h^{t-1}), y^{t-1}).$$

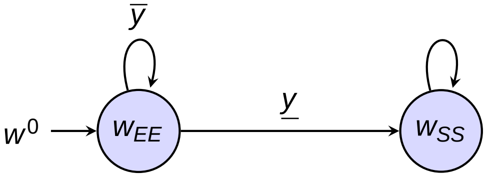
The induced strategy $s$ is given by $s(\varnothing) = f(w^0)$ and

$$s(h^t) = f(\tau(w^0, h^t)), \quad \forall h^t \in H \backslash \{\varnothing\}.$$

Every public profile can be represented by an automaton (set $\mathcal{W} = H$).

# Prisoners' Dilemma with Noisy Monitoring

## Grim Trigger




This is an eq if

$$\begin{aligned} V &= (1-\delta)2 + \delta[pV + (1-p)\times 0] \\ &\geq (1-\delta)3 + \delta[qV + (1-q)\times 0] \\ &\Rightarrow \frac{2\delta(p-q)}{(1-\delta p)} \geq 1 \quad \iff \quad \delta \geq \frac{1}{3p-2q}. \end{aligned}$$

Note that

$$V = \frac{2(1-\delta)}{(1-\delta p)},$$

and so $\lim_{\delta\to 1} V = 0$.

# Equilibrium Notion

- Game has no proper subgames, so how to usefully capture sequential rationality?

# Equilibrium Notion

- Game has no proper subgames, so how to usefully capture sequential rationality?
- A public strategy for an individual ignores that individual's private actions, so that behavior only depends on public information. Every player has a public strategy best response when all other players are playing public strategies.

**Definition**

The automaton $(\mathcal{W}, w^0, f, \tau)$ is a perfect public equilibrium (PPE) if for all states $w \in \mathcal{W}(w^0)$, the automaton $(\mathcal{W}, w, f, \tau)$ is a Nash equilibrium.

# Principle of No One-Shot Deviation

## Definition

Player $i$ has a profitable one-shot deviation from $(\mathcal{W}, w^0, f, \tau)$, if there is a state $w \in \mathcal{W}(w^0)$ and some action $a_i \in A_i$ such that

$$V_i(w) < (1-\delta)u_i(a_i, f_{-i}(w)) + \delta \sum_y V_i(\tau(w,y))\rho(y \mid (a_i, f_{-i}(w))).$$

# Principle of No One-Shot Deviation

## Definition

Player $i$ has a profitable one-shot deviation from $(\mathcal{W}, w^0, f, \tau)$, if there is a state $w \in \mathcal{W}(w^0)$ and some action $a_i \in A_i$ such that

$$V_i(w) < (1-\delta)u_i(a_i, f_{-i}(w)) + \delta \sum_y V_i(\tau(w, y))\rho(y \mid (a_i, f_{-i}(w))).$$

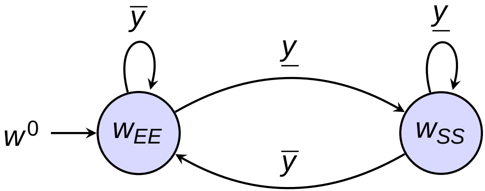
## Theorem

*The automaton $(\mathcal{W}, w^0, f, \tau)$ is a PPE iff there are no profitable one-shot deviations, i.e, for all $w \in \mathcal{W}(w^0)$, $f(w)$ is a Nash eq of the normal form game with payoff function $g^w : A \to \mathbb{R}^n$, where*

$$g_i^w(a) = (1-\delta)u_i(a) + \delta \sum_y V_i(\tau(w, y))\rho(y \mid a).$$

# Prisoners' Dilemma with Noisy Monitoring

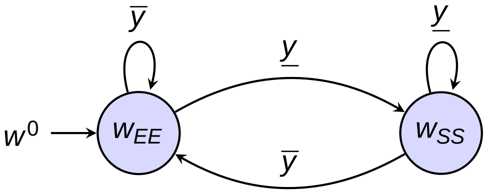
## Bounded Recall




- $V(w_{EE}) = (1-\delta)2 + \delta\{pV(w_{EE}) + (1-p)V(w_{SS})\}$
  $V(w_{SS}) = \delta\{rV(w_{EE}) + (1-r)V(w_{SS})\}$
- $V(w_{EE}) > V(w_{SS})$, but $V(w_{EE}) - V(W_{SS}) \to 0$ as $\delta \to 1$.
- At $w_{EE}$, $EE$ is a Nash eq of $g^{w_{EE}}$ if $\delta \geq (3p - 2q - r)^{-1}$.
- At $w_{SS}$, $SS$ is a Nash eq of $g^{w_{SS}}$ if $\delta \leq (p + 2q - 3r)^{-1}$.

# Prisoners' Dilemma with Noisy Monitoring

## Bounded Recall




- $V(w_{EE}) = (1-\delta)2 + \delta\{pV(w_{EE}) + (1-p)V(w_{SS})\}$
  $V(w_{SS}) = \delta\{rV(w_{EE}) + (1-r)V(w_{SS})\}$
- $V(w_{EE}) > V(w_{SS})$, but $V(w_{EE}) - V(W_{SS}) \to 0$ as $\delta \to 1$.
- At $w_{EE}$, $EE$ is a Nash eq of $g^{w_{EE}}$ if $\delta \geq (3p - 2q - r)^{-1}$.
- At $w_{SS}$, $SS$ is a Nash eq of $g^{w_{SS}}$ if $\delta \leq (p + 2q - 3r)^{-1}$.
- PPE if $(3p - 2q - r)^{-1} \leq \delta \leq (p + 2q - 3r)^{-1}$.

# Characterizing PPE

- A major conceptual breakthrough was to focus on continuation values in the description of equilibrium, rather than focusing on behavior directly.
- This yields a more transparent description of incentives, and an informative characterization of equilibrium payoffs.
- The cost is that we know little about the details of behavior underlying most of the equilibria, and so have little sense which of these equilibria are plausible descriptions of behavior.

# Enforceability and Decomposability

## Definition

An action profile $a' \in A$ is enforced by the continuation promises $\gamma : Y \to \mathbb{R}^n$ if $a'$ is a Nash eq of the normal form game with payoff function $g^\gamma : A \to \mathbb{R}^n$, where

$$g_i^\gamma(a) = (1 - \delta)u_i(a) + \delta \sum_y \gamma_i(y)\rho(y \mid a).$$

# Enforceability and Decomposability

## Definition

An action profile $a' \in A$ is enforced by the continuation promises $\gamma : Y \to \mathbb{R}^n$ if $a'$ is a Nash eq of the normal form game with payoff function $g^\gamma : A \to \mathbb{R}^n$, where

$$g_i^\gamma(a) = (1-\delta)u_i(a) + \delta \sum_y \gamma_i(y)\rho(y \mid a).$$

## Definition

A payoff $v$ is decomposable on a set of payoffs $\mathcal{V}$ if there exists an action profile $a'$ enforced by some continuation promises $\gamma : Y \to \mathcal{V}$ satisfying, for all $i$,

$$v_i = (1-\delta)u_i(a') + \delta \sum_y \gamma_i(y)\rho(y \mid a').$$

# Characterizing PPE

## The Role of Continuation Values

- Let $\mathcal{E}^p(\delta) \subset \mathcal{F}^*$ be the set of (pure strategy) PPE.
- If $v \in \mathcal{E}^p(\delta)$, then there exists $a' \in A$ and $\gamma : Y \to \mathcal{E}^p(\delta)$ so that, for all $i$,

$$\begin{aligned} v_i &= (1-\delta)u_i(a') + \delta \sum\nolimits_y \gamma_i(y)\rho(y \mid a') \\ &\geq (1-\delta)u_i(a_i, a'_{-i}) + \delta \sum\nolimits_y \gamma_i(y)\rho(y \mid a_i, a'_{-i}) \quad \forall a_i \in A_i. \end{aligned}$$

  That is, $v$ is decomposed on $\mathcal{E}^p(\delta)$.

# Characterizing PPE

## The Role of Continuation Values

- Let $\mathcal{E}^p(\delta) \subset \mathcal{F}^*$ be the set of (pure strategy) PPE.
- If $v \in \mathcal{E}^p(\delta)$, then there exists $a' \in A$ and $\gamma : Y \to \mathcal{E}^p(\delta)$ so that, for all $i$,

$$v_i = (1-\delta)u_i(a') + \delta \sum_y \gamma_i(y)\rho(y \mid a')$$
$$\geq (1-\delta)u_i(a_i, a'_{-i}) + \delta \sum_y \gamma_i(y)\rho(y \mid a_i, a'_{-i}) \quad \forall a_i \in A_i.$$

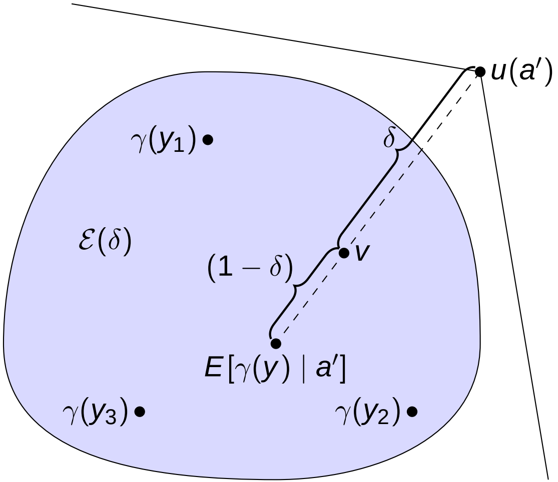
That is, $v$ is decomposed on $\mathcal{E}^p(\delta)$.

**Theorem (Self-generation, Abreu, Pearce, Stacchetti, 1990)**

*$B \subset \mathcal{E}^p(\delta)$ if and only if for all $v \in B$, $B$ bounded, there exists $a' \in A$ and $\gamma : Y \to B$ so that, for all $i$,*

$$v_i = (1-\delta)u_i(a') + \delta \sum_y \gamma_i(y)\rho(y \mid a')$$
$$\geq (1-\delta)u_i(a_i, a'_{-i}) + \delta \sum_y \gamma_i(y)\rho(y \mid a_i, a'_{-i}) \quad \forall a_i \in A_i.$$

# Decomposability




$$v - E[\gamma(y) \mid a'] = (1-\delta)(u(a') - E[\gamma(y) \mid a'])$$
$$u(a') - v = \delta(u(a') - E[\gamma(y) \mid a'])$$

# Impact of Increased Precision

- Let $R$ be the $|A| \times |Y|$-matrix, $[R]_{ay} := \rho(y \mid a)$.
- $(Y, \rho')$ is a garbling of $(Y, \rho)$ if there exists a stochastic matrix $Q$ such that
$$R' = RQ.$$
That is, the "experiment" $(Y, \rho')$ is obtained from $(Y, \rho)$ by first drawing $y$ according to $\rho$, and then adding noise.
- If $\mathcal{W}$ can be decomposed on $\mathcal{W}'$ under $\rho'$, then $\mathcal{W}$ can be decomposed on the convex hull of $\mathcal{W}'$ under $\rho$. And so the set of PPE payoffs is weakly increasing as the monitoring becomes more precise.

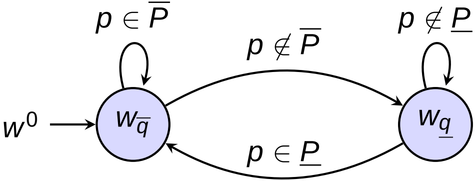
# Bang-Bang

- Suppose $A$ is finite and the signals $y$ are distributed absolutely continuously with respect to Lebesgue measure on a subset of $\mathbb{R}^k$. Every pure strategy eq payoff can be achieved by $(\mathcal{W}, w^0, f, \tau)$ with the bang-bang property:

$$V(w) \in \text{ext}\, \mathcal{E}^p(\delta) \quad \forall w \neq w^0,$$

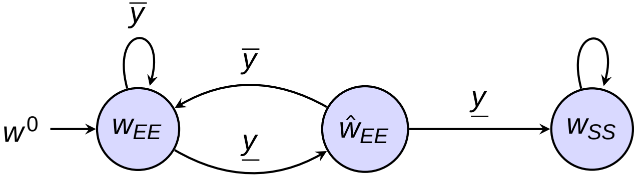
where ext $\mathcal{E}^p(\delta)$ is the set of extreme points of $\mathcal{E}^p(\delta)$.

- (Green-Porter) If $(\mathcal{W}, w^0, f, \tau)$ is strongly symmetric, then ext $\mathcal{E}^p(\delta) = \{\underline{V}, \overline{V}\}$, where $\underline{V} := \min \mathcal{E}^p(\delta)$, $\overline{V} := \max \mathcal{E}^p(\delta)$.



# Prisoners' Dilemma with Noisy Monitoring

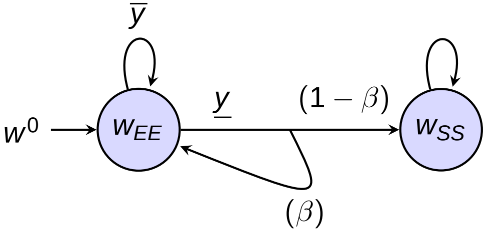
## The value of "forgiveness" I




- This has a higher value than grim trigger, since permanent *SS* is only triggered after two consecutive $\underline{y}$.
- But the limiting value (as $\delta \to 1$) is still zero. As players become more patient, the future becomes more important, and smaller variations in continuation values suffice to enforce *EE*.
- *EE* can be enforced by more forgiving specifications as $\delta \to 1$.

# Prisoners' Dilemma with Noisy Monitoring

## The value of "forgiveness" II




- Public correlating device: $\beta$.
- This is an eq if

$$V = (1-\delta)2 + \delta(p + (1-p)\beta)V$$
$$\geq (1-\delta)3 + \delta(q + (1-q)\beta)V$$

- In the efficient eq (requires $p > q$ and $\delta(3p-2q) > 1$),

$$\beta = \frac{\delta(3p-2q)-1}{\delta(3p-2q-1)} \quad \text{and} \quad V = 2 - \frac{1-p}{p-q} < 2.$$

# Prisoners' Dilemma with Noisy Monitoring

## The value of "forgiveness" III

- Public correlating device is not necessary: Every pure strategy strongly symmetric PPE has payoff no larger than

$$2 - \frac{1-p}{p-q} =: \overline{\gamma}.$$

# Prisoners' Dilemma with Noisy Monitoring

## The value of "forgiveness" III

- Public correlating device is not necessary: Every pure strategy strongly symmetric PPE has payoff no larger than

$$2 - \frac{1-p}{p-q} =: \overline{\gamma}.$$

- Moreover, the upper bound is achieved: For sufficiently large $\delta$, both $[0, \overline{\gamma}]$ and $(0, \overline{\gamma}]$ are self-generating.

# Prisoners' Dilemma with Noisy Monitoring

## The value of "forgiveness" III

- Public correlating device is not necessary: Every pure strategy strongly symmetric PPE has payoff no larger than

$$2 - \frac{1-p}{p-q} =: \overline{\gamma}.$$

- Moreover, the upper bound is achieved: For sufficiently large $\delta$, both $[0, \overline{\gamma}]$ and $(0, \overline{\gamma}]$ are self-generating.
- The use of payoff 0 is Nash reversion.

# Prisoners' Dilemma with Noisy Monitoring

## The value of "forgiveness" III

- Public correlating device is not necessary: Every pure strategy strongly symmetric PPE has payoff no larger than

$$2 - \frac{1-p}{p-q} =: \bar{\gamma}.$$

- Moreover, the upper bound is achieved: For sufficiently large $\delta$, both $[0, \bar{\gamma}]$ and $(0, \bar{\gamma}]$ are self-generating.
- The use of payoff 0 is Nash reversion.
- Forgiving grim trigger: the set $\mathcal{W} = \{0\} \cup [\underline{\gamma}, \bar{\gamma}]$, where

$$\underline{\gamma} := \frac{2(1-\delta)}{1-\delta p},$$

is, for large $\delta$, self-generating with all payoffs $> 0$ decomposed using *EE*.

# Implications

- Providing intertemporal incentives requires imposing punishments on the equilibrium path.
- These punishments may generate inefficiencies, and the greater the noise, the greater the inefficiency.
- How to impose punishments without creating inefficiencies: transfer value rather than destroying it.
- In PD example, impossible to distinguish $ES$ from $SE$.
- Efficiency requires the monitoring be statistically sufficiently informative.
- Other examples reveal the need for asymmetric/ nonstationary behavior in symmetric stationary environments.

# Statistically Informative Monitoring

## Rank Conditions

### Definition

The profile $\alpha$ has individual full rank for player $i$ if the $|A_i| \times |Y|$-matrix $R_i(\alpha_{-i})$, with

$$[R_i(\alpha_{-i})]_{a_i y} := \rho(y \mid a_i \alpha_{-i}),$$

has full row rank.

# Statistically Informative Monitoring

## Rank Conditions

### Definition

The profile $\alpha$ has individual full rank for player $i$ if the $|A_i| \times |Y|$-matrix $R_i(\alpha_{-i})$, with

$$[R_i(\alpha_{-i})]_{a_i y} := \rho(y \mid a_i \alpha_{-i}),$$

has full row rank.

The profile $\alpha$ has pairwise full rank for players $i$ and $j$ if the $(|A_i| + |A_j|) \times |Y|$-matrix

$$R_{ij}(\alpha) := \begin{bmatrix} R_i(\alpha_{-i}) \\ R_j(\alpha_{-j}) \end{bmatrix}$$

has rank $|A_i| + |A_j| - 1$.

# Another Folk Theorem

**The Public Monitoring Folk Theorem (Fudenberg, Levine, and Maskin 1994)**

Suppose the set of feasible and individually rational payoffs has nonempty interior, and that all action profiles satisfy pairwise full rank for all players. Every strictly individually rational and feasible payoff is a perfect public equilibrium payoff, provided players are patient enough.

# Another Folk Theorem

**The Public Monitoring Folk Theorem (Fudenberg, Levine, and Maskin 1994)**

Suppose the set of feasible and individually rational payoffs has nonempty interior, and that all action profiles satisfy pairwise full rank for all players. Every strictly individually rational and feasible payoff is a perfect public equilibrium payoff, provided players are patient enough.

- Pairwise full rank fails for our prisoners' dilemma example (can be satisfied if there are three signals).
- Also fails for Green Porter noisy oligopoly example, since distribution of the market clearing price only depends on total market quantity.
- Folk theorem holds under weaker assumptions.

# Role of Patience

- The monitoring can be arbitrarily noisy, as long as it remains statistically informative.
- But, the noisier the monitoring the more patient the players must be.

# Role of Patience

- The monitoring can be arbitrarily noisy, as long as it remains statistically informative.
- But, the noisier the monitoring the more patient the players must be.
- Suppose time is continuous, and decisions are taken at points $\Delta$, $2\Delta$, $3\Delta$,....
- If $r$ is continuous rate of time discounting, then $\delta = e^{-r\Delta}$.
- As $\Delta \to 0$, $\delta \to 1$.
  - For games of perfect monitoring, high $\delta$ can be interpreted as $\Delta$.
  - But, this is problematic for games of imperfect monitoring: As $\Delta \to 0$, the monitoring becomes increasingly precise over a fixed time interval.

# Repeated Games and Reputations: Private Monitoring

George J. Mailath

University of Pennsylvania
and
Australian National University

CEMMAP Lectures
November 17-18, 2016

The slides and associated bibliography
are on my webpage
http://economics.sas.upenn.edu/∼gmailath

# Games with Private Monitoring

- Intertemporal incentives arise when public histories coordinate continuation play.
- Can intertemporal incentives be provided when the monitoring is private?
- Stigler (1964) suggested that that answer is often NO, and so collusion is not likely to be a problem when monitoring problems are severe.

# The Problem

- Fix a strategy profile $\sigma$. Player *i*'s strategy is sequentially rational if, after all private histories, the continuation strategy is a best reply to the other players' continuation strategies (which depend on their private histories).
- That is, player *i* is best responding to the other players' behavior, given his beliefs over the private histories of the other players.
- While player *i* knows his/her beliefs, we typically do not.
- Most researchers thought this problem was intractable,

# The Problem

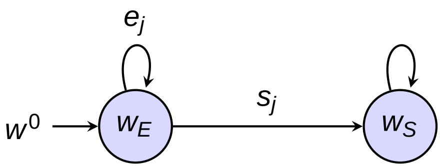
- Fix a strategy profile $\sigma$. Player $i$'s strategy is sequentially rational if, after all private histories, the continuation strategy is a best reply to the other players' continuation strategies (which depend on their private histories).
- That is, player $i$ is best responding to the other players' behavior, given his beliefs over the private histories of the other players.
- While player $i$ knows his/her beliefs, we typically do not.
- Most researchers thought this problem was intractable, until Sekiguchi, in 1997, showed:

There exists an almost efficient eq for the PD with conditionally-independent almost-perfect private monitoring.

# Prisoners' Dilemma

## Conditionally Independent Private Monitoring

| | $E$ | $S$ |
|---|---|---|
| $E$ | $2, 2$ | $-1, 3$ |
| $S$ | $3, -1$ | $0, 0$ |




- Rather than observing the other player's action for sure, player $i$ observes a noisy signal: $\pi_i(y_i = a_j) = 1 - \varepsilon$.
- Grim trigger is not an equilibrium: at the end of the first period, it is not optimal for player $i$ to play $S$ after observing $y_i = s_j$ (since in eq, player $j$ played $E$ and so with high prob, observed $y_j = e_i$).
- Sekiguchi (1997) avoided this by having players randomize (we will see how later).

# Almost Public Monitoring

- How robust are PPE in the game with public monitoring to the introduction of a little private monitoring?
- Perturb the public signal, so that player $i$ observes the conditionally (on $y$) independent signal $y_i \in \{\underline{y}, \bar{y}\}$, with probabilities given by

$$\pi(y_1, y_2 \mid y) = \pi_1(y_1 \mid y)\pi_2(y_2 \mid y),$$

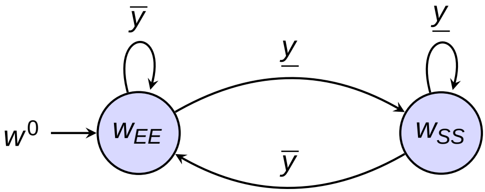
  and

$$\pi_i(y_i \mid y) = \begin{cases} 1 - \varepsilon, & \text{if } y_i = y, \\ \varepsilon, & \text{if } y_i \neq y. \end{cases}$$

- Ex post payoffs are now $u_i^*(a_i, y_i)$.

# Prisoners' Dilemma with Noisy Monitoring

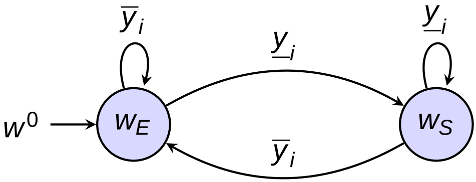
## Bounded Recall-public monitoring




- Suppose $(3p - 2q - r)^{-1} < \delta < (p + 2q - 3r)^{-1}$, so profile is strict PPE in game with public monitoring.
- $V_i(w)$ is $i$'s value from being in public state $w$.

# Prisoners' Dilemma with Noisy Monitoring

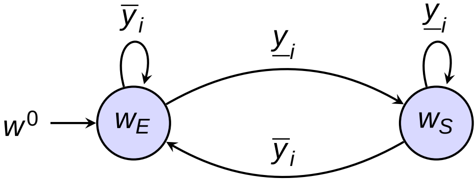
## Bounded Recall-private (almost-public) monitoring




- In period $t$, player $i$'s continuation strategy after private history $h_i^t = (a_i^0, a_i^1, \ldots, a_i^{t-1})$ is completely determined by $i$'s private state $w_i^t \in \mathcal{W}$.
- In period $t$, $j$ sees private history $h_j^t$, and forms belief $\beta_j(h_j^t) \in \mathcal{W}$ over the period $t$ state of player $i$.

# Prisoners' Dilemma with Noisy Monitoring

## Bounded Recall-Best Replies




- For all $y$, $\Pr(y_i \neq y_j \mid y) = 2\varepsilon(1-\varepsilon)$, and so

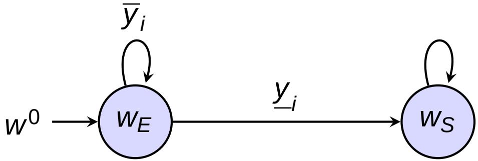
$$\Pr(w_j^t \neq w_i^t(h_i^t) \mid h_i^{t'}) = 2\varepsilon(1-\varepsilon) \quad \forall t' \leq t.$$

- For $\varepsilon$ sufficiently small, incentives from public monitoring carry over to game with almost public monitoring, and profile is an equilibrium.

# Prisoners' Dilemma with Noisy Monitoring

## Grim Trigger

- Suppose $\frac{1}{2p-q} < \delta < 1$, so grim trigger is a strict PPE.
- Strategy in game with private monitoring is




- If $1 > p > q > r > 0$, profile is not a Nash eq (for any $\varepsilon > 0$).
- If $1 > p > r > q > 0$, profile is a Nash eq (but not sequentially rational).

# Prisoners' Dilemma with Noisy Monitoring

Grim Trigger, $1 > p > q > r > 0$

- Consider private history $h_1^t = (E\underline{y}_1, S\bar{y}_1, S\bar{y}_1, \cdots, S\bar{y}_1)$.
- Associated beliefs of 1 about $w_2^t$:

$$\Pr(w_2^0 = w_E) = 1,$$

$$\Pr(w_2^1 = w_S \mid E\underline{y}_1) = \Pr(y_2^1 = \underline{y}_2 \mid E\underline{y}_1, w_2^0 = w_E) \approx 1 - \varepsilon < 1,$$

but

$$\begin{aligned}\Pr(w_2^t = w_S \mid h_1^t) &\\ = \Pr(w_2^t = w_S \mid w_2^{t-1} = w_S) \Pr(w_2^{t-1} = w_S \mid h_1^t) &\\ + \underbrace{\Pr(y_2^t = \underline{y} \mid w_2^{t-1} = w_E, h_1^t)}_{\approx 0} \Pr(w_2^{t-1} = w_E \mid h_1^t),&\end{aligned}$$

and $\Pr(w_2^{t-1} = w_S \mid h_1^t) < \Pr(w_2^{t-1} = w_S \mid h_1^{t-1})$, and so $\Pr(w_2^t = w_S \mid h_1^t) \to \approx 0$, as $t \to \infty$ .

# Prisoners' Dilemma with Noisy Monitoring

Grim Trigger, $1 > p > r > q > 0$

- Consider private history $h_1^t = (E\underline{y}_1, S\bar{y}_1, S\bar{y}_1, \cdots, S\bar{y}_1)$.
- Associated beliefs of 1 about $w_2^t$:

$$\Pr(w_2^0 = w_E) = 1,$$

$$\Pr(w_2^1 = w_S \mid E\underline{y}_1) = \Pr(y_2^1 = \underline{y}_2 \mid E\underline{y}_1, w_2^0 = w_E) \approx 1 - \varepsilon < 1,$$

but

$$\begin{aligned} &\Pr(w_2^t = w_S \mid h_1^t) \\ &\quad = \Pr(w_2^t = w_S \mid w_2^{t-1} = w_S) \Pr(w_2^{t-1} = w_S \mid h_1^t) \\ &\qquad + \underbrace{\Pr(y_2^t = \underline{y} \mid w_2^{t-1} = w_E, h_1^t)}_{\approx 0} \Pr(w_2^{t-1} = w_E \mid h_1^t), \end{aligned}$$

and $\Pr(w_2^{t-1} = w_S \mid h_1^t) > \Pr(w_2^{t-1} = w_S \mid h_1^{t-1})$, and so $\Pr(w_2^t = w_S \mid h_1^t) \approx 1$ for all $t$.

# Prisoners' Dilemma with Noisy Monitoring

Grim Trigger, $1 > p > r > q > 0$

- Consider private history $h_1^t = (E\underline{y}_1, E\bar{y}_1, E\bar{y}_1, \cdots, E\bar{y}_1)$.
- Associated beliefs of 1 about $w_2^t$:

$$\Pr(w_2^0 = w_E) = 1,$$

$$\Pr(w_2^1 = w_S \mid E\underline{y}_1) = \Pr(y_2^1 = \underline{y}_2 \mid E\underline{y}_1, w_2^0 = w_E) \approx 1 - \varepsilon < 1,$$

but

$$\begin{aligned}\Pr(w_2^t = w_S \mid h_1^t) \\ &= \Pr(w_2^t = w_S \mid w_2^{t-1} = w_S) \Pr(w_2^{t-1} = w_S \mid h_1^t) \\ &\quad + \underbrace{\Pr(y_2^t = \underline{y} \mid w_2^{t-1} = w_E, h_1^t)}_{\approx 0} \Pr(w_2^{t-1} = w_E \mid h_1^t),\end{aligned}$$

and $\Pr(w_2^{t-1} = w_S \mid h_1^t) < \Pr(w_2^{t-1} = w_S \mid h_1^{t-1})$, and so $\Pr(w_2^t = w_S \mid h_1^t) \to \approx 0$, as $t \to \infty$ .

# Automaton Representation of Strategies

An automaton is the tuple $(\mathcal{W}_i, w_i^0, f_i, \tau_i)$, where

- $\mathcal{W}_i$ is set of states,
- $w_i^0$ is initial state,
- $f_i : \mathcal{W} \to A_i$ is output function (decision rule), and
- $\tau_i : \mathcal{W}_i \times A_i \times Y_i \to \mathcal{W}_i$ is transition function.

Any automaton $(\mathcal{W}_i, w_i^0, f_i, \tau_i)$ induces a strategy for $i$. Define

$$\tau_i(w_i, h_i^t) := \tau_i(\tau_i(w_i, h_i^{t-1}), a_i^{t-1}, y_i^{t-1}).$$

The induced strategy $s_i$ is given by $s_i(\varnothing) = f_i(w_i^0)$ and

$$s_i(h_i^t) = f_i(\tau_i(w_i^0, h_i^t)), \quad \forall h_i^t.$$

Every strategy can be represented by an automaton.

# Almost Public Monitoring Games

- Fix a game with imperfect full support public monitoring, so that for all $y \in Y$ and $a \in A$, $\rho(y \mid a) > 0$.
- Rather than observing the public signal directly, each player $i$ observes a private signal $y_i \in Y$.
- The game with private monitoring is $\varepsilon$-close to the game with public monitoring if the joint distribution $\pi$ on the private signal profile $(y_1, \ldots, y_n)$ satisfies

  $$|\pi((y, y, \ldots, y) \mid a) - \rho(y \mid a)| < \varepsilon.$$

  Such a game has almost public monitoring.
- Any automaton in the game with public monitoring describes a strategy profile in all $\varepsilon$-close almost public monitoring games.

# Almost Pubic Monitoring

## Rich Private Monitoring

- Fix a game with imperfect full support public monitoring, so that for all $y \in Y$ and $a \in A$, $\rho(y \mid a) > 0$.
- Each player $i$ observes a private signal $z_i \in Z_i$, with $(z_1, \ldots, z_n)$ distributed according to the joint dsn $\pi$.
- The game with rich private monitoring is $\varepsilon$-close to the game with public monitoring if there are mappings $\xi_i : Z_i \to Y$ such that

$$\left| \sum_{\xi_1(z_1)=y,\ldots,\xi_n(z_n)=y} \pi((z_1, \ldots, z_n) \mid a) - \rho(y \mid a) \right| < \varepsilon.$$

  Such a game has almost public monitoring.
- Any automaton in the game with public monitoring describes a strategy profile in all $\varepsilon$-close almost public monitoring games with rich private monitoring.

# Behavioral Robustness

### Definition

An eq of a game with public monitoring is behaviorally robust if the same automaton is an eq in all $\varepsilon$-close games to the game with public monitoring for $\varepsilon$ sufficiently small.

### Definition

A public automaton $(\mathcal{W}, w^0, f, \tau)$ has bounded recall if there exists $L$ such that after any history of length at least $L$, continuation play only depends on the last $L$ periods of the public history (i.e., $\tau(w, h^L) = \tau(w', h^L)$ for all $w, w' \in \mathcal{W}$).

# Behavioral Robustness

An eq is behaviorally robust if the same profile is an eq in near-by games.

A public profile has bounded recall if there exists $L$ such that after any history of length at least $L$, continuation play only depends on the last $L$ periods of the public history.

**Theorem (Mailath and Morris, 2002)**

A strict PPE with bounded recall is behaviorally robust to private monitoring that is almost public.

# Behavioral Robustness

An eq is behaviorally robust if the same profile is an eq in near-by games.

A public profile has bounded recall if there exists $L$ such that after any history of length at least $L$, continuation play only depends on the last $L$ periods of the public history.

**Theorem (Mailath and Morris, 2002)**

A strict PPE with bounded recall is behaviorally robust to private monitoring that is almost public.

**“Theorem” (Mailath and Morris, 2006)**

If the private monitoring is sufficiently rich, a strict PPE is behaviorally robust to private monitoring that is almost public if and only if it has bounded recall.

# Bounded Recall

It is tempting to think that bounded recall provides an attractive restriction on behavior. But:

**Folk Theorem II (Hörner and Olszewski, 2009)**

The public monitoring folk theorem holds using bounded recall strategies. The folk theorem also holds using bounded recall strategies for games with almost-public monitoring.

- This private monitoring folk theorem is not behaviorally robust.

# Bounded Recall

It is tempting to think that bounded recall provides an attractive restriction on behavior. But:

**Folk Theorem II (Hörner and Olszewski, 2009)**

The public monitoring folk theorem holds using bounded recall strategies. The folk theorem also holds using bounded recall strategies for games with almost-public monitoring.

- This private monitoring folk theorem is not behaviorally robust.

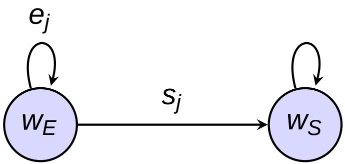
**Folk Theorem III (Mailath and Olszewski, 2011)**

The perfect monitoring folk theorem holds using bounded recall strategies with uniformly strict incentives. Moreover, the resulting equilibrium is behaviorally robust to almost-perfect almost-public monitoring.

# Prisoners' Dilemma

## Conditionally Independent Private Monitoring

| | $E$ | $S$ |
|---|---|---|
| $E$ | $2, 2$ | $-1, 3$ |
| $S$ | $3, -1$ | $0, 0$ |




Player $i$ observes a noisy signal: $\pi_i(y_i = a_j) = 1 - \varepsilon$.

**Theorem (Sekiguchi, 1997)**

*For all $\psi > 0$, there exists $\eta'' > \eta' > 0$ such that for all $\delta \in (1/3 + \eta', 1/3 + \eta'')$, there is a Nash equilibrium in which each player randomizing over the initial state, with the probability on $w_E$ exceeding $1 - \psi$.*

# Proof and extend to all high $\delta$

## Proof of theorem

Optimality of grim trigger after different histories:

- *Es*: updating given original randomization $\implies$ S optimal.

# Proof and extend to all high $\delta$

## Proof of theorem

Optimality of grim trigger after different histories:

- *Es*: updating given original randomization $\implies$ S optimal.
- $Ee, Ee, \ldots, Ee$: perpetual *e* reassures *i* that *j* is still in $w_E$.

# Proof and extend to all high $\delta$

## Proof of theorem

Optimality of grim trigger after different histories:

- $Es$: updating given original randomization $\implies$ S optimal.
- $Ee, Ee, \ldots, Ee$: perpetual $e$ reassures $i$ that $j$ is still in $w_E$.
- $Ee, Ee, \ldots, Ee, Es$. Most likely events: either $j$ is still in $w_E$ and $s$ is a mistake, or $j$ received an erroneous signal in the previous period. Odds slightly favor $j$ receiving the erronous signal, and because $\delta$ low, $S$ is optimal.

# Proof and extend to all high $\delta$

**Proof of theorem**

Optimality of grim trigger after different histories:

- $Es$: updating given original randomization $\implies$ S optimal.
- $Ee, Ee, \ldots, Ee$: perpetual $e$ reassures $i$ that $j$ is still in $w_E$.
- $Ee, Ee, \ldots, Ee, Es$. Most likely events: either $j$ is still in $w_E$ and $s$ is a mistake, or $j$ received an erroneous signal in the previous period. Odds slightly favor $j$ receiving the erronous signal, and because $\delta$ low, $S$ is optimal.
- $Ee, Ee, \ldots, Ee, Es, Se, \ldots, Se$. This period's $S$ will trigger $j$'s switch to $w_S$, if not there already.

# Proof and extend to all high $\delta$

## Proof of theorem

Optimality of grim trigger after different histories:

- $Es$: updating given original randomization $\implies$ S optimal.
- $Ee, Ee, \ldots, Ee$: perpetual $e$ reassures $i$ that $j$ is still in $w_E$.
- $Ee, Ee, \ldots, Ee, Es$. Most likely events: either $j$ is still in $w_E$ and $s$ is a mistake, or $j$ received an erroneous signal in the previous period. Odds slightly favor $j$ receiving the erronous signal, and because $\delta$ low, $S$ is optimal.
- $Ee, Ee, \ldots, Ee, Es, Se, \ldots, Se$. This period's $S$ will trigger $j$'s switch to $w_S$, if not there already.

To extend to all high $\delta$, lower effective discount factor by dividing games into $N$ interleaved games.

# Belief-Free Equilibria

Another approach is to specify behavior in such a way that the beliefs are irrelevant.
Suppose $n = 2$.

## Definition

The profile $((\mathcal{W}_1, w_1^0, f_1, \tau_1), (\mathcal{W}_2, w_2^0, f_2, \tau_2))$ is a belief-free eq if for all $(w_1, w_2) \in \mathcal{W}_1 \times \mathcal{W}_1$, $(\mathcal{W}_i, w_i, f_i, \tau_i)$ is a best reply to $(\mathcal{W}_j, w_j, f_j, \tau_j)$, all $i \neq j$.

This approach is due to Piccione (2002), with a refinement by Ely and Valimaki (2002). Belief-free eq are characterized by Ely, Hörner, and Olszewski (2005).

# Illustration of Belief Free Eq

## The product-choice game

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

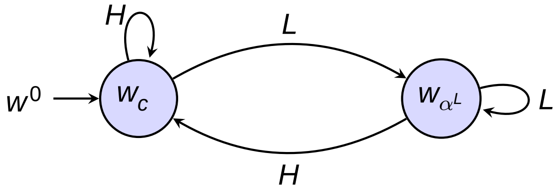
- Row player is a firm choosing *H*igh or *L*ow quality.
- Column player is a short-lived customer choosing the *c*ustomized or *s*tandard product.
- In the game with perfect monitoring, grim trigger (play $Hc$ till 1 plays $L$, then revert to perpetual $Ls$) is an eq if $\delta \geq \frac{1}{2}$.

The belief-free eq that achieves a payoff of 2 for the row player:

- Row player always plays $\frac{1}{2} \circ H + \frac{1}{2} \circ L$. (Trivial automaton)
- Column player's strategy has one period memory. Play $c$ for sure after $H$ in the previous period, and play

$$\alpha^L := \left(1 - \frac{1}{2\delta}\right) \circ c + \frac{1}{2\delta} \circ s$$

after $L$ in the previous period. Player 2's automaton:



- Let $V_1(w; a_1)$ denote player 1's payoff when 2 is in state $w$, and 1 plays $a_1$. Then (where $\alpha = 1 - 1/(2\delta)$),

$$\begin{aligned} V_1(w_c; H) &= (1-\delta)2 + \delta V_1(w_c) \\ = V_1(w_c; L) &= (1-\delta)3 + \delta V_1(w_{\alpha^L}), \\ V_1(w_{\alpha^L}; a_1 = H) &= (1-\delta)2\alpha + \delta V_1(w_c) \\ = V_1(w_{\alpha^L}; a_1 = L) &= (1-\delta)(2\alpha + 1) + \delta V_1(w_{\alpha^L}). \end{aligned}$$

- Then, $V_1(w_c) - V_1(w_{\alpha^L}) = (1-\delta)/\delta$.
- Which is true when $\alpha = 1 - 1/(2\delta)$.

# Belief-Free Eq in the Prisoners' Dilemma

Ely and Valimaki (2002)

| | $E$ | $S$ |
|---|---|---|
| $E$ | $2, 2$ | $-1, 3$ |
| $S$ | $3, -1$ | $0, 0$ |

- Perfect monitoring.
- Player $i$'s automaton, $(\mathcal{W}_i, w_i, f_i, \tau_i)$:

$$\mathcal{W} = \{w_i^E, w_i^S\},$$

$$f_i(w_i^a) = \begin{cases} 1, & a = E, \\ \alpha \circ E + (1-\alpha) \circ S, & a = S, \end{cases}$$

$$\tau_i(w_i, a_i a_j) = w_i^{a_j},$$

  where $\alpha := 1 - 1/(3\delta)$.
- Both $(\mathcal{W}_1, w_1^E, f_1, \tau_1)$ and $(\mathcal{W}_1, w_1^S, f_1, \tau_1)$ are best replies to both $(\mathcal{W}_2, w_2^E, f_2, \tau_2)$ and $(\mathcal{W}_2, w_2^S, f_2, \tau_2)$.

# Belief-Free in the Prisoners' Dilemma-Proof

- Let $V_1(aa')$ denote player 1's payoff when 1 is in state $w_1^a$ and 2 is in state $w_2^{a'}$. Then

$$
\begin{aligned}
V_1(EE) &= (1-\delta)2 + \delta V_1(EE), \\
V_1(ES) &= (1-\delta)(3\alpha - 1) \\
&\qquad + \delta[\alpha V_1(EE) + (1-\alpha)V_1(SE)], \\
V_1(SE; a_1 = E) &= (1-\delta)2 + \delta V_1(EE) \\
= V_1(SE; a_1 = S) &= (1-\delta)3 + \delta V_1(ES), \\
V_1(SS : a_1 = E) &= (1-\delta)(-1) \\
&\qquad + \ \delta[\alpha V_1(EE) + (1-\alpha)V_1(SE)] \\
= V_1(SS : a_1 = S) &= \delta[\alpha V_1(ES) + (1-\alpha)V_1(SS)].
\end{aligned}
$$

- Then, $V_1(EE) - V_1(ES) = V_1(SE) - V_1(SS) = (1-\delta)/\delta$.
- Which is true when $\alpha = 1 - 1/(3\delta)$.

# Belief-Free in the Prisoners' Dilemma

## Private Monitoring

- Suppose we have conditionally independent private monitoring.
- For $\varepsilon$ small, there is a value of $\alpha$ satisfying the analogue of the indifference conditions for perfect monitoring (the system of equations is well-behaved, and so you can apply the implicit function theorem).
- These kinds of strategies can be used to construct equilibria with payoffs in the square $(0, 2) \times (0, 2)$ for sufficiently patient players.

- Histories are not being used to coordinate play! There is no common understanding of continuation play.
- This is to be contrasted with strict PPE.
- Rather, lump sum taxes are being imposed after “deviant” behavior is “suggested.”
- This is essentially what do in the repeated prisoners’ dilemma.
- Folk theorems for games with private monitoring have been proved using belief free constructions.
- These equilibria seem crazy, yet Kandori and Obayashi (2014) report suggestive evidence that in some community unions in Japan, the behavior accords with such an equilibrium.

# Imperfect Monitoring

- This works for public and private monitoring.
- No hope for behavioral robustness.

**"Theorem" (Hörner and Olszewski, 2006)**

The folk theorem holds for games with private almost-perfect monitoring.

- Result uses belief-free ideas in a central way, but the equilibria constructed are not belief free.

# Purifiability

- Belief-free equilibria typically have the property that players randomize the same way after different histories (and so with different beliefs over the private states of the other player(s)).
- Harsanyi (1973) purification (robustness to private payoff shocks) is perhaps the best rationale for randomizing behavior in finite normal form games.
- Can we purify belief-free equilibria (Bhaskar, Mailath, and Morris, 2008)?
  - The one period memory belief free equilibria of Ely and Valimaki (2002), as exemplified above, is not purifiable using one period memory strategies.
  - They are purifiable using unbounded memory strategies.
  - Open question: can they be purified using bounded memory strategies? (It turns out that for sequential games, only Markov equilibria can be purified using bounded memory strategies, Bhaskar, Mailath, and Morris 2013).

# What about noisy monitoring?

Current best result is Sugaya (2013):

**"Theorem"**

The folk theorem generically holds for the repeated two-player prisoners' dilemma with private monitoring if the support of each player's signal distribution is sufficiently large. Neither cheap talk communication nor public randomization is necessary, and the monitoring can be very noisy.

# Ex Post Equilibria

- The belief-free idea is very powerful.
- Suppose there is an unknown state determining payoffs and monitoring.

| $\omega_E$ | $E$ | $S$ |
|---|---|---|
| $E$ | $1, 1$ | $-1, 2$ |
| $S$ | $2, -1$ | $0, 0$ |

| $\omega_S$ | $E$ | $S$ |
|---|---|---|
| $E$ | $0, 0$ | $2, -1$ |
| $S$ | $-1, 2$ | $1, 1$ |

- Let $\Gamma(\delta; \omega)$ denote the complete-information repeated game when state $\omega$ is common knowledge. Monitoring may be perfect or imperfect public.

# Perfect Public Ex Post Equilibria

$\Gamma(\delta;\omega)$ is complete-information repeated game at $\omega$.

**Definition**

The profile of public strategies $\sigma^*$ is a perfect public ex post eq if $\sigma^*|_{h^t}$ is a Nash eq of $\Gamma(\delta;\omega)$ for all histories $h^t \in H$, where $\sigma^*|_{h^t}$ is the continuation public profile induced by $h^t$.

- These equilibria can be strict; histories do coordinate play.
- But the eq are belief free.

# Perfect Public Ex Post Equilibria

$\Gamma(\delta;\omega)$ is complete-information repeated game at $\omega$.

**Definition**

The profile of public strategies $\sigma^*$ is a perfect public ex post eq if $\sigma^*|_{h^t}$ is a Nash eq of $\Gamma(\delta;\omega)$ for all histories $h^t \in H$, where $\sigma^*|_{h^t}$ is the continuation public profile induced by $h^t$.

- These equilibria can be strict; histories do coordinate play.
- But the eq are belief free.

**"Theorem" (Fudenberg and Yamamoto 2010)**

Suppose the signals are statistically informative (about actions and states). The folk theorem holds state-by-state.

These ideas also can be used in some classes of reputation games (Hörner and Lovo, 2009) and in games with private monitoring (Yamamoto, 2014).

# Conclusion

- The current theory of repeated games shows that the long relationships can discourage opportunistic behavior, it does not show that long run relationships will discourage opportunistic behavior.
- Incentives can be provided when histories coordinate continuation play.
- Punishments must be credible, and this can limit their scope.
- Some form of monitoring is needed to punish deviators.
- This monitoring can occur through communication networks.
- Intertemporal incentives can also be provided in situations when there is no common understanding of histories, and so of continuation play.

# What is left to understand

- Which behaviors in long-run relationships are plausible?
- Why are formal institutions important?
- Why do we need formal institutions to protect property rights, for example?
- Communication is not often modelled explicitly, and it should be. Communication make things significantly easier (see Compte, 1998, and Kandori and Matsushima, 1998).
- Too much focus on patient players ($\delta$ close to 1).

# Repeated Games and Reputations: Reputations I

George J. Mailath

University of Pennsylvania
and
Australian National University

CEMMAP Lectures
November 17-18, 2016

The slides and associated bibliography
are on my webpage
http://economics.sas.upenn.edu/∼gmailath

# Introduction

- Repeated games have many equilibria. At the same time, certain plausible outcomes are not consistent with equilibrium. Illustrate with product-choice game.
- Reputation effects: the impact on the set of equilibria (typically of a repeated game) of perturbing the game by introducing incomplete information of a particular kind.
- Reputation effects bound eq payoffs in a natural way. First illustrate again using the product choice game, and then give a complete proof in the canonical model of Fudenberg and Levine (1989, 1992), using the tool of relative entropy introduced by Gossner (2011), and
- outline the temporary reputation results of Cripps, Mailath, and Samuelson (2004, 2007).

# Introduction

## The product choice game

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

- Row player is a firm, choosing between high ($H$) and low ($L$) effort.
- Column player is a customer, choosing between a customized ($c$) and standard ($s$) product.
- Game has a unique Nash equilibrium: $Ls$.

# Perfect Monitoring

Suppose firm is long-lived, playing the product-choice game with a sequence of short-lived customers.
Suppose moreover that

- monitoring is perfect (everyone sees all past actions) and
- the firm has unbounded lifespan, with a discount factor $\delta$.

Then

- for $\delta \geq \frac{1}{2}$, there is a subgame perfect eq in which *Hc* is played in every period (any deviation results in *Ls* forever).
- for $\delta \geq \frac{2}{3}$, every payoff in $[1, 2]$ is the payoff of some pure strategy subgame perfect eq.

# Perfect Monitoring

Suppose firm is long-lived, playing the product-choice game with a sequence of short-lived customers.
Suppose moreover that

- monitoring is perfect (everyone sees all past actions) and
- the firm has unbounded lifespan, with a discount factor $\delta$.

Then

- for $\delta \geq \frac{1}{2}$, there is a subgame perfect eq in which *Hc* is played in every period (any deviation results in *Ls* forever).
- for $\delta \geq \frac{2}{3}$, every payoff in $[1, 2]$ is the payoff of some pure strategy subgame perfect eq.

But for all $\delta$, the profile in which history is ignored and *Ls* is played in every period is an eq.

# Imperfect Monitoring

Suppose now that the actions of the firm are imperfectly observed. There is a signal $y \in \{\bar{y}, \underline{y}, \}$ (good experience, bad experience) with distribution

$$\rho(\bar{y} \mid a_1) = \begin{cases} p, & \text{if } a_1 = H, \\ q, & \text{if } a_1 = L, \end{cases}$$

where $0 < q < p < 1$.

- If $2p - q \leq 1$, the only pure strategy PPE is perpetual $Ls$ (and as under perfect monitoring, this is always an eq).
- The maximum payoff the firm can achieve in any PPE is

  $$2 - \frac{(1-p)}{(p-q)} < 2.$$

  (Achieving this bound requires $\delta$ close to 1.)
- Payoffs are bounded away from payoff from perpetual $Hc$.

# The issue

Repeated games have too many equilibria and not enough:

- In the finitely horizon product choice, the unique Nash eq is *Ls* in every period, irrespective of the length of the horizon.
- In the finitely repeated prisoners' dilemma, the unique Nash eq is always defect, irrespective of the number of repetitions.
- In the chain store paradox, the chain store cannot deter entry no matter how many entrants it is facing.

It seems counter-intuitive that observing a sufficiently long history of $H$'s (or sufficiently high fraction of $\bar{y}$'s) in our example would not convince customers that the firm will play $H$.

# Incomplete Information

Suppose the customers are not completely certain of all the characteristics of the firm. That is, the game has incomplete information, with the firm's characteristics (type) being private information to the firm.

Suppose that the customers assign some (small) chance to the firm being a behavioral (commitment) type $\xi(H)$ who always plays $H$.

Then, if the normal type firm is sufficiently patient, its payoff is close to 2.

# A simple reputation result

## A preliminary lemma

### Lemma

*Suppose prob assigned to* $\xi(H)$, $\mu(\xi(H)) =: \mu_0$, *is strictly positive. Fix a Nash equilibrium. Let* $h^t$ *be a positive probability period-*$t$ *history in which* $H$ *is always played. The number of periods in* $h^t$ *in which a customer plays* $s$ *is no larger than*

$$k^* := -\frac{\log \mu_0}{\log 2}.$$

# A simple reputation result

## A preliminary lemma

**Lemma**

*Suppose prob assigned to $\xi(H)$, $\mu(\xi(H)) =: \mu_0$, is strictly positive. Fix a Nash equilibrium. Let $h^t$ be a positive probability period-t history in which $H$ is always played. The number of periods in $h^t$ in which a customer plays $s$ is no larger than*

$$k^* := -\frac{\log \mu_0}{\log 2}.$$

$q_\tau$ is 2's prob that firm plays $H$ in period $\tau$ conditional on $h^\tau$.
In eq, if customer $\tau$ does choose $s$, then

$$q_\tau \leq \tfrac{1}{2}.$$

So, would like an upper bound on

$$k(t) := \#\{\tau \leq t : q_\tau \leq \tfrac{1}{2}\}.$$

Let $\mu_\tau := \Pr\{\xi(H)|h^\tau\}$ be the posterior assigned to $\xi(H)$ after $h^\tau$, and since $h^\tau$ is an initial segment of $h^t$,

$$\begin{aligned}\mu_{\tau+1} = \Pr\{\xi(H)|h^\tau, H\} &= \frac{\Pr\{\xi(H), H|h^\tau\}}{\Pr\{H|h^\tau\}} \\ &= \frac{\Pr\{H|\xi(H), h^\tau\}\Pr\{\xi(H)|h^\tau\}}{\Pr\{H|h^\tau\}} \\ &= \frac{\mu_\tau}{q_\tau} \implies \mu_\tau = q_\tau \mu_{\tau+1}.\end{aligned}$$

Let $\mu_\tau := \Pr\{\xi(H)|h^\tau\}$ be the posterior assigned to $\xi(H)$ after $h^\tau$, and since $h^\tau$ is an initial segment of $h^t$,

$$\begin{aligned}\mu_{\tau+1} = \Pr\{\xi(H)|h^\tau, H\} &= \frac{\Pr\{\xi(H), H|h^\tau\}}{\Pr\{H|h^\tau\}} \\ &= \frac{\Pr\{H|\xi(H), h^\tau\}\Pr\{\xi(H)|h^\tau\}}{\Pr\{H|h^\tau\}} \\ &= \frac{\mu_\tau}{q_\tau} \implies \mu_\tau = q_\tau \mu_{\tau+1}.\end{aligned}$$

Then,

$$\mu_0 = q_0\mu_1 = q_0 q_1 \mu_2 = \mu_t \prod_{\tau=0}^{t-1} q_\tau \le \mu_t \prod\nolimits_{\{\tau : q_\tau \le \frac{1}{2}\}} q_\tau \le \left(\tfrac{1}{2}\right)^{k(t)}.$$

Taking logs, $\log \mu_0 \le k(t) \log \frac{1}{2}$, and so

$$k(t) \le -\frac{\log \mu_0}{\log 2}.$$

# The Theorem

**Theorem (Fudenberg and Levine 1989)**

*Suppose $\xi(H)$ receives positive prior probability $\mu_0 > 0$. In any Nash equilibrium, the normal type's expected payoff is at least $2\delta^{k^*}$. Thus, for all $\varepsilon > 0$, there exists $\bar{\delta}$ such that for all $\delta \in (\bar{\delta}, 1)$, the normal type's payoff in any Nash equilibrium is at least $2 - \varepsilon$.*

Normal type can always playing $H$.
Applying Lemma yields a lower bound on payoffs of

$$\sum_{\tau=0}^{k^*-1} (1-\delta)\delta^\tau 0 + \sum_{\tau=k^*}^{\infty} (1-\delta)\delta^\tau 2 = 2\delta^{k^*}.$$

This can be made arbitrarily close to 2 by choosing $\delta$ close to 1.

# Comments

- This result made few assumptions on the nature of the incomplete information. In particular, the type space $\Xi$ can be infinite (even uncountable), as long as there is a grain of truth on the commitment type ($\mu(\xi(H)) > 0$).
- The result also holds for finite horizons. If firm payoffs are the average of the flow (static) payoffs, then average payoffs are close to 2 for sufficiently long horizons.
- Perfect monitoring of the behavioral type's action is critical.
- Above argument cannot be extended to either imperfect monitoring or mixed behavior types (and yet the intuition is compelling).

  A new argument is needed.

# The Canonical Reputation Model

## The Complete Information 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 max's her (expected) stage game payoff $u_2$.

- Player 1's ex post stage game payoff is $u_1^*(a_1, a_2, y)$, and 1's ex ante payoff is

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

- Player 1 max's the expected value of

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

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

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

- A strategy for player 2:

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

# 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 max's her (expected) 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 $\mu$ on $\Xi$.
- Type space is partitioned into two sets, $\Xi = \Xi_1 \cup \Xi_2$, where
  - $\Xi_1$ is the set of payoff types and
  - $\Xi_2$ 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

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

- Each type $\xi \in \Xi_1$ of player 1 max's the expected value of

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

- 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 there exists unique $\hat{\alpha}_1 \in \Delta(A_1)$ 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 $(\sigma_1, \sigma_2)$ with prior $\mu$ induces the unconditional distribution $\mathbf{P} \in \Delta(\Omega)$.
- For a fixed simple type $\hat{\xi} = \xi(\hat{\alpha}_1)$, the probability measure on $\Omega$ conditioning on $\hat{\xi}$ (and so induced by $\hat{\alpha}_1$ in every period and $\sigma_2$), is denoted $\widehat{\mathbf{P}} \in \Delta(\Omega)$.
- Denoting by $\widetilde{\mathbf{P}}$ the measure induced by $(\sigma_1, \sigma_2)$ and conditioning on $\xi \neq \hat{\xi}$, we have

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

- Given a strategy profile $\sigma$, $U_1(\sigma, \xi)$ denotes the type-$\xi$ 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 $(\sigma_1', \sigma_2')$ is a Nash equilibrium of the game $\Gamma(\mu, \delta)$ if, for all $\xi \in \Xi_1$, $\sigma_1'$ 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 $(\sigma_1', \sigma_2')$ and $\mu$ (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 $(\sigma_1', \sigma_2')$ is a Nash equilibrium of the game $\Gamma(\mu, \delta)$ if, for all $\xi \in \Xi_1$, $\sigma_1'$ 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 $(\sigma_1', \sigma_2')$ and $\mu$ (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 $\sigma'$ and $\Xi$, 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*

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

# 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.**

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

□

# 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.**

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

□

# 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.**

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

□

# 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.

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

□

# 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

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

# 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 $\varepsilon$-entropy confirming best response to $\alpha_1 \in \Delta(A_1)$ if there exists $\alpha_1' \in \Delta(A_1)$ such that

1. $\alpha_2$ is a best response to $\alpha_1'$; and
2. $d(\rho(\cdot|\alpha_1)\|\rho(\cdot|\alpha_1')) \leq \varepsilon$.

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

# From prediction bounds to payoff bounds

## Definition

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

1. $\alpha_2$ is a best response to $\alpha_1'$; and
2. $d(\rho(\cdot|\alpha_1)\|\rho(\cdot|\alpha_1')) \leq \varepsilon$.

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

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 $\varepsilon$-entropy confirming best response to $\alpha_1 \in \Delta(A_1)$ if there exists $\alpha_1' \in \Delta(A_1)$ such that

1. $\alpha_2$ is a best response to $\alpha_1'$; and
2. $d(\rho(\cdot|\alpha_1) \| \rho(\cdot|\alpha_1')) \leq \varepsilon$.

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

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

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

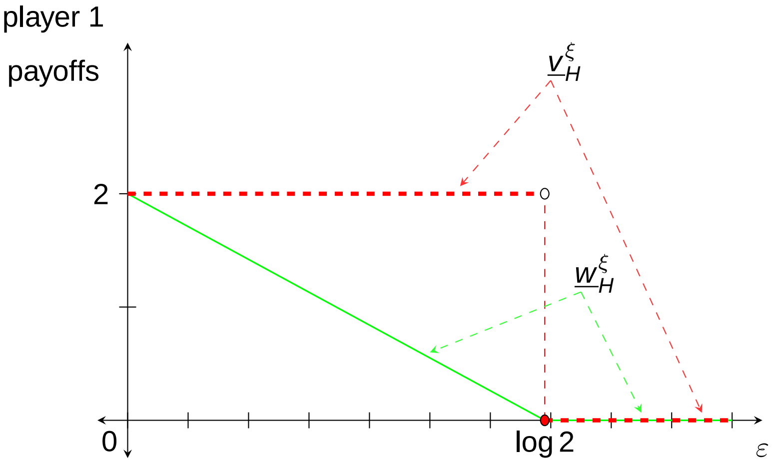
and denote by $\underline{w}_{\alpha_1}^{\xi}$ the largest convex function below $\underline{v}_{\alpha_1}^{\xi}$.

# The product-choice game I

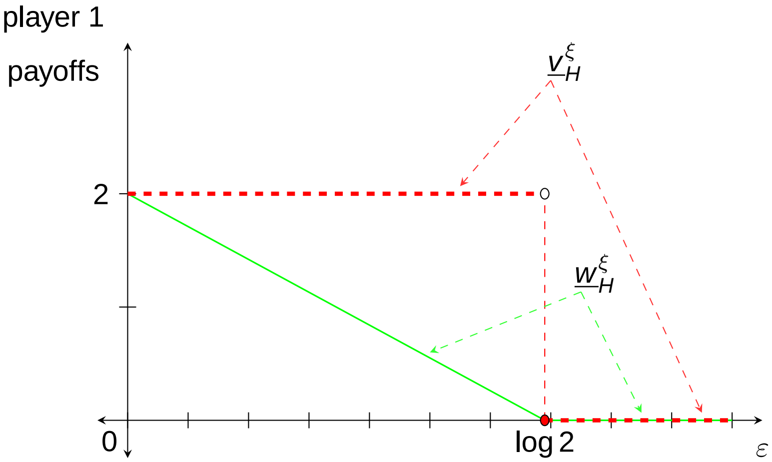
| | $c$ | $s$ |
|---|---|---|
| $H$ | $2, 3$ | $0, 2$ |
| $L$ | $3, 0$ | $1, 1$ |

- Suppose $\hat{\alpha}_1 = 1 \circ H$.
- $c$ is unique BR to $\alpha_1$ if $\alpha_1(H) > \frac{1}{2}$.
- $s$ is also a BR to $\alpha_1$ 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



# A picture is worth a thousand words



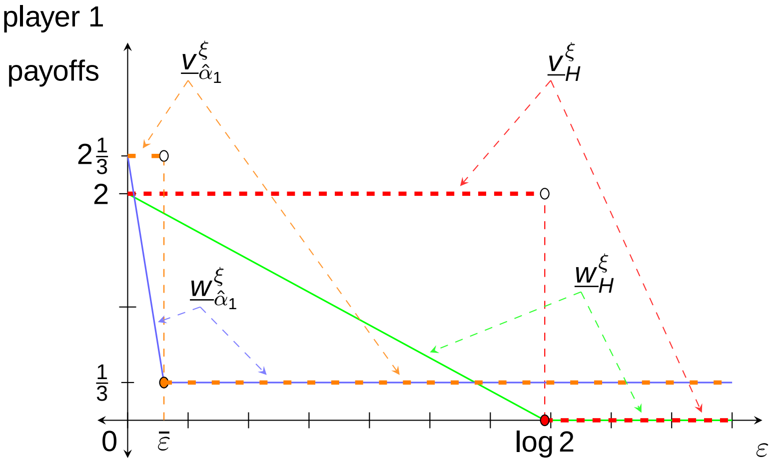

Diff between this and the earlier simple rep. bound is $o(1-\delta)$.

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

# 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* $\xi$*'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 $\mu$.

# The proof

Since in any Nash equilibrium $(\sigma_1', \sigma_2')$, each payoff type $\xi$ has the option of playing $\hat{\alpha}_1$ in every period, we have

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

# The proof

Since in any Nash equilibrium $(\sigma_1', \sigma_2')$, each payoff type $\xi$ has the option of playing $\hat{\alpha}_1$ in every period, we have

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

# The proof

Since in any Nash equilibrium $(\sigma_1', \sigma_2')$, each payoff type $\xi$ has the option of playing $\hat{\alpha}_1$ in every period, we have

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

# The proof

Since in any Nash equilibrium $(\sigma_1', \sigma_2')$, each payoff type $\xi$ has the option of playing $\hat{\alpha}_1$ in every period, we have

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

# The proof

Since in any Nash equilibrium $(\sigma_1', \sigma_2')$, each payoff type $\xi$ 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* $\xi$*'s payoff in any Nash equilibrium of the game* $\Gamma(\mu, \delta)$ *is greater than or equal to*

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

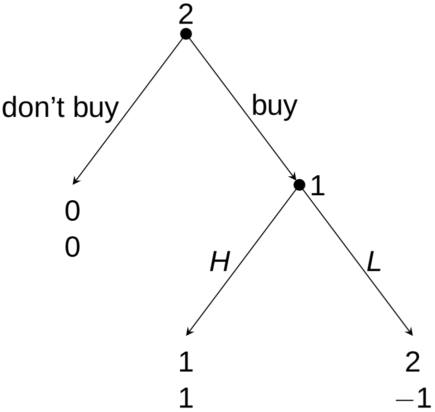
# When does $B_0^d(\alpha_1) = BR(\alpha_1)$?

- 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 subset of $\Delta(A_1)$.

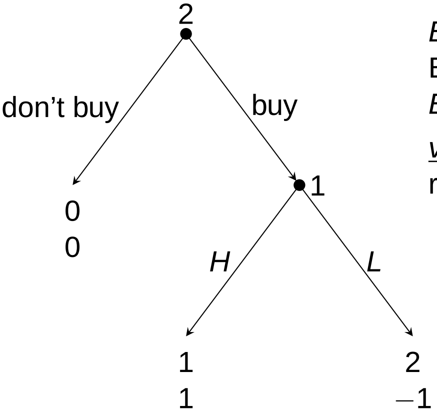
# How general is the result?

- 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.

# The Purchase Game



# The Purchase Game




$BR(H) = \{b\}$.
But $\rho(\cdot|Hd) = \rho(\cdot|Ld)$ and so $B_0^d(Hd) = \{d, b\}$, implying $\underline{v}_H^{\xi(H)}(0) = 0$, and no useful reputation bound.

# Repeated Games and Reputations: Reputations II

George J. Mailath

University of Pennsylvania
and
Australian National University

CEMMAP Lectures
November 17-18, 2016

The slides and associated bibliography
are on my webpage
http://economics.sas.upenn.edu/∼gmailath

# Impermanent Reputations under Imperfect Monitoring

- Imperfect monitoring of long-lived players is not an impediment for reputation effects.
- But it does have implications for its permanance:

  Reputation effects are necessarily temporary in the presence of imperfect monitoring.

  (Under perfect monitoring, permanent reputation effects are trivially possible.)

# Imperfect Monitoring

Suppose only two types, the normal type $\xi_0$ and the simple action type $\hat{\xi} := \xi(\hat{\alpha}_1)$.
Allow signal dsn to depend on $a_1$ and $a_2$.
Maintain assumption that player 1 observes past $a_2$.

**Assumption: Full support**

$\rho(y|\hat{\alpha}_1, a_2) > 0$ for all $y \in Y$ and $a_2 \in A_2$.

**Assumption: Identifiability**

$[\rho(\cdot|\cdot, \alpha_2)]_{y,a_1}$ has full column rank.

Identifiability implies $B_0^d(\alpha_1) = BR(\alpha_1)$.

# Disappearing Reputations

Given a strategy profile $(\sigma_1, \sigma_2)$ of the incomplete information game, the short-lived player's belief in period $t$ that player 1 is type $\hat{\xi}$ is

$$\mu^t(h_2^t) := \mathbf{P}(\hat{\xi} | h_2^t),$$

and so $\mu^0$ is the period 0, or prior, probability assigned to $\hat{\xi}$.

**Proposition (Cripps, Mailath, Samuelson 2004)**

*Suppose player 2 has a unique best response $\hat{a}_2$ to $\hat{\alpha}_1$ and $(\hat{\alpha}_1, \hat{a}_2)$ is not a Nash equilibrium of the stage game. If $(\sigma_1, \sigma_2)$ is a Nash equilibrium of the game $\Gamma(\mu, \delta)$, then*

$$\mu^t \to 0, \qquad \widetilde{\mathbf{P}}\text{-a.s.}$$

# Intuition

Bayes' rule determines $\mu^t$ after all histories (of 1's actions).

At any Nash eq, $\mu^t$ is a bounded martingale and so

$$\exists \mu^\infty : \quad \mu^t \to \mu^\infty \quad \textbf{P}\text{-a.s. (and hence } \widetilde{\textbf{P}}\text{- and } \widehat{\textbf{P}}\text{-a.s.).}$$

# Intuition

Bayes' rule determines $\mu^t$ after all histories (of 1's actions).
At any Nash eq, $\mu^t$ is a bounded martingale and so

$$\exists \mu^\infty : \quad \mu^t \to \mu^\infty \quad \mathbf{P}\text{-a.s. (and hence } \widetilde{\mathbf{P}}\text{- and } \widehat{\mathbf{P}}\text{-a.s.).}$$

1. Suppose result is false. Then, there is a positive $\widetilde{\mathbf{P}}$-probability event on which $\mu^\infty$ is strictly positive.

# Intuition

Bayes' rule determines $\mu^t$ after all histories (of 1's actions).
At any Nash eq, $\mu^t$ is a bounded martingale and so

$$\exists \mu^\infty : \quad \mu^t \to \mu^\infty \quad \textbf{P}\text{-a.s. (and hence } \widetilde{\textbf{P}}\text{- and } \widehat{\textbf{P}}\text{-a.s.).}$$

1. Suppose result is false. Then, there is a positive $\widetilde{\textbf{P}}$-probability event on which $\mu^\infty$ is strictly positive.
2. On this event, player 2 believes that both types of player 1 are eventually choosing the same distribution over actions $\hat{\alpha}_1$ (because otherwise player 2 could distinguish them).

# Intuition

Bayes' rule determines $\mu^t$ after all histories (of 1's actions).

At any Nash eq, $\mu^t$ is a bounded martingale and so

$$\exists \mu^\infty : \quad \mu^t \to \mu^\infty \quad \mathbf{P}\text{-a.s. (and hence } \widetilde{\mathbf{P}}\text{- and } \widehat{\mathbf{P}}\text{-a.s.).}$$

1. Suppose result is false. Then, there is a positive $\widetilde{\mathbf{P}}$-probability event on which $\mu^\infty$ is strictly positive.
2. On this event, player 2 believes that both types of player 1 are eventually choosing the same distribution over actions $\hat{\alpha}_1$ (because otherwise player 2 could distinguish them).
3. Consequently, on a positive $\widetilde{\mathbf{P}}$-probability set of histories, eventually, player 2 will always play a best response to $\hat{\alpha}_1$.

# Intuition

Bayes' rule determines $\mu^t$ after all histories (of 1's actions).
At any Nash eq, $\mu^t$ is a bounded martingale and so

$$\exists \mu^\infty : \quad \mu^t \to \mu^\infty \quad \mathbf{P}\text{-a.s. (and hence } \widetilde{\mathbf{P}}\text{- and } \widehat{\mathbf{P}}\text{-a.s.).}$$

1. Suppose result is false. Then, there is a positive $\widetilde{\mathbf{P}}$-probability event on which $\mu^\infty$ is strictly positive.
2. On this event, player 2 believes that both types of player 1 are eventually choosing the same distribution over actions $\hat{\alpha}_1$ (because otherwise player 2 could distinguish them).
3. Consequently, on a positive $\widetilde{\mathbf{P}}$-probability set of histories, eventually, player 2 will always play a best response to $\hat{\alpha}_1$.
4. Since player 1 is more informed than player 2, player 1 knows this.

# Intuition

Bayes' rule determines $\mu^t$ after all histories (of 1's actions).
At any Nash eq, $\mu^t$ is a bounded martingale and so

$$\exists \mu^\infty : \quad \mu^t \to \mu^\infty \quad \mathbf{P}\text{-a.s. (and hence } \widetilde{\mathbf{P}}\text{- and } \widehat{\mathbf{P}}\text{-a.s.).}$$

1. Suppose result is false. Then, there is a positive $\widetilde{\mathbf{P}}$-probability event on which $\mu^\infty$ is strictly positive.
2. On this event, player 2 believes that both types of player 1 are eventually choosing the same distribution over actions $\hat{\alpha}_1$ (because otherwise player 2 could distinguish them).
3. Consequently, on a positive $\widetilde{\mathbf{P}}$-probability set of histories, eventually, player 2 will always play a best response to $\hat{\alpha}_1$.
4. Since player 1 is more informed than player 2, player 1 knows this.
5. This yields the contradiction, since player 1 has a strict incentive to play differently from $\hat{\alpha}_1$.

# Player 2 either learns the type is normal or doesn't believe it matter-I

**Lemma**

*At any Nash eq,*

$$\lim_{t\to\infty} \mu^t(1-\mu^t)\left\|\hat{\alpha}_1 - \widetilde{E}[\sigma_1(h_1^t,\xi_0)|(h_2^t]\right\| = 0, \quad \mathbf{P}\text{-}a.s.$$

# Player 2 either learns the type is normal or doesn't believe it matter-II

For $\varepsilon > 0$ small, on the event

$$X^t := \left\{ \left\| p(h_2^t) - \hat{p}(h_2^t) \right\| < \varepsilon_1 \right\},$$

player 2 best responds to $\hat{\alpha}_1$, i.e., $\sigma_2(h_2^t) = \hat{a}_2$.

Player 2 cannot have too many $\widetilde{\mathbf{P}}$-expected surprises (i.e., periods in which player 2 both assigns a nontrivial probability to player 1 being $\hat{\xi}$ and believes $p(h_2^t)$ is far from $\hat{p}(h_2^t)$):

**Lemma**

$$\sum_{t=0}^{\infty} E_{\widetilde{\mathbf{P}}} \left[ (\mu^t)^2 (1 - \mathbb{1}_{X^t}) \right] \leq -\frac{2 \log(1 - \mu^0)}{\varepsilon_1^2},$$

*where $\mathbb{1}_{X^t}$ is the indicator function for the event $X^t$.*

# Implications of Permanent Reputations

If reputations do not disappear almost surely under $\widetilde{\mathbf{P}}$, then

$$\widetilde{\mathbf{P}}(\mu^\infty = 0) < 1,$$

and so there exists a $\lambda > 0$ and $T_0$ such that

$$0 < \widetilde{\mathbf{P}}(\mu^t \geq \lambda, \forall t \geq T_0) =: \widetilde{\mathbf{P}}(F).$$

# Implications of Permanent Reputations

If reputations do not disappear almost surely under $\widetilde{\mathbf{P}}$, then

$$\widetilde{\mathbf{P}}(\mu^\infty = 0) < 1,$$

and so there exists a $\lambda > 0$ and $T_0$ such that

$$0 < \widetilde{\mathbf{P}}(\mu^t \geq \lambda, \forall t \geq T_0) =: \widetilde{\mathbf{P}}(F).$$

On $F$, eventually player 2 believes $\xi_0$ plays $\hat{\alpha}_1$:

**Lemma**

*Suppose $\mu^t \not\to 0$ $\widetilde{\mathbf{P}}$-a.s. There exists $T_1$ such that for*

$$B := \bigcap_{t \geq T_1} X^t,$$

*we have*

$$\widetilde{\mathbf{P}}(B) \geq \widetilde{\mathbf{P}}(F \cap B) > 0.$$

# Conclusion of Argument

- On $B$, not only is player 2 always playing $\hat{a}_2$, the BR to $\hat{\alpha}_1$, but player 1 eventually is confident that 2 is doing so.

# Conclusion of Argument

- On $B$, not only is player 2 always playing $\hat{a}_2$, the BR to $\hat{\alpha}_1$, but player 1 eventually is confident that 2 is doing so.
- Moreover, again on $B$, for all $\tau$, for sufficiently large $t$, 1 is confident that 2 is doing so in periods, $t, t+1, \ldots, t+\tau$, irrespective of the signals 2 observes in periods $t, t+1, \ldots, t+\tau$.

# Conclusion of Argument

- On $B$, not only is player 2 always playing $\hat{a}_2$, the BR to $\hat{\alpha}_1$, but player 1 eventually is confident that 2 is doing so.
- Moreover, again on $B$, for all $\tau$, for sufficiently large $t$, 1 is confident that 2 is doing so in periods, $t, t+1, \ldots, t+\tau$, irrespective of the signals 2 observes in periods $t, t+1, \ldots, t+\tau$.

  Imperfect monitoring is key here: The minimum prob of any $\tau$ sequence of signals under $\hat{\alpha}_1$ is bounded away from zero.

# Conclusion of Argument

- On $B$, not only is player 2 always playing $\hat{a}_2$, the BR to $\hat{\alpha}_1$, but player 1 eventually is confident that 2 is doing so.
- Moreover, again on $B$, for all $\tau$, for sufficiently large $t$, 1 is confident that 2 is doing so in periods, $t, t+1, \ldots, t+\tau$, irrespective of the signals 2 observes in periods $t, t+1, \ldots, t+\tau$.

  Imperfect monitoring is key here: The minimum prob of any $\tau$ sequence of signals under $\hat{\alpha}_1$ is bounded away from zero.
- Contradiction: Player 1 best responding to player 2 cannot play $\hat{\alpha}_1$.

# Comments

- Result is very general. Holds if:
  - there are many types,
  - under private monitoring of both players' actions, as long as an identifiability condition holds on both players' actions (Cripps, Mailath, and Samuelson 2007, Mailath and Samuelson 2014).

# Asymptotic Restrictions on Behavior I

Result is on beliefs. What about behavior? If player 2's actions are observed by player 1, then:

For any Nash eq of the incomplete information game and for all $\widetilde{\mathbf{P}}$-almost all sequences of histories $\{h_t\}$, every cluster point of the sequence of continuation profiles is a Nash eq of the complete information game with normal type player 1.

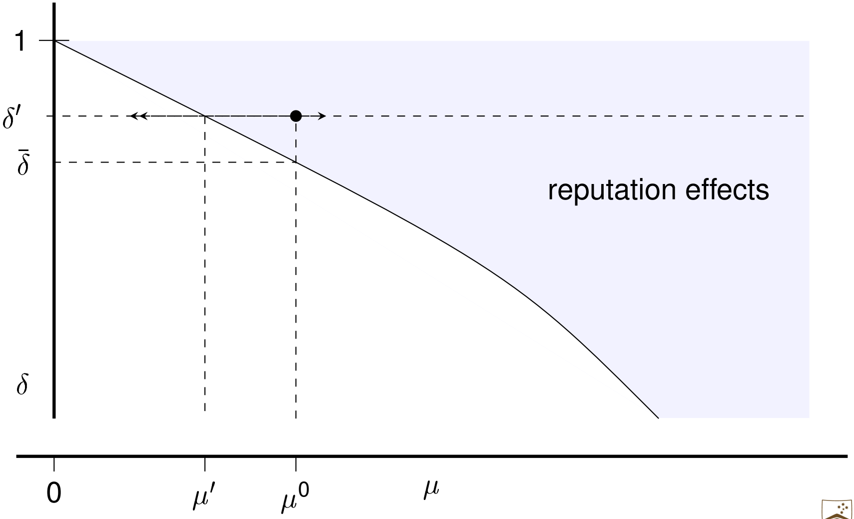
If player 2 is imperfectly monitored, then need to replace the second Nash with correlated.

# Asymptotic Restrictions on Behavior II

- Suppose player 2's actions are perfectly monitored.
- Suppose the stage game has a strict Nash equilibrium $a^*$.
- Suppose for all $\varepsilon > 0$, there exists $\eta > 0$ and an eq of the complete information game $\sigma(0)$ such that for all $\mu_0 \in (0, \eta)$ the incomplete information game with prior $\mu_0$ has an eq with player 1 payoff within $\varepsilon$ of $u_1(\sigma(0))$.

Given any prior $\mu_0$ and any $\delta$, for all $\varepsilon > 0$, there exists a Nash eq of the incomplete information game in which the $\widetilde{\mathbf{P}}$-probability of the event that eventually $a^*$ is played in every period is at least $1 - \varepsilon$.

# Interpretation



# Reputation Effects with Long-lived Player 2?

- Simple types no longer provide the best bounds on payoffs. For the repeated PD, a reputation for tit-fot-tat is valuable (while a reputation for always cooperate is not!), Kreps, Milgrom, Roberts, and Wilson (1982).

# Reputation Effects with Long-lived Player 2?

- Simple types no longer provide the best bounds on payoffs. For the repeated PD, a reputation for tit-fot-tat is valuable (while a reputation for always cooperate is not!), Kreps, Milgrom, Roberts, and Wilson (1982).
- The bound of surprises arguments still hold with long-lived player 2 (as does the disappearing reputation result), but player need not best respond to the belief that on the equilibrium path, player 1 plays like an action type.

  There are some positive results, but few and make strong assumptions.

# Persistent Reputations

How to rescue reputations?

- Limited observability
  Suppose short-lived players can only observe the last $L$ periods. Then reputations can persist and may cycle (Liu 2011, Liu and Skrzypacz 2014).
- Changing types
  Yields both cyclical reputations (Phelan 2006) and permanent reputations (Ekmekci, Gossner, and Wilson 2012).

# Reputation as Separation

- Are reputations always about scenarios where uninformed players assign positive probability to "good" types?
- Sometimes reputations are about behavior where informed players are trying to avoid a "bad" reputation.
- But avoidance of bad reputations is hard: Mailath and Samuelson (2001), Morris (2001), and Ely and Valimaki (2003).

# Further Reading