Justifying Optimal Play via Consistency

Felix Brandt
(joint work with Florian Brandl)

Sorbonne Economics Centre, Paris, March 2021

DSS
Decision Sciences & Systems

	$\frac{1}{3}$	$\frac{2}{3}$	
$\frac{1}{3}$	2	0	3
$\frac{2}{3}$	0	1	0
	1	0	5

Why should one play maximin strategies in two-player zero-sum games?

John von Neumann

- Von Neumann's minimax theorem (1928) shows that the best outcome that the row player can guarantee coincides with the best outcome the column player can guarantee.
  - All pairs of maximin strategies are Nash equilibria, which furthermore yield the same payoff.
  - The set of Nash equilibria is convex.
  - Nash equilibria of zero-sum games can be efficiently computed.
  - "Every two-person zero-sum game is determined [...] it has precisely one individually rational payoff vector" (Aumann, 1987)
- Yet, providing normative foundations for maximin play turns out to be surprisingly difficult.


Robert Aumann

Related Work

- Epistemic approaches
  - Bayesian belief hierarchies, which capture players' knowledge about each other (e.g., Aumann & Brandenburger, 1995; Aumann & Drèze, 2008)
- Characterizations of the value
  - Typically not motivated on normative grounds; value is devoid of any strategic content (e.g., Vilkas, 1963; Tijs, 1981; Hart et al., 1994; Norde & Voorneveld, 2004)
- Characterizations of Nash equilibrium
  - Consistency axiom for variable number of players (Peleg & Tijs, 1996, Norde et al., 1996)

Summary

- *Our approach*: Characterize maximin strategies via decision-theoretic axioms that require players to behave coherently across hypothetical games.

- *Our result*: A rational and consistent consequentialist who ascribes the same properties to his opponent must play maximin strategies.

- The result can be turned into a characterization of Nash equilibrium in unrestricted (non-zero-sum) games.

The Model

- U: Infinite universal set of actions
  - $\mathscr{F}(U)$: set of *finite* subsets of U
- $M \in \mathbb{Q}^{A \times B}$: zero-sum game with action sets $A, B \in \mathscr{F}(U)$
- $\Delta(A)$: set of *rational-valued* strategies over $A \in \mathscr{F}(U)$
- f: solution concept mapping a game M to a set of recommended strategies $f(M) \subseteq \Delta(A)$ for the row player
  - $maximin(M) = \arg \max_{p \in \Delta(A)} \min_{q \in \Delta(B)} p^t M q$

$U = \{a, b, c, \ldots\}$

$A = \{a, b\} \in \mathscr{F}(U)$

$$M = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

$p = (1/2, 1/2) \in \Delta(A)$

$maximin(M) = \{(2/3, 1/3)\}$

Consequentialism

Players do not distinguish between payoff-equivalent actions.

- Decision-theoretic precursors
  - Chernoff (1954)'s *Postulate 6 (cloning of player's actions)* and *Postulate 9 (cloning of nature's states)*
  - *Column duplication* (Milnor, 1954)
  - *Deletion of repetitious states* (Arrow and Hurwicz, 1972; Maskin, 1979)
- Implies invariance w.r.t. permutations of actions
  - Chernoff (1954)'s *Postulate 3*
  - *Symmetry* (Milnor, 1954)

Consequentialism

Players do not distinguish between payoff-equivalent actions.

- Let $A, B \in \mathscr{F}(U)$, $\hat{A} \subseteq A$, $\hat{B} \subseteq B$, $M \in \mathbb{Q}^{A \times B}$, and $\hat{M} \in \mathbb{Q}^{\hat{A} \times \hat{B}}$ such that there exist surjective functions
  $\alpha \colon A \to \hat{A}$ and
  $\beta \colon B \to \hat{B}$ with
  $M_{ab} = \hat{M}_{\alpha(a)\beta(b)}$ for all $(a, b) \in A \times B$.

- Then,
  $$f(M) = \bigcup_{\hat{p} \in f(\hat{M})} \{p \in \Delta(A) \colon \sum_{a \in \alpha^{-1}(\hat{a})} p(a) = \hat{p}(\hat{a}) \text{ for all } \hat{a} \in \hat{A}\}.$$

- Let $A, B \in \mathscr{F}(U)$, $\hat{A} \subseteq A$, $\hat{B} \subseteq B$, $M \in \mathbb{Q}^{A\times B}$, and $\hat{M} \in \mathbb{Q}^{\hat{A}\times\hat{B}}$ such that there exist surjective functions
  $\alpha\colon A \to \hat{A}$ and
  $\beta\colon B \to \hat{B}$ with
  $M_{ab} = \hat{M}_{\alpha(a)\beta(b)}$ for all $(a,b) \in A \times B$.

- Then,
$$f(M) = \bigcup_{\hat{p}\in f(\hat{M})} \{p \in \Delta(A)\colon \sum_{a\in\alpha^{-1}(\hat{a})} p(a) = \hat{p}(\hat{a}) \text{ for all } \hat{a} \in \hat{A}\}.$$

- Example:

$$M = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 2 \end{pmatrix} \qquad \hat{M} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

$$f(M) = \{(2/3, \lambda, 1/3 - \lambda)\colon \lambda \in [0, 1/3]\} \qquad f(\hat{M}) = \{(2/3, 1/3)\}$$

Consistency

A strategy recommended for two different games will also be recommended if there is uncertainty which of the games will be played.

- Let $A, B \in \mathscr{F}(U)$, and $\hat{M}, \bar{M} \in \mathbb{Q}^{A \times B}$, $\lambda \in [0,1] \cap \mathbb{Q}$.
- If $f(\hat{M}) \cap f(\bar{M}) \neq \emptyset$ and $f(-\hat{M}^t) \cap f(-\bar{M}^t) \neq \emptyset$, then

$$f(\hat{M}) \cap f(\bar{M}) \subseteq f(\lambda\hat{M} + (1-\lambda)\bar{M}).$$

Consistency

- Let $A, B \in \mathscr{F}(U)$, and $\hat{M}, \bar{M} \in \mathbb{Q}^{A\times B}$, $\lambda \in [0,1] \cap \mathbb{Q}$.
- If $f(\hat{M}) \cap f(\bar{M}) \neq \emptyset$ and $f(-\hat{M}^t) \cap f(-\bar{M}^t) \neq \emptyset$, then

$$f(\hat{M}) \cap f(\bar{M}) \subseteq f(\lambda \hat{M} + (1-\lambda)\bar{M}).$$

- Example:

$$\hat{M} = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 2 \\ 0 & 0 & 4 \end{pmatrix} \quad \bar{M} = \begin{pmatrix} 2 & 4 & 0 \\ 0 & 0 & 4 \\ 4 & 0 & 0 \end{pmatrix} \quad \tfrac{1}{2}\hat{M} + \tfrac{1}{2}\bar{M} = \begin{pmatrix} 3 & 2 & 0 \\ 0 & 2 & 3 \\ 2 & 0 & 2 \end{pmatrix}$$

$$f(\hat{M}) = f(\bar{M}) = \{(\tfrac{2}{5}, \tfrac{2}{5}, \tfrac{1}{5})\}$$
$$f(-\hat{M}^t) = f(-\bar{M}^t) = \{(\tfrac{2}{5}, \tfrac{1}{5}, \tfrac{2}{5})\}$$

$$(\tfrac{2}{5}, \tfrac{2}{5}, \tfrac{1}{5}) \in f(\tfrac{1}{2}\hat{M} + \tfrac{1}{2}\bar{M})$$

Rationality

Strictly dominated actions are not recommended.

- Classic axiom from decision theory
  - *Strong domination* (Milnor, 1954)
  - *Property (5)* (Maskin, 1979)
  - weaker than Chernoff (1954)'s *Postulate 2*

Rationality

Strictly dominated actions are not recommended.

- Let $A, B \in \mathscr{F}(U)$ and $M \in \mathbb{Q}^{A \times B}$.
- $f(M) \subseteq \{p \in \Delta(A) : \forall a \in A\ \exists \hat{a} \in A\ \forall b \in B,\ M_{ab} < M_{\hat{a}b} \Rightarrow p(a) \neq 1\}$
- Example:

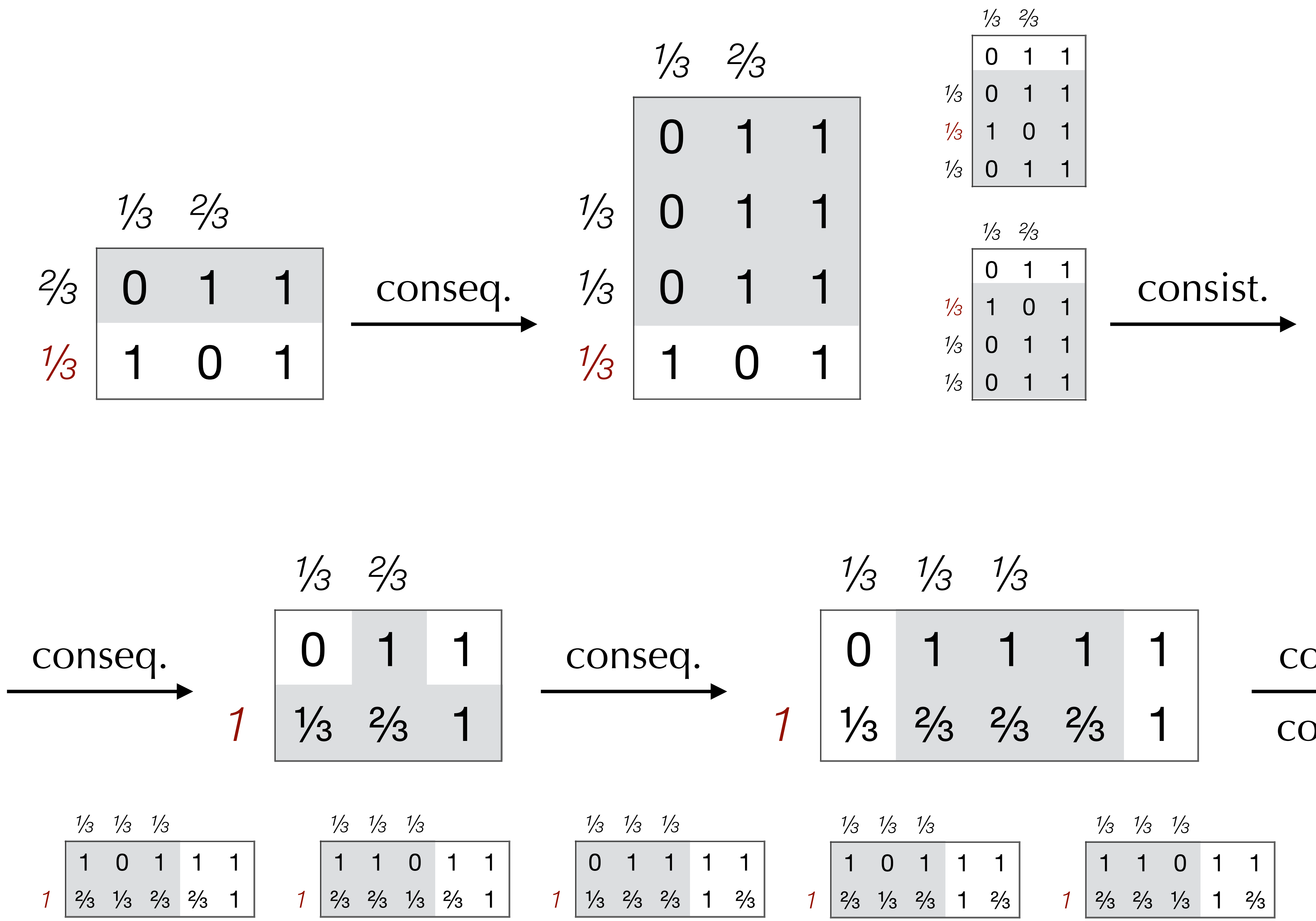
$$M = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \qquad f(M) \subseteq \{(\lambda, 1-\lambda) : \lambda \in (0,1]\}$$

The Result

- If f satisfies consequentialism, consistency, and rationality, then $f(M) \subseteq maximin(M)$ for all $A, B \in \mathcal{F}(U)$, $M \in \mathbb{Q}^{A \times B}$.

- Proof idea:
  - If one of the players does not play a maximin strategy, their strategies do not constitute a Nash equilibrium.
  - Use consequentialism and consistency to construct a game in which the player who has a profitable deviation plays a dominated action with probability 1.
  - This contradicts rationality.

Proof Sketch



Row player plays a dominated action ↯

Independence of Axioms

- All axioms are required for the characterization of *maximin*.
  - The solution concept that returns all lotteries violates rationality.
  - *maximax* (returns all randomizations over rows that contain a maximal entry of the game matrix) violates consistency.
    - $\hat{M} = \begin{pmatrix} 5 & 1 & 0 \\ 4 & 4 & 0 \end{pmatrix}$ $\bar{M} = \begin{pmatrix} 1 & 5 & 0 \\ 4 & 4 & 0 \end{pmatrix}$ $½\,\hat{M} + ½\,\bar{M} = \begin{pmatrix} 3 & 3 & 0 \\ 4 & 4 & 0 \end{pmatrix}$
  - *average* (all randomizations over rows with maximal average payoff) violates consequentialism.
    - $\hat{M} = \begin{pmatrix} 0 & 2 & 2 \\ 3 & 0 & 0 \end{pmatrix}$ $\bar{M} = \begin{pmatrix} 0 & 2 \\ 3 & 0 \end{pmatrix}$

Strong Consistency

- *maximin* violates strong consistency: $f(\hat{M}) \cap f(\bar{M}) \neq \emptyset$ implies $f(\hat{M}) \cap f(\bar{M}) \subseteq f(\lambda\hat{M} + (1-\lambda)\bar{M})$.
  - (Consistency additionally requires $f(-\hat{M}^t) \cap f(-\bar{M}^t) \neq \emptyset$.)

$$\hat{M} = \begin{pmatrix} 0 & 4 \\ 2 & 0 \end{pmatrix} \quad \bar{M} = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix} \quad M = 1/2\,\hat{M} + 1/2\,\bar{M} = \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix}$$

- The characterization also holds in the domain of symmetric zero-sum games (via a simpler proof).
  - In this case, consistency and strong consistency coincide.

John Nash

Extensions

- Assuming that f is upper hemi-continuous allows to
  - extend the result to games with real-valued payoffs,
  - show that $f(M) = maximin(M)$,
  - weaken consistency by fixing $\lambda = 1/2$, and
  - weaken rationality by restricting it to 2x1 games.
- When considering general (non-zero-sum) multi-player games and solution concepts that return strategy *profiles*, one obtains a characterization of Nash equilibrium.
  - However, recommendations are not independent anymore!