Strategic Distinguishability

Dirk Bergemann, Stephen Morris and Satoru Takahashi

October 2009

Summary

1. Canonical and universal "revealed preference" description of expected utility types for strategic analysis
2. Two types are "equilibrium strategically distinguishable" (there exists a mechanism where they must take different equilibrium actions) if and only if they map to different canonical types
3. Characterization works also with various versions of rationalizability.....

Outline of Talk

1. Introduction
2. An Example, Motivation, Discussion and Related Literature
3. The Universal Space of Expected Utility Preferences
4. Statement of The Main Result
5. Other Stuff: Strategic Equivalence, Another Example, Yet More "Redundancy", Rationalizability and Issues in the Proof

Outline of Talk

1. Introduction **[informal]**
2. An Example, Motivation, Discussion and Related Literature **[informal]**
3. The Universal Space of Expected Utility Preferences **[formal]**
4. Statement of The Main Result **[formal]**
5. Other Stuff: Strategic Equivalence, Another Example, Yet More "Redundancy", Rationalizability and Issues in the Proof **[informal]**

Example 1: Common Values and Private Signals

- Two agents
- Two equally likely states $\Omega = \{L, H\}$
- Common value of an object is 0 in state L, 90 in state H
- Each agent i observes conditional independent signal $s_i \in \{l, h\}$ with $\Pr(l|L) = \Pr(h|H) = \frac{2}{3}$

Type Space Representation 1: "Natural"

- $T_1 = T_2 = \{l, h\}$
- State space is $T_1 \times T_2 \times \Omega$ with 8 states
- Common prior:

$\omega = L$:

$t_1 \backslash t_2$	l	h
l	$\frac{2}{9}$	$\frac{1}{9}$
h	$\frac{1}{9}$	$\frac{1}{18}$

$\omega = H$:

$t_1 \backslash t_2$	l	h
l	$\frac{1}{18}$	$\frac{1}{9}$
h	$\frac{1}{9}$	$\frac{2}{9}$

- Common Valuation

$\omega = L$:

$t_1 \backslash t_2$	l	h
l	0	0
h	0	0

$\omega = H$:

$t_1 \backslash t_2$	l	h
l	90	90
h	90	90

Type Space Representation 2: Integrate out Unobserved States

- $T_1 = T_2 = \{l, h\}$
- State space is $T_1 \times T_2$ with 4 states
- Common prior:

$t_1 \backslash t_2$	l	h
l	$\frac{5}{18}$	$\frac{2}{9}$
h	$\frac{2}{9}$	$\frac{5}{18}$

- The common valuation of object is

$t_1 \backslash t_2$	l	h
l	18	45
h	45	72

Type Space Representation 3: Independent Beliefs WLOG

- $T_1 = T_2 = \{l, h\}$
- State space is $T_1 \times T_2$ with 4 states
- Common prior:

$t_1 \backslash t_2$	l	h
l	$\frac{1}{4}$	$\frac{1}{4}$
h	$\frac{1}{4}$	$\frac{1}{4}$

- The common value of object is

$t_1 \backslash t_2$	l	h
l	20	40
h	40	80

Related Literature 1

- Gul-Pesendorfer 07: A Canonical Space of Interdependent Types
- all interdependence can be mapped into psychological effects
- without loss of generality assume degenerate beliefs

Type Space Representation 4: Mertens Zamir

- $V_1 = V_2 = \{0, 90\}$
- $T_1 = T_2 = \{l, h\}$.
- State space is $T_1 \times V_1 \times T_2 \times V_2$ with 16 states.

Type h of agent 1 has beliefs:

$t_2 = l$	0	90
0	$\frac{2}{9}$	0
90	0	$\frac{2}{9}$

$t_2 = h$	0	90
0	$\frac{1}{9}$	0
90	0	$\frac{4}{9}$

Type Space Representation 5: Mertens Zamir with ex post expected valuations

- $V_1 = V_2 = \{18, 45, 72\}$
- $T_1 = T_2 = \{l, h\}$.
- State space is $T_1 \times V_1 \times T_2 \times V_2$ with 36 states.
- Type h of agent 1 has beliefs:

$t_2 = l$	18	45	72
18	0	0	0
45	0	$\frac{4}{9}$	0
72	0	0	0

$t_2 = h$	18	45	72
18	0	0	0
45	0	0	0
72	0	0	$\frac{5}{9}$

What is "Observable" about Types?

- "First level" observation about type h of agent 1
  - unconditional willingness to pay for the object is 60 ($= \frac{2}{3} \times 90$)
- "Second level" observation:
  - what is willingness to pay for the object conditional on agent 2's unconditional willingness to pay being x?
  - 20 ($= \frac{4}{9} \times 45$) if $x = 30$; 40 ($= \frac{5}{9} \times 72$) if $x = 60$; 0 otherwise

Contribution 1: Universal Expected Utility Preference Space

- Hierarchical Description:
  1. Unconditonal preferences over lottery space.
  2. Preferences over lotteries conditional on unconditional preferences of other agents....
  3. etc....
- Countable Closure, Universality, etc...
- Technically close to classic belief hierarchies of Mertens-Zamir 85, Brandenburger-Dekel 93
  - use signed measures to represent expected utility preferences plus "Kolmogorov for signed measures"

Related Literature 2: Universal Spaces of Preferences

- Epstein-Wang 96:
  - universal preference hierarchy without independence (expected utility) but with monotonicity
- Di Tillio 08:
  - universal preference hierarchy without independence or monotonicity but restricted to finite preferences

Contribution 2: A Characterization of Strategic Distinguishability

DEFINITION. A mechanism consists of a finite set of actions for each agent and an outcome function mapping action profiles to lotteries.

DEFINITION. Two types are equilibrium strategically distinguishable if there exists a mechanism for which the set of (Bayesian Nash) equilibrium actions of the two agents are disjoint.

THEOREM. Two (bounded and countable) types are equilibrium strategically distinguishable if and only if they map to distinct points in the universal (EU) preference space.

Related Literature 3: Measurability

- Abreu-Matsushima (unpublished 1992) show that a "measurability condition" characterizes when two types are strategically distinguishable (in the course of characterizing necessary and sufficient conditions for virtual Bayesian implementation under incomplete information)
- Their characterization is dependent on the finite type space in which agents live.

Related Literature 3: Measurability

- Abreu-Matsushima (unpublished 1993) show that a "measurability condition" characterizes when two types are strategically distinguishable (in the course of characterizing necessary and sufficient conditions for virtual Bayesian implementation under incomplete information)
- Their characterization is dependent on the finite type space in which agents live.
- Bergemann-Morris 09: fix "payoff type environment;" agent i knows his "payoff type" θ_i but also has "belief type" π_i (beliefs and higher order beliefs about θ_{-i}); for fixed θ_i and θ_i', BM characterize when there exist π_i and π_i' such that (θ_i, π_i) is strategically indistinguishable from (θ_i', π_i') (answer: when there is lot of interdependence in preferences).

Motivation 1: Strategic Revealed Preference

- Don't seem to have a well developed strategic analogue to single agent choice revealed preference theory
    - Single person expected utility preferences
- "Psychological/Behavioral" theories of interdependent choices cannot be distinguished from "informational" theories using this strategic revealed preference data
    - richer information or dynamic settings (with sequential rationality) required to distinguish them / extract counterfactuals

Motivation 2: Talking about Types and Implicit Common Certainty Assumptions

- A canonical language to represent types is useful in understanding our modelling assumptions
- Dasgupta-Maskin 00 and others consider two agent situation where the value of an object to agent i is $v_i = \theta_i + \frac{1}{2}\theta_j$, where θ_i is agent i's "type" or "payoff type" with $\theta_1, \theta_2 \in [0, 1]$. Agent i knows own payoff type but the planner knows nothing about what agents do or do not know or believe about other agent's payoff type.
- What is the "detail-free" content of the above assumption?

Motivation 2: Talking about Types and Implicit Common Knowledge Assumptions

- What is the "detail-free" content of the above assumption?
- Since $v_1 = \theta_1 + \frac{1}{2}\theta_2$ and $v_2 = \theta_2 + \frac{1}{2}\theta_1$, elementary linear algebra tells us that $v_1 - \frac{1}{2}v_2 = \frac{3}{4}\theta_1$ and $v_2 - \frac{1}{2}v_1 = \frac{3}{4}\theta_2$. Thus we seem to be assuming common certainty that each agent i knows the value of $v_i - \frac{1}{2}v_j$.
- But do we mean by agent i's "valuation"? We saw in the example, that there are multiple ways of representing his value using the MZ approach.
- Operational meaning of agent i's valuation: valuation conditional on strategically distinguishable types of all agents.
- Operational meaning of agent i's preferences: preferences conditional on strategically distinguishable types of all agents.

Signed Measures

- X: measurable space
- μ: signed measure (real-valued set function with σ-additivity) on X.
- $||\mu||$: total variation of μ

$$||\mu|| = \sup \left\{ \sum_{k=1}^{n} |\mu(E_k)| : \{E_1, \ldots, E_n\} \text{ is a partition of } X \right\}.$$

- $ca(X)$: set of all signed measures on X with $||\mu|| < \infty$.

Anscombe-Aumann Acts

- Z: finite set of outcomes
- $f : X \to \Delta(Z)$: measurable function (Anscombe-Aumann act)
- $F(X)$: set of all acts over X

State-Dependent Preferences

- $P(X)$: set of all binary relations $\succsim$ over $F(X)$ that are represented by $\mu \in ca(X \times Z)$:

$$f \succsim f' \Leftrightarrow \int f(x)(z)d\mu(x,z) \geq \int f'(x)(z)d\mu(x,z).$$

- (Axiomatization: Drop monotonicity (and non-degeneracy) from Anscombe-Aumann's set of axioms)
- μ is decomposed into payoffs u and beliefs ν:

$$\text{“}\mu(x,z) = \underbrace{u(x,z)}_{\text{payoff}} \times \underbrace{\nu(x)}_{\text{belief}}\text{”}$$

Preference Type Spaces

- Environment
  - finite players $\mathcal{I}$
  - finite outcomes Z
  - compact and metrizable "observable states" Θ
    - can have $\#\Theta = 1$ as in example 1
- Preference Type Space $\mathcal{T} = (T_i, \pi_i)_{i \in \mathcal{I}}$
  - T_i: measurable space of player i's types
  - $\pi_i : T_i \to P(\Theta \times T_{-i})$: measurable function that maps each type to his interdependent preference
    - example maps into this framework

Induced Preferences, Marginal Preferences

- $\varphi: X \to Y$ *induces* $\varphi^P: P(X) \to P(Y)$ by

$$\succsim \in P(X),\ f\ \phi^P(\succsim)\ f' \Leftrightarrow f \circ \varphi \succsim f' \circ \varphi.$$

- $proj_X : X \times Y \to X$ induces

$$marg_X = (proj_X)^P : P(X \times Y) \to P(X).$$

- $\text{marg}_X \succsim$ is the restriction of $\succsim$ to acts that are independent of the Y coordinate.

Construction of Hierarchies

- For simplicity, state for $I = 2$
- each $\mathcal{T} = (T_i, \pi_i)_{i \in \mathcal{I}}$, $i \in \mathcal{I}$, and $t_i \in T_i$,

$$\begin{aligned}
\hat{\pi}_{i,1}(t_i) &= \mathrm{marg}_{\Theta} \pi_i(t_i) \in P(\Theta), \\
\hat{\pi}_{i,2}(t_i) &= \left(\mathrm{id}_{\Theta} \times \hat{\pi}_{-i,1}\right)^P (\pi_i(t_i)) \in P(\Theta \times P(\Theta)), \\
\hat{\pi}_{i,3}(t_i) &= \left(\mathrm{id}_{\Theta} \times (\hat{\pi}_{-i,1}, \hat{\pi}_{-i,2})\right)^P (\pi_i(t_i)) \\
&\in P(\Theta \times P(\Theta) \times P(\Theta \times P(\Theta))), \\
&\vdots \\
\hat{\pi}_{i,n}(t_i) &= \left(\mathrm{id}_{\Theta} \times (\hat{\pi}_{-i,1}, \dots, \hat{\pi}_{-i,n-1})\right)^P (\pi_i(t_i)), \\
&\vdots
\end{aligned}$$

Construction of Hierarchies

- For each $\mathcal{T} = (T_i, \pi_i)_{i \in \mathcal{I}}$, $i \in \mathcal{I}$, and $t_i \in T_i$,

$$\hat{\pi}_{i,1}(t_i) = \mathrm{marg}_\Theta \pi_i(t_i) \in P(\Theta),$$

$$\hat{\pi}_{i,2}(t_i) = (\mathrm{id}_\Theta \times \hat{\pi}_{-i,1})^P(\pi_i(t_i)) \in P(\Theta \times P(\Theta)),$$

$$\hat{\pi}_{i,3}(t_i) = (\mathrm{id}_\Theta \times (\hat{\pi}_{-i,1}, \hat{\pi}_{-i,2}))^P(\pi_i(t_i)) \in P(\Theta \times P(\Theta) \times P(\Theta \times P(\Theta))),$$

$$\vdots$$

$$\hat{\pi}_{i,n}(t_i) = (\mathrm{id}_\Theta \times (\hat{\pi}_{-i,1}, \ldots, \hat{\pi}_{-i,n-1}))^P(\pi_i(t_i)),$$

$$\vdots$$

- $\hat{\pi}_{i,n}(t_i)$: the *n-th order preference of* t_i.
- $\hat{\pi}_i(t_i) = (\hat{\pi}_{i,1}(t_i), \hat{\pi}_{i,2}(t_i), \ldots)$: the *hierarchy of preferences of* t_i.

The Universal Type Space

- T^*: the set of all hierarchies of preferences that can arise from type spaces

Proposition

There is a "natural" Borel isomorphism

$$\pi^* \colon T^* \to P(\Theta \times T^*).$$

- $\mathcal{T}^* = (T^*, \pi^*)$: the *universal type space*.

So Far....

- Environment
  - finite players $\mathcal{I}$
  - finite outcomes Z
  - compact and metrizable observable states Θ
- Preference Type Space $\mathcal{T} = (T_i, \pi_i)_{i \in \mathcal{I}}$
  - T_i: measurable space of player i's types
  - $\pi_i : T_i \to P(\Theta \times T_{-i})$: measurable function that maps each type to his interdependent preference

Mechanisms

- Environment
- Preference Type Space $\mathcal{T} = (T_i, \pi_i)_{i \in \mathcal{I}}$
- Mechanism $\mathcal{M} = ((A_i)_{i \in \mathcal{I}}, g)$
  - A_i: finite set of player i's actions/messages
  - $g : \Theta \times A \to \Delta(Z)$: outcome function

Equilibrium

- Incomplete Information Game $(\mathcal{T}, \mathcal{M})$
- Strategy Profile $\sigma = (\sigma_i)_{i \in \mathcal{I}}$
  - $\sigma_i \colon T_i \to \Delta(A_i)$ measurable
- σ is an equilibrium of $(\mathcal{T}, \mathcal{M})$ if, for every $i \in \mathcal{I}$, $t_i \in T_i$ and $a_i \in A_i$,

$$g(\cdot, \sigma_i(t_i), \cdot) \circ (id_\Theta \times \sigma_{-i}) \quad \pi_i(t_i) \quad g(\cdot, a_i, \cdot) \circ (id_\Theta \times \sigma_{-i})$$

- $E_i(t_i, \mathcal{T}, \mathcal{M})$: set of all pure actions type $t_i \in T_i$ plays with positive probability in some equilibrium of $(\mathcal{T}, \mathcal{M})$.

Definition of Strategic Distinguishability

- Two types t_i in $\mathcal{T}$ and t_i' in $\mathcal{T}'$ are strategically indistinguishable if, for every mechanism, there exists some action that can be chosen by both types: $E_i(t_i, \mathcal{T}, \mathcal{M}) \cap E_i(t_i', \mathcal{T}', \mathcal{M}) \neq \emptyset$ for every $\mathcal{M}$.

- Conversely, t_i and t_i' are strategically distinguishable if there exists a mechanism in which no action can be chosen by both types: $E_i(t_i, \mathcal{T}, \mathcal{M}^*) \cap E_i(t_i', \mathcal{T}', \mathcal{M}^*) = \emptyset$ for some $\mathcal{M}^*$.

Necessary Condition

PROPOSITION. If $\mathcal{T}$ and $\mathcal{T}'$ are countable, then

$$\hat{\pi}_i(t_i, \mathcal{T}) = \hat{\pi}_i(t_i', \mathcal{T}')$$
$$\Rightarrow E_i(t_i, \mathcal{T}, \mathcal{M}) \cap E_i(t_i', \mathcal{T}', \mathcal{M}) \neq \emptyset.$$

Necessary Condition

PROPOSITION. If $\mathcal{T}$ and $\mathcal{T}'$ are countable, then

$$\hat{\pi}_i(t_i, \mathcal{T}) = \hat{\pi}_i(t_i', \mathcal{T}')$$
$$\Rightarrow E_i(t_i, \mathcal{T}, \mathcal{M}) \cap E_i(t_i', \mathcal{T}', \mathcal{M}) \neq \emptyset.$$

PROOF: We have

$$E_i(t_i, \mathcal{T}, \mathcal{M}) \supseteq E_i(\hat{\pi}_i(t_i, \mathcal{T}), \mathcal{T}^*, \mathcal{M}),$$
$$E_i(t_i', \mathcal{T}', \mathcal{M}) \supseteq E_i(\hat{\pi}_i(t_i', \mathcal{T}'), \mathcal{T}^*, \mathcal{M}),$$

thus

$$E_i(t_i, \mathcal{T}, \mathcal{M}) \cap E_i(t_i', \mathcal{T}', \mathcal{M}) \supseteq E_i(\hat{\pi}_i(t_i, \mathcal{T}), \mathcal{T}^*, \mathcal{M}),$$

which is non-empty by Kakutani's fixed point theorem.

Boundedness

- Let $\mu_i(t_i) \in ca(\Theta \times T_{-i} \times Z)$ be a signed measure that represents $\pi_i(t_i)$.
- $\mathcal{T}$ is *bounded by K* $< \infty$ if, for every $i \in \mathcal{I}$ and $t_i \in T_i$, we have

$$||\mu_i(t_i)|| \leq K||\mathrm{marg}_{\Theta \times Z}\mu_i(t_i)|| \neq 0.$$

Sufficient Condition

- Let d^* be a metric on T^* compatible with its product topology

PROPOSITION. For every $\varepsilon > 0$ and $K < \infty$, there exists a mechanism M^* such that, for every pair of type spaces T and T' bounded by K, we have

$$d^*(\hat{\pi}_i(t_i, \mathcal{T}), \hat{\pi}_i(t_i', \mathcal{T}')) > \varepsilon$$
$$\Rightarrow E_i(t_i, \mathcal{T}, \mathcal{M}^*) \cap E_i(t_i', \mathcal{T}', \mathcal{M}^*) = \emptyset.$$

- Discuss Proof Later

Example 2: "Redundancy," Strategic Distinguishability and Strategic Equivalence

- As in example 1....
    - Two agents
    - Two equally likely states $\Omega \in \{L, H\}$
    - Common value is 0 in state L, 90 in state H
    - Each agent i observes conditional independent signal $s_i \in \{l, h\}$ with $\Pr(l|L) = \Pr(h|H) = \frac{2}{3}$
- But now assume agent i's valuation is common value component plus private value component x_i, where $x_i = 0$ if $s_i = h$ and $x_i = 30$ if $s_i = l$

Example 1 Repeated:

Prior on type profiles (t_1, t_2);

$t_1 \backslash t_2$	l	h
l	$\frac{5}{18}$	$\frac{2}{9}$
h	$\frac{2}{9}$	$\frac{5}{18}$

and valuations

$t_1 \backslash t_2$	l	h
l	$18, 18$	$45, 45$
h	$45, 45$	$72, 72$

Example 2: "Redundancy"

Prior on type profiles (t_1, t_2);

$t_1 \backslash t_2$	l	h
l	$\frac{5}{18}$	$\frac{2}{9}$
h	$\frac{2}{9}$	$\frac{5}{18}$

and valuations

$t_1 \backslash t_2$	l	h
l	$48, 48$	$75, 45$
h	$45, 75$	$72, 72$

Example 2: "Redundancy"

- Now observe that each agent's unconditional valuation of the object is 60 independent of his type.
- Thus they map to the same type in the universal preference space
- Thus they are equilibrium strategically indistinguishable
- And they are indistinguishable from any "complete information" type with common certainty that the unconditional valuation in 60
- Analogous to (but different from) (i) Mertens-Zamir redundancy; (ii) extended redundancy in Ely-Peski 06 and Dekel-Fudenberg-Morris 07.

Example 2: Game

- Now consider the two player game where each agent can
  1. opt out; or
  2. opt in and pay 1 (for sure) and get the object and pay another 72 only if the other agent opts out.
- On the "reduced" complete information type space (without redundant types), each agent must opt out in equilibrium.
- But on the "rich" type space (with redundant types), there will be an strict equilibrium type *h* opts out and type *l* opts in.

Example 2: Game

	in	out
in	pay 1	pay 73 and get object
out	-	-

Strategic Equivalence

RECALL DEFINITION. Two types are equilibrium strategically distinguishable if there exists a mechanism for which the set of (Bayesian Nash) equilibrium actions of the two agents are disjoint. Thus two types are equilibrium strategically indistinguishable if, for every mechanism, their sets of (Bayesian Nash) equilibrium actions have a non-empty intersection.

NEW DEFINITION. Two types are equilibrium strategically equivalent if, for every mechanism, their sets of (Bayesian Nash) equilibrium actions are the same.

Back to Example 2

- Types l and h are equilibrium strategically indistinguishable
- Types l and h are *not* equilibrium strategically equivalent

Other Solution Concepts

DEFINITION. Two types are equilibrium strategically distinguishable if there exists a mechanism for which the set of (Bayesian Nash) equilibrium actions of the two agents are disjoint. Thus two types are equilibrium strategically indistinguishable if, for every mechanism, their sets of (Bayesian Nash) equilibrium actions have a non-empty intersection.

DEFINITION. Two types are equilibrium strategically equivalent if, for every mechanism, their sets of (Bayesian Nash) equilibrium actions are the same.

We can also substitute any other solution concept in the above two definitions.

Common Certainty of von-Neumann Morgenstern Utility Indices

In addition to assuming that there is common certainty expected utility maximization, we assume

1. common certainty of von Neumann-Morgenstern utility indices mapping outcomes to "utility"
2. observable states Θ (that the mechanism can condition on)
3. private goods / rich preferences

Relevant "type" is now a Θ-Mertens-Zamir type (i.e., belief and higher order beliefs about Θ)

Interim Correlated Rationalizability

- Fix a type space
- Iteratively delete actions for each type that are not best responses for any beliefs about the observable states and type/action profiles of other players that (1) are consistent with type's fixed beliefs over states and other players' types; and (2) put zero probability on deleted type/action profiles.
- Allow unexplained correlation between others' actions and the state.
- Equivalent to iterated deletion of dominated strategies / captures (a version of) common certainty of rationality.

Common Certainty of vN-M Indices / Related Literature 4

- Dekel-Fudenberg-Morris 06+07 show that two types are "interim correlated rationalizability" (ICR) strategically equivalent if and only if they have same Mertens-Zamir type
- Ely-Peski 06 gives a characterization of when two types are "interim independent rationalizability" (IIR) strategically equivalent (in terms of a richer hierarchy)
- Sadzik 07 gives charactization of when two types are equilibrium strategically equivalent
- "Redundant types" are key to these distinctions

Common Certainty of vN-M Indices

OBSERVATION. The following are equivalent:

1. Two types are equilibrium strategically indistinguishable
2. Two types are IIR strategically indistinguishable
3. Two types are ICR strategically indistinguishable
4. Two types map to the same MZ type

"PROOF" (1) $\Rightarrow$ (2) because equilibrium is refinement of IIR; (2) $\Rightarrow$ (3) because IIR is refinement of ICR; (3) $\Rightarrow$ (4) follows an adaption of DFM argument; (4) $\Rightarrow$ (1) because there always exists an equilibrium where strategies are measurable w.r.t. MZ types.

Back to General Case (w/o CC of vN-M Indices...)

- Conjecture: universal preference space characterizes ICR strategic equivalence (without common certainty of vN-M indices)

Back to General Case (w/o CC of vN-M Indices...)

- Conjecture: universal preference space characterizes ICR strategic equivalence (without common certainty of vN-M indices)
- False....

Back to Example 2

- Conjecture: universal preference space characterizes ICR strategic equivalence
- False....
- in example 2, opt out is unique ICR action on "reduced" complete information type space without redundant types
- in example 2, opt in is ICR (and IIR and equilibrium action) for type I in type space with "redundant" types

Interim Preference Correlated Rationalizability

- Fix a type space
- Iteratively delete actions for each type that are not best responses for any preferences over lotteries conditional on observable states and type/action profiles of other players that (1) are consistent with type's fixed preferences over lotteries conditional on observable states and other players' types; and (2) put "zero probability" on deleted type/action profiles.
- Allow unexplained correlation between others' actions and the player's utility from outcomes.
- Captures (a version of) common certainty of rationality.

Interim Preference Correlated Rationalizability is Permissive

	a_2	a_2'
a_1	z_0	z_1
a_1'	z_1	z_2

- Complete information
- Player 1's vN-M utility index $(u_1(z_0), u_1(z_1), u_1(z_2)) = (0, 1, 2)$
- a_1' is unique interim (belief) correlated rationalizable action
- but a_1 is a preference correlated best response, given uniform prior and utility over outcomes is $(0, -2, 2)$ conditional on a_2 but $(0, 4, 2)$ conditional on a_2'
- this is consistent with common certainty preferences $(0, 1, 2)$.

Interim Preference Correlated Rationalizability is very Permissive

- An action is preference correlated rationalizable if and only if it is a preference correlated best response.

Related Literature 5

Ledyard (1986) "The Scope of the Hypothesis of Bayesian Equilibrium" JET 1986

A Family of Rationalizability Notions

- Fix a type space
- Iteratively delete actions for each type that are not best responses for any preferences over lotteries conditional on observable states and type/action profiles of other players that (1) are consistent with type's fixed preferences over lotteries conditional on observable states and other players' types; (2) put "zero probability" on deleted type/action profiles; and (3) belong to set Φ.
- Φ is a collection of preferences. Larger or smaller Φ capture more permissive and less permissive version of rationalizability. Φ is rich if includes preferences given behavioral strategy of opponent.
- An action is Φ-rationalizable for a type if it survives this iterative process.

Related Literature 6

Battigalli and Siniscalchi BJATE 03 Δ-rationalizability.

Necessary Condition

PROPOSITION. Two (countable, bounded) types corresponding to the same type in the universal preference space have (i) non-empty intersection of (i) equilibrium actions and thus (ii) Φ-rationalizable actions for each rich Φ.
PROOF. We already proved (i). (ii) follows from rationalizability of equilibrium actions.

Sufficient Condition

- Φ is uniformly bounded if there is a uniform bound on how much others' actions can change a player's preferences; intuitively, not too much redundancy is built into the solution concept.

PROPOSITION. If there is common certainty of no complete indifference, for a fixed uniform bound K, for any two types corresponding to distinct types in the universal preference space, there is a mechanism such that those two types have no Φ-rationalizable actions in common, for each rich Φ bounded by K, and thus no equilibrium actions in common.

Issues in the Proof of Sufficient Condition

- compare Abreu-Matsushima 92, DFM 07, BM 09 and this paper
- all will construct canonical mechanism with players reporting 1st level preferences/beliefs, 2nd level preferences/beliefs, etc...
- for each player i and each $k = 1, 2, \ldots$, there will be (with some positive probability) a lottery y_{ik} chosen that depends on kth level report of player i and the $(k-1)$th reports of players other than i
- this should give player i an incentive to report his kth level preferences/beliefs correctly if he thinks others are reporting their $(k-1)$th level preferences/beliefs correctly.

Issues in the Proof of Sufficient Condition

- all will construct canonical mechanism with players reporting 1st level preferences/beliefs, 2nd level preferences/beliefs, etc...
- for each player i and each $k = 1, 2, ...,$ there will be (with some positive probability) a lottery y_{ik} chosen that depends on kth level report of player i and the $(k-1)$th reports of players other than i
- this should give player i an incentive to report his kth level preferences/beliefs correctly if he thinks others are reporting their $(k-1)$th level preferences/beliefs correctly.
- key problem: ensuring that player i does not have incentive to mis-report his kth level preferences/beliefs in order to manipulate $y_{j,k+1}$ for $j \neq i$

Issues in the Proof of Result 2

- key problem: ensuring that player i does not have incentive to mis-report his kth level preferences/beliefs in order to manipulate $y_{j,k+1}$ for $j \neq i$
- Abreu-Matsushima 93: exploit finiteness of types
- BM 09: exploit finiteness of "payoff types"
- DFM 07 exploit private goods (agent i is indifferent about $y_{j,k+1}$)
- this paper: boundedness need for continuity argument, care in order of limits