Dynamic Games and Applications
https://doi.org/10.1007/s13235-024-00595-2

Learning Decision Criteria from Play

Paolo Galeazzi[1] · **Mathias W. Madsen**[2]

Accepted: 5 September 2024

Abstract
This paper investigates population games under ambiguity in which players may adopt decision criteria different from one another. After defining equilibria for these situations by extending well-known decision-theoretic criteria to the game-theoretic context, we apply these concepts to examine the case of two-person games played within a population whose relative proportions of decision criteria are unknown to the players. We state necessary and sufficient conditions under which such games prompt the players to reveal their decision criterion through their actions, and we show when the relative proportions may be learned by observing the increasingly informed agents play.

Keywords Decision criteria · Learning · Population games

1 Introduction

In the last decades, the necessity of elevating the analysis of the individuals' behavior from observed actions to underlying mechanisms has emerged in various ways in different fields, from psychology to ecology to evolutionary game theory [1, 9, 14, 16, 17, 19–21, 23, 31]. The general idea, as expressed e.g. by [14], is that

> Natural environments are so complex, dynamic, and unpredictable that natural selection cannot possibly furnish an animal with an appropriate, specific behavior pattern for every conceivable situation it might encounter. Instead, we should expect animals to have evolved a set of psychological mechanisms which enable them to perform well on average across a range of different circumstances.

Here we take this idea seriously and consider a game-theoretic model where a population of agents inhabits an environment consisting of a multitude of different games (which we call multigame), and each individual in the population is endowed with a "psychological mechanism" that produces a specific behavior for any possible game in the environment. The main research question that we want to investigate then concerns the possibility of discerning the different underlying mechanisms by observing the agents' expressed behaviors only. That

✉ Paolo Galeazzi
pagale87@gmail.com

1 University of Bayreuth, Bayreuth, Germany

2 Micropsi industries, Berlin, Germany

Published online: 26 September 2024

is: Under which conditions is it possible to distinguish the agents' underlying mechanisms given their behavior?

In this work, we present a case where the agents' behavior-generating mechanisms are represented by different decision criteria and each of such criteria makes the agent act in a certain way when faced with a specific game in the environment. The decision criteria that we consider in the following are arguably the two main criteria for choice under ambiguity from the decision-theoretic literature, i.e., maxmin expected utility and regret minimization.

The following example introduces a simple instance of the population model that we study in this paper. Consider a population living in an environment consisting of the three games below. The first is a Prisoner's Dilemma (PD), the second is a Stag Hunt (SH), and the third is an anti-coordination game (AG).

PD	I	II
I	2,2	0,3
II	3,0	1,1

SH	I	II
I	3,3	0,2
II	2,0	2,2

AG	I	II
I	1,1	2,5
II	5,2	0,0

Individuals from such population randomly meet and play one of these three possible games. Crucially, however, each individual plays the game based on her own behavior-generating mechanism, namely, her own decision criterion—which we also refer to as the individual's type. A maxmin player by definition chooses an action that guarantees the highest minimum payoff. In the SH game, for instance, the minimum payoff that one can get from action I is 0, while the minimum payoff from action II is 2, and a maxmin player would therefore choose action II in the SH game. A regret-minimizing player instead aims to choose an action that minimizes the regret, defined as the maximum amount possibly given up by playing a certain action. For instance, the maximum regret from action I in the SH game is 2, which is the payoff given up by playing action I when the opponent plays II. By similar reasoning, the regret from action II in the SH game is 1, which is the payoff given up by choosing II when the opponent chooses I. A regret minimizer would hence play II in the SH game.

Similar computations lead to the conclusion that both a maxmin player and a regret-minimizing player would choose action II in the PD game too. In an environment consisting uniquely of one or both of these two games the two player types would thus be behaviorally indistinguishable. Looking at the AG game, however, one can compute that a maxmin player would play action I while a regret minimizer would play action II. In the environment including all three games, the different types are distinguishable.

In the following, we study the conditions on the games in the environment ensuring that the types in the population are distinguishable. To do that, we first define the concepts of games with ambiguity on the decision criteria and of equilibrium in such games in Sect. 2, and then we state the conditions for a 2×2 game to be informative, that is, to allow telling different types apart, in Sect. 3. In Sect. 4, we introduce the population multigame, we study the properties of the environment that guarantee that informative 2×2 games can always occur with positive probability, and we generalize the results to the case of $n \times n$ games. Section 5 then considers a specific instance of population multigame and shows how to compute the probability of informative games in that particular case and that the agents can asymptotically learn the precise proportions of types in the population. Section 6 instead considers a variety of different multigames and computes the probability of strongly informative games in different cases by means of computer simulations. Finally, Sect. 7 concludes. Before moving to Sect. 2, however, in the next subsection we say a few words on the literature related to the present work.

1.1 Related Literature

Although, as already mentioned above, the necessity of developing models with agents characterized by different behavior-generating mechanisms has been explicitly advocated especially by biologists and ecologists [14, 21, 23], the game-theoretic literature in economics has almost always focused on models with homogeneous decision criteria—i.e., models where all the agents follow the same decision criterion [3, 22, 24–28, 30, 33]. A few exceptions are the following. [2] study necessary and sufficient conditions for the existence of an equilibrium in games under ambiguity where the agents can have very general subjective choice preferences. [10, 12, 13] and [18] introduce epistemic type spaces that allow the agents to follow different decision criteria and to reason strategically about each other's criteria, but their results are purely on the epistemic side. On the evolutionary side, the evolution of preferences [9, 11] is a branch of evolutionary game theory that studies the evolutionary fitness of different subjective preferences, but the focus there is on the players' subjective utility functions rather than on the players' decision criteria. Moreover, here we are primarily interested in the learning and not in the evolution of the players. The idea of investigating multigame models has sometimes appeared in other fields too. [16] and [17] study the evolution of decision criteria in an environment similar to the one we consider here. [4] too consider evolutionary processes driven by a multigame environment, but with the difference that their agents are defined by automata rather than by decision criteria. [38] exploits a multigame consisting of three different games to explain the evolution of fairness, but the types there are decision rules specific to those games and hence simpler than the decision criteria examined here.

2 Games with Criterion Ambiguity

In the simple example from the previous section, the players were implicitly assumed to hold uncertainty over the opponents' actions and to use possibly different decision criteria to cope with such uncertainty and to pick an action for any given game in the environment. In population games, however, it is natural to imagine that the players' uncertainty comes from the distribution of the different types in the population. In this section, we formally introduce population games with criterion ambiguity and show how the uncertainty on the opponent's actions is derived from the uncertainty on the type distribution in the population. We interpret these games as modeling a situation in which players are drawn at random from a large and mixed population where maxmin types (M) and regret-minimizing types (R) coexist in unknown proportions. The concept of game with criterion ambiguity and that of equilibrium in such games can then be formalized as follows.

Definition 1 A *game with criterion ambiguity* consists of the following components:

- a set of worlds Ω;
- a set of parameters Λ;
- a set $I = \{1, 2, 3, \ldots, N\}$ of players;
- for each player i,
  - an action set A_i;
  - a set of (criterion) types T_i;
  - a set of signals S_i;
  - a criterion-assignment function $\tau_i : \Omega \to T_i$;

  - a signal function $\varsigma_i : \Omega \to S_i$;
  - a utility function $u_i : A_1 \times \cdots \times A_n \times \Omega \to \mathbb{R}$;

- for each $\lambda \in \Lambda$, a probability distribution P_λ over Ω.

A *policy* of player i in such a game is a function $\sigma_i : T_i \times S_i \to A_i$ that associates each decision criterion and signal with an action.[1]

Note that if the set Λ is equipped with a probability distribution known to all the players, this definition describes a Bayesian game [29]. However, we are interested in situations where this is not the case and the players hold ambiguity (i.e., non-probabilistic uncertainty) about the parameter λ.

Throughout the rest of the paper we denote typical profiles of signals and types by $s \in S := S_1 \times S_2 \times \cdots \times S_N$ and $t \in T := T_1 \times T_2 \times \cdots \times T_N$, respectively. For profiles of types, signals, or policies, we use the notation $x_{-i} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_N)$ for the $(N-1)$-dimensional vector that results by removing the ith coordinate from the vector x.

In games where the uncertainty only concerns the criteria adopted by the others, no private information other than their own decision criterion is revealed to the agents. In the notation above, this can be modeled by assuming that the sets S_i are all singletons. To ease notation, we therefore dispense with the specification of the private signals s_i in the following discussion.

For any fixed value of λ, the type of a player is a random variable. Hence, given policies $(\sigma_i)_{i\in I}$, the action $a_i = \sigma_i(t_i)$ too is a random variable, and $(a_1, a_2, \ldots, a_N)$ is a random vector. The utility of each player is therefore also a random variable, which has an expected value for any fixed λ. Since all of these distributions depend on λ, the value of this expected utility is a function of λ. We then let each player resolve the ambiguity about this expected utility by means of the decision criterion given by his or her type.

In formulas, the expected utility that follows from type t_i using action a_i, given that the other players adopt policies σ_{-i}, is

$$E_\lambda[u_i \mid a_i, t_i, \sigma_{-i}] = \int_\Omega u_i\,(a_i, \sigma_{-i}(\tau_{-i}(\omega)), \omega)\; P_\lambda(d\omega \mid t_i). \tag{1}$$

Definition 2 Let an incomplete information game with criterion ambiguity be given as above, with $T_i = \{M, R\}$ for all $i \in I$. We say that a policy profile $(\sigma_i^*)_{i\in I}$ is an *equilibrium* if, for all $i \in I$, the policy σ_i^* maximizes the function

$$\sigma_i \;\mapsto\; \inf_{\lambda\in\Lambda} E_\lambda[u_i \mid \sigma_i(M), M, \sigma_{-i}^*]$$

and minimizes the function

$$\sigma_i \;\mapsto\; \sup_{\lambda\in\Lambda} \left\{ \sup_{a_i\in A_i} E_\lambda[u_i \mid a_i, R, \sigma_{-i}^*] - E_\lambda[u_i \mid \sigma_i(R), R, \sigma_{-i}^*] \right\}.$$

2.1 Population Games with Criterion Ambiguity

As a main source of ambiguity about the criteria, we consider a large population of agents using different decision criteria that randomly meet to play games. Since it is often unlikely to know the precise distribution of types in a large population, we allow the agents to hold unmeasurable uncertainty about the distribution of criteria in the population they are part

[1] There is no consensus in game theory about the possibility of using mixed actions (see for instance [32] for a discussion). Here we follow the position of [13] and [12].

of. In particular, the rest of this paper is concerned with a class $\mathcal{G}$ of two-player population games with criterion ambiguity. At first, we focus on 2×2 symmetric games, and we later show how the results can be generalized to $n \times n$ symmetric games.[2]

In the following, we assume that each of the two players $i \in \{1, 2\}$ has his or her own decision criterion $t_i \in \{M, R\}$ revealed, but holds unmeasurable uncertainty about the opponent's criterion. Specifically, we assume that

$$P_\lambda(t_{3-i} = R \mid t_i = R) = P_\lambda(t_{3-i} = R \mid t_i = M) = \lambda$$
$$P_\lambda(t_{3-i} = M \mid t_i = R) = P_\lambda(t_{3-i} = M \mid t_i = M) = 1 - \lambda$$

where $\lambda \in [\underline{\lambda}, \overline{\lambda}] \subseteq [0, 1]$ is subject to unmeasurable uncertainty. In the case of 2×2 games, the agents are equipped with the binary action sets $A_1 = A_2 = \{I, II\}$, and their utility functions are defined in terms of the symmetric 2×2 game matrix

	I	II
I	a, a	b, c
II	c, b	d, d

for all $\omega \in \Omega$, where $(a, b, c, d) \in \mathbb{R}^4$. The two players of this game are then agents sampled at random from a large population characterized by the unknown parameter λ. For convenience, we use the vector notation $\sigma_i = (\sigma_i(M), \sigma_i(R))$ to specify the policy function of each agent $i \in \{1, 2\}$.

Example 3 Suppose two agents are randomly drawn from a population consisting of a proportion λ of regret minimizers and $1-\lambda$ of maximinimizers. The exact value of λ is unknown and subject to ambiguity, with $\lambda \in [1/5, 1/2]$. The two agents thus play the following coordination game under ambiguity about the value of λ.

	I	II
I	1, 1	0, 0
II	0, 0	2, 2

What are the equilibria of this game? Each player in this game is a regret type with probability λ and a maxmin type with probability $1 - \lambda$, and their policy functions take the form of pairs of pure actions, with one action for each of these two types. By inspecting each of the 16 policy profiles, we can reject the ones in which any of the two players is not using a best reply.

Consider first the case in which the row player faces the column policy $\sigma_2 = (I, I)$. Given such a homogeneous policy, the action of the column player is deterministic and hence not subject to uncertainty, unmeasurable or otherwise. We therefore find that the unique best reply to this policy is $\sigma_1 = (I, I)$. We similarly find that the unique best reply to $\sigma_2 = (II, II)$ is $\sigma_1 = (II, II)$. Since the exact same argument holds for the column player, it follows that the policies

$$(\sigma_1, \sigma_2) = ((I, I), (I, I))$$
$$(\sigma_1, \sigma_2) = ((II, II), (II, II))$$

are both equilibria of this game, and that the homogeneous policies (I, I) and (II, II) appear in no other equilibria.

[2] Having symmetric games only allows us to stick with the single-population model, as we need not consider a different population for each role in the game.

Suppose now that the column player uses the policy $\sigma_2 = (I, II)$. Then the conditional expected utilities of the row player given λ are

$$E_\lambda[u_1 \mid (I, (I, II))] = 1 - \lambda$$
$$E_\lambda[u_1 \mid (II, (I, II))] = 2\lambda$$

For the maxmin type of the row player, action I is a best reply to $\sigma_2 = (I, II)$ since the inequality

$$\min_{\lambda \in [1/5, 1/2]} 1 - \lambda \ \geq \ \min_{\lambda \in [1/5, 1/2]} 2\lambda$$

reduces to the true statement $1/2 \geq 2/5$. For the regret-minimizing type of the row player, on the other hand, action II is a best reply to $\sigma_2 = (I, II)$. This follows from the fact that the conditional regrets of action I and II given λ are

$$\max_{a_1 \in A_1} E_\lambda[u_1 \mid (a_1, (I, II))] - E_\lambda[u_1 \mid (I, (I, II))] = \max\{0, 3\lambda - 1\}$$
$$\max_{a_1 \in A_1} E_\lambda[u_1 \mid (a_1, (I, II))] - E_\lambda[u_1 \mid (I, (I, II))] = \max\{1 - 3\lambda, 0\},$$

and

$$\max_{\lambda \in [1/5, 1/2]} \{\max\{0, 3\lambda - 1\}\} \ \geq \ \max_{\lambda \in [1/5, 1/2]} \{\max\{1 - 3\lambda, 0\}\}$$

reduces to the true statement $1/2 \geq 2/5$. The policy $\sigma_1 = (I, II)$ is thus a best reply to the policy $\sigma_2 = (I, II)$.

Suppose now that the column player uses $\sigma_2 = (II, I)$. We then find that

$$E_\lambda[u_1 \mid (I, (II, I))] = \lambda$$
$$E_\lambda[u_1 \mid (II, (II, I))] = 2(1 - \lambda)$$

It follows that action II is a best reply for the regret type, since

$$\max_{a_1 \in A_1} E_\lambda[u_1 \mid (a_1, (II, I))] - E_\lambda[u_1 \mid (I, (II, I))] = \max\{0, 2 - 3\lambda\}$$
$$\max_{a_1 \in A_1} E_\lambda[u_1 \mid (a_1, (II, I))] - E_\lambda[u_1 \mid (II, (II, I))] = \max\{3\lambda - 2, 0\}$$

and

$$\max_{\lambda \in [1/5, 1/2]} \{\max\{0, 2 - 3\lambda\}\} \geq \max_{\lambda \in [1/5, 1/2]} \{\max\{3\lambda - 2, 0\}\}$$

reduces to the true statement $7/5 \geq 0$. Action I is thus not a regret-minimizing reply to the policy $\sigma_1 = (II, I)$, and hence $\sigma_2 = (II, I)$ cannot be a best reply to $\sigma_1 = (II, I)$. Since we have already ruled out the options $\sigma_2 = (I, I)$ and $\sigma_2 = (II, II)$ above, the only remaining option is $\sigma_2 = (I, II)$. However, we have also seen that $(\sigma_1, \sigma_2) = ((I, II), (II, I))$ is not an equilibrium, and since the game is symmetric, neither is $(\sigma_1, \sigma_2) = ((II, I), (I, II))$.

In sum, we have that the three policy profiles

$$(\sigma_1, \sigma_2) = ((I, I), (I, I))$$
$$(\sigma_1, \sigma_2) = ((II, II), (II, II))$$
$$(\sigma_1, \sigma_2) = ((I, II), (I, II))$$

are the only equilibria of the game.

We are interested here in symmetric equilibria, i.e., equilibria such that $\sigma_1 = \sigma_2$, as this is the only type of equilibrium that can be interpreted as a population adaptive strategy. In particular, in the case of 2×2 games we are interested in the policy profiles

$$(\sigma_1, \sigma_2) = ((I, II), (I, II))$$
$$(\sigma_1, \sigma_2) = ((II, I), (II, I))$$

since a population playing according to any of these profiles allows the observers to infer the decision criteria of the players from their actions. In games where exactly one of these profiles is the sole symmetric equilibrium, the players necessarily reveal their decision criterion. We then say that such games are *strongly informative*, and in the next section we provide necessary and sufficient conditions for a game to be strongly informative.

3 Strongly Informative Games

Strongly informative games are the key to the learning of the actual proportions of decision criteria in the population, because only by playing strongly informative games the players unambiguously reveal their type. For any interval $[\underline{\lambda}, \overline{\lambda}]$ with $\underline{\lambda} < \overline{\lambda}$, it is solely the positive probability of a strongly informative game to occur that can give the players relevant information about the proportions in the population. In this section, we characterize the region of strongly informative games in $\mathbb{R}^4$ and we next provide conditions on the distribution of possible games in the class $\mathcal{G}$ that guarantee the occurrence of strongly informative games.

3.1 Strong Informativity with Full Uncertainty

We first formulate the conditions under which a game (a, b, c, d) is strongly informative given that all the players have full unmeasurable uncertainty about the proportions of decision criteria, i.e., for $\lambda \in [0, 1]$. As a first step, we can immediately reduce the set of strongly informative games by discarding all games that are not anti-coordination games. In the following, informativity in 2×2 games will be enough for most of our purposes. However, we can prove the following result for general $n \times n$ games. To that aim, we partition the class of symmetric $n \times n$ games into three sets:

- coordination games: $a_i \in br(a_i)$ for all pure actions a_i,
- anti-coordination games: $a_i \notin br(a_i)$ for all pure actions a_i,
- mixed games: $a_i \in br(a_i)$ for some a_i, and $a_i' \notin br(a_i')$ for some a_i',

where $br(a_i)$ is the set of best replies to action a_i. Then the following result holds.

Proposition 4 *Only anti-coordination games can be strongly informative.*

Proof See Appendix A. □

In the case of 2×2 symmetric games, let us denote the two revealing policy functions by $\sigma^\circ := (I, II)$ and $\sigma^\bullet := (II, I)$. We can then give necessary and sufficient conditions in terms of the utility values (a, b, c, d) for the policy profiles $(\sigma^\circ, \sigma^\circ)$ and $(\sigma^\bullet, \sigma^\bullet)$ to be strict equilibria in anti-coordination games, and hence for a game to be strongly informative under full unmeasurable uncertainty.

Theorem 5 *Suppose that* (a, b, c, d) *is an anti-coordination population game and that* $\lambda \in [0, 1]$. *Then the policy profile* $(\sigma^\circ, \sigma^\circ)$ *defines a strict equilibrium if and only if*

$$a > d \quad and \quad c - a > b - d.$$

The profile $(\sigma^\bullet, \sigma^\bullet)$ *defines a strict equilibrium if and only if*

$$a < d \quad and \quad c - a < b - d.$$

Proof See Appendix A. □

It is also crucial to note that the profile of policy functions $(\sigma^\circ, \sigma^\circ)$ and the dual profile $(\sigma^\bullet, \sigma^\bullet)$ cannot be both strict equilibria of the same game.

Corollary 6 *At most one of the two profiles* $(\sigma^\circ, \sigma^\circ)$ *and* $(\sigma^\bullet, \sigma^\bullet)$ *can be a strict equilibrium.*

Proof See Appendix A. □

3.2 Strong Informativity with Partial Uncertainty

In this section, we find necessary and sufficient conditions for $(\sigma^\circ, \sigma^\circ)$ and $(\sigma^\bullet, \sigma^\bullet)$ to be strict equilibria in situations where the agents' uncertainty is not described by the full unit interval [0, 1], but by some non-empty subinterval $[\underline{\lambda}, \overline{\lambda}] \subseteq [0, 1]$. We derive our results by showing that a game (a, b, c, d) played in this state of partial uncertainty is equivalent to an "inner game" played in a state of full uncertainty. For brevity, we write $\langle (a, b, c, d), [\underline{\lambda}, \overline{\lambda}] \rangle$ for the game with game matrix (a, b, c, d) played in the information state described by the uncertainty interval $[\underline{\lambda}, \overline{\lambda}]$. Notice that each game (a, b, c, d) and uncertainty interval $[\underline{\lambda}, \overline{\lambda}]$, together with a policy profile $(\sigma_i)_{i \in \{1,2\}}$, define an inner game as follows.

Definition 7 For $\langle (a, b, c, d), [\underline{\lambda}, \overline{\lambda}] \rangle$ and policy function $\sigma_{3-i} : \{M, R\} \to \{I, II\}$, the corresponding *inner game* (a', b', c', d') for player i is given by

$$\begin{aligned}
a' &= E_{\underline{\lambda}}[u_i \mid I, M, \sigma_{3-i}] = E_{\underline{\lambda}}[u_i \mid I, R, \sigma_{3-i}] \\
b' &= E_{\overline{\lambda}}[u_i \mid I, M, \sigma_{3-i}] = E_{\overline{\lambda}}[u_i \mid I, R, \sigma_{3-i}] \\
c' &= E_{\underline{\lambda}}[u_i \mid II, M, \sigma_{3-i}] = E_{\underline{\lambda}}[u_i \mid II, R, \sigma_{3-i}] \\
d' &= E_{\overline{\lambda}}[u_i \mid II, M, \sigma_{3-i}] = E_{\overline{\lambda}}[u_i \mid II, R, \sigma_{3-i}]
\end{aligned}$$

Since we are interested in games where players reveal their decision criterion, the relevant inner games are those given by the revealing policy $\sigma^\circ = (I, II)$ and by the revealing policy $\sigma^\bullet = (II, I)$. For $\langle (a, b, c, d), [\underline{\lambda}, \overline{\lambda}] \rangle$ and revealing policy σ° the corresponding inner game $(a^\circ, b^\circ, c^\circ, d^\circ)$ is thus defined by

$$\begin{aligned}
a^\circ &= a + \underline{\lambda}(b - a) \\
b^\circ &= a + \overline{\lambda}(b - a) \\
c^\circ &= c + \underline{\lambda}(d - c) \\
d^\circ &= c + \overline{\lambda}(d - c)
\end{aligned}$$

Similarly, the inner game $(a^\bullet, b^\bullet, c^\bullet, d^\bullet)$ corresponding to $\sigma^\bullet = (II, I)$ is defined by

$$\begin{aligned}
a^\bullet &= b + \overline{\lambda}(a - b) \\
b^\bullet &= b + \underline{\lambda}(a - b) \\
c^\bullet &= d + \overline{\lambda}(c - d) \\
d^\bullet &= d + \underline{\lambda}(c - d)
\end{aligned}$$

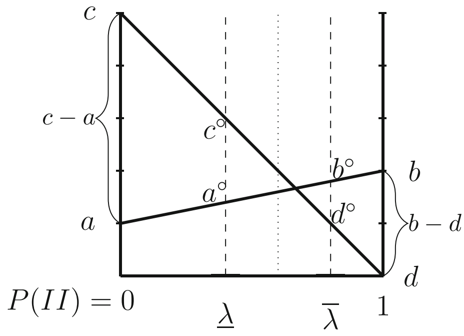

Fig. 1 An inner anti-coordination game

Note that the two inner games are identical when $[\underline{\lambda}, \overline{\lambda}] = [0, 1]$. Figure 1 shows an example of inner game $(a^\circ, b^\circ, c^\circ, d^\circ)$ for $[\underline{\lambda}, \overline{\lambda}] = [0.4, 0.8]$ and $(a = 1, b = 2, c = 5, d = 0)$.

Looking at Fig. 1, it should be clear that only anti-coordination games can give rise to inner games that are also anti-coordination games, but some anti-coordination games give rise to inner games with weakly or strictly dominant actions, depending on where the uncertainty interval $[\underline{\lambda}, \overline{\lambda}]$ is located with respect to the intersection point of the expected-utility lines of the two actions. Intuitively, if the intersection point of the two expected-utility lines in Fig. 1 is not inside the uncertainty interval $[\underline{\lambda}, \overline{\lambda}]$, then one of the two actions dominates the other in the inner game $(a^\circ, b^\circ, c^\circ, d^\circ)$. More precisely, we need the following additional concept.

Definition 8 For an inner game $\langle (a, b, c, d), [\underline{\lambda}, \overline{\lambda}] \rangle$, the *indifference point* associated with the policy $\sigma^\circ = (I, II)$ is the quantity

$$\lambda^\circ = \frac{c - a}{c - a + b - d}.$$

The indifference point associated with $\sigma^\bullet = (II, I)$ is

$$\lambda^\bullet = \frac{b - d}{c - a + b - d}.$$

The following theorem then provides the necessary and sufficient conditions for strong informativity under partial uncertainty.

Theorem 9 *The policy profile* $(\sigma^\circ, \sigma^\circ)$ *is the unique strict equilibrium of* $\langle (a, b, c, d), [\underline{\lambda}, \overline{\lambda}] \rangle$ *if and only if*

$$\underline{\lambda} < \lambda^\circ < \overline{\lambda}, \quad a^\circ > d^\circ \quad \textit{and} \quad c^\circ - a^\circ > b^\circ - d^\circ.$$

The policy profile $(\sigma^\bullet, \sigma^\bullet)$ *is the unique strict equilibrium of* $\langle (a, b, c, d), [\underline{\lambda}, \overline{\lambda}] \rangle$ *if and only if*

$$\underline{\lambda} < \lambda^\bullet < \overline{\lambda}, \quad a^\bullet < d^\bullet \quad \textit{and} \quad c^\bullet - a^\bullet < b^\bullet - d^\bullet.$$

Proof See Appendix A. □

Corollary 10 *Any positive affine transformation of a strongly informative game is still a strongly informative game.*

Proof See Appendix A. □

3.3 Geometric Conditions of Strong Informativity

In the previous subsection, we have seen that the policy profile $(\sigma^\circ, \sigma^\circ)$ is the sole symmetric equilibrium of the anti-coordination game (a, b, c, d) if and only if

$$\underline{\lambda} < \lambda^\circ < \overline{\lambda}, \quad c^\circ - a^\circ > b^\circ - d^\circ \quad \text{and} \quad a^\circ > d^\circ, \tag{2}$$

while $(\sigma^\bullet, \sigma^\bullet)$ is the sole symmetric equilibrium if and only if

$$\underline{\lambda} < \lambda^\bullet < \overline{\lambda}, \quad c^\bullet - a^\bullet < b^\bullet - d^\bullet \quad \text{and} \quad a^\bullet < d^\bullet. \tag{3}$$

Given an uncertainty interval $[\underline{\lambda}, \overline{\lambda}]$, the game (a, b, c, d) is thus informative if it satisfies one of these two mutually exclusive conditions. These two sets of strongly informative games admit a compact and instructive graphical representation. If we set

$$x = \frac{b-a}{c-a+b-d}$$
$$y = \frac{c-a}{c-a+b-d}$$

then the three conditions which ensure that $(\sigma^\circ, \sigma^\circ)$ is the sole equilibrium of (a, b, c, d) can be written as the three linear inequalities

$$\underline{\lambda} < y < \overline{\lambda}, \quad \frac{\underline{\lambda} + \overline{\lambda}}{2} < y \quad \text{and} \quad y < \overline{\lambda} - (\overline{\lambda} - \underline{\lambda})x,$$

each of which is exactly equivalent to one of the original conditions. Similarly, the conditions ensuring that $(\sigma^\bullet, \sigma^\bullet)$ is the sole equilibrium can be written as

$$1 - \overline{\lambda} < y < 1 - \underline{\lambda}, \quad y < \frac{1 - \underline{\lambda} + 1 - \overline{\lambda}}{2} \quad \text{and} \quad y > (1 - \underline{\lambda}) - (\overline{\lambda} - \underline{\lambda})x.$$

Assuming that (a, b, c, d) is an anti-coordination game, the set of informative games for the uncertainty interval $[\underline{\lambda}, \overline{\lambda}]$ can thus be represented as the intersection of three linearly delimited regions in two-dimensional space, as shown in Fig. 2.[3]

4 The Population Multigame

In classic population games, agents from a population are randomly matched to repeatedly play a unique fixed game

$$G^0, G^0, G^0, \ldots.$$

If agents in our population are faced with such a situation they will almost never be able to learn about the actual proportions of decision criteria. We thus extend standard population game models to a framework where at each time $t = 1, 2, 3, \ldots$ members of the population are randomly matched to play a randomly selected game

$$G^1, G^2, G^3, \ldots.$$

[3] Besides allowing for a compact graphical representation, the two variables x and y carry game-theoretic significance too: y is the indifference point, which corresponds to the point where the two lines cross in Fig. 2, while x encodes the slopes of the two lines. For $x \leq 0$, both lines have nonpositive slope; for $x \geq 1$, both lines have nonnegative slope; for $0 < x < 1$, one line has positive slope and one negative.

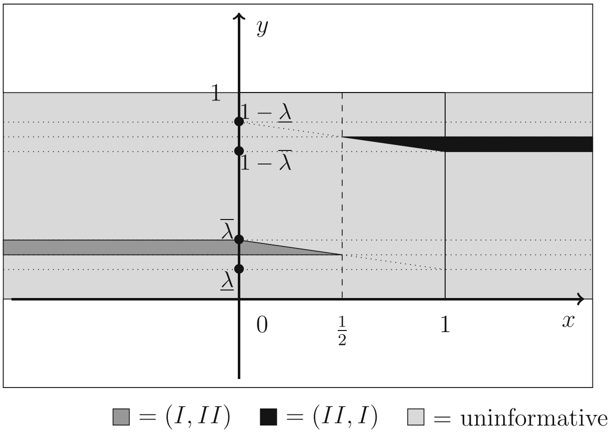

Fig. 2 Strong informativity in anti-coordination games

We call this model the population *multigame*.[4] Moreover, we assume that all the agents in the population are born equally ignorant, i.e., at time $t = 0$ (when no game has been played yet), all players hold full uncertainty about the proportions in the population $[\underline{\lambda}, \overline{\lambda}] = [0, 1]$. We also assume that, at any time $t \geq 0$, the outcomes of the games which have been played are commonly known within the population.

4.1 The Probability of Strong Informativity

When a population is engaged in an infinite sequence of different interactions, the players will be able to refine their uncertainty interval $[\underline{\lambda}, \overline{\lambda}]$ to a single point in the limit only if they will keep receiving informative evidence about the population composition, i.e., only if strongly informative games will keep arising. If we interpret the games in the infinite sequence as a product of the environment where the population lives, it is meaningful to investigate the properties of the environment that ensure positive probability of strongly informative games to occur for any $[\underline{\lambda}, \overline{\lambda}]$ with $\underline{\lambda} < \overline{\lambda}$. In particular, we assume that the game matrix (a, b, c, d) is a random vector sampled from some probability distribution on $\mathbb{R}^4$, which is supposed to encode the properties of the environment. It then makes sense to ask for the probability that the game (a, b, c, d) has some property. In the following, we formulate conditions under which the game is strongly informative with a positive (although possibly small) probability.

Definition 11 We say that a measure P is *absolutely continuous* on a set J if $P(I) > 0$ for any open set $I \subseteq J$.

It would be more precise to say "absolutely continuous with respect to the Lebesgue measure", but since we only need this particular instantiation of the concept, we dispense with formalities.

[4] Note that the concept of random-payoff games has been introduced in game theory and studied from different angles by [5–8, 35]. It is also possible to relate such multigame to the literature on stochastic games [34], with the caveat that the sequence of random games in the multigame does not depend on the actions played by the agents in the games.

Proposition 12 *Suppose that the joint distribution of* (a, b, c, d) *is absolutely continuous on some non-empty hypercube* $(L, H)^4$*. Then there is positive probability that the random game* (a, b, c, d) *is strongly informative with respect to the uncertainty interval* $[0, 1]$.

Proof See Appendix A. □

Example 13 Suppose that a, b, c, d are independently sampled from the uniform distribution on the unit interval. Then the probability that the random game (a, b, c, d) is strongly informative for $(\sigma^\circ, \sigma^\circ)$ with respect to the uncertainty interval $[0, 1]$ is

$$\int_{a=0}^{a=1} \int_{d=0}^{d=a} \int_{c=a}^{c=1} \int_{b=d}^{b=d+c-a} 1 \; dP_a \, dP_d \, dP_c \, dP_b \;=\; \frac{1}{24}.$$

By symmetry, the total probability that (a, b, c, d) is strongly informative with respect to $[0, 1]$ is $1/12$.

The next result connects the positive probability of strong informativity with full uncertainty to the positive probability of strong informativity with partial uncertainty, $[\underline{\lambda}, \overline{\lambda}] \subseteq [0, 1]$.

Proposition 14 *If the joint distribution of* (a, b, c, d) *is absolutely continuous on a non-empty hypercube* $(L, H)^4 \subset \mathbb{R}^4$*, then the distribution of* $(a^\circ, b^\circ, c^\circ, d^\circ)$ *is absolutely continuous on a non-empty hypercube* $(L^\circ, H^\circ)^4 \subseteq (L, H)^4$*, and the distribution of* $(a^\bullet, b^\bullet, c^\bullet, d^\bullet)$ *is absolutely continuous on a non-empty hypercube* $(L^\bullet, H^\bullet)^4 \subseteq (L, H)^4$.

Proof See Appendix A. □

The following theorem then makes use of the last two propositions to prove a general result about the positive probability of strong informativity.

Theorem 15 *Suppose that the joint distribution of* (a, b, c, d) *is absolutely continuous on some non-empty hypercube* $(L, H)^4$*. Then there is positive probability that the random game* (a, b, c, d) *is strongly informative with respect to* $[\underline{\lambda}, \overline{\lambda}]$ *for any* $\underline{\lambda} < \overline{\lambda}$.

Proof See Appendix A. □

Theorem 15 also entails that whenever the four entries a, b, c, d have independent and absolutely continuous distributions on a common interval (L, H), the probability of strongly informative games is positive. For concreteness, in Sect. 5 we give a detailed analysis of the case where a, b, c, d are independent random variables uniformly distributed over a common interval. Before doing that, we take a look at $n \times n$ games and show how our sufficiency results for 2×2 games extend to $n \times n$ games too.

4.2 The $n \times n$ Case

Symmetric $n \times n$ games can be represented as in the following game matrix, where the entry u_{ij} is the row player's utility from playing the i^{th} action against column player's j^{th} action.

	I	II	...	n
I	u_{11}	u_{12}	...	u_{1n}
II	u_{21}	u_{22}	...	...
...	...	...	...	...
n	u_{n1}	...	...	u_{nn}

Theorem 16 *If the joint distribution of* $(u_{ij})_{1\leq i,j\leq n}$ *is absolutely continuous on a non-empty hypercube* $(L, H)^{n\times n} \subset \mathbb{R}^{n\times n}$*, then strongly informative* $n \times n$ *games have positive probability for* $\underline{\lambda} \neq \overline{\lambda}$.

Proof See Appendix A. □

It follows from Theorem 16 that whenever the utility values of the $n \times n$ game matrix are sampled independently from a distribution that is absolutely continuous on a common interval, then strongly informative games have positive probability for any $\underline{\lambda} \neq \overline{\lambda}$.

After having stated sufficiency conditions for strong informativity in 2×2 games in Theorem 15 and in $n \times n$ games in Theorem 16, in the next two sections we compute the probability of strong informativity in some concrete cases, i.e., for some fixed distributions of the utility values. Specifically, in Sect. 5 we analytically compute the probability of strongly informative 2×2 games for utility values that are i.i.d. from the uniform distribution on the unit interval, and we also analyze the speed of learning of the agents in that case. In Sect. 6, we compute the probability of strong informativity by means of computer simulations for games with more than two actions and for utility values between 0 and 1 coming from a range of different distributions.

5 Limit-Learnability in the Uniform Case

In this section, we focus on the case in which a, b, c, d are drawn independently from a uniform distribution on the unit interval. The two constants x and y which played a crucial role in determining whether the game (a, b, c, d) was strongly informative in Sect. 3.3 will then vary randomly, and we will now treat them as random variables X and Y.

In this section, we will compute the conditional distribution of X and Y given that (a, b, c, d) is an anti-coordination game, and use this conditional distribution to approximate the probability of encountering a strongly informative game. Having found this probability, we can make some observations on the interaction between the level of uncertainty and the rate at which new strongly informative games appear, and we can then draw conclusions about the asymptotic rate of decrease in $|\overline{\lambda} - \underline{\lambda}|$ over time.

5.1 The Density of *X* and *Y*

When the entries of the game matrix $G = (a, b, c, d)$ are sampled independently from a uniform distribution on the unit interval $[0, 1]$, the probability density at G is $p(G) = 1$ everywhere inside $[0, 1]^4$ and $p(G) = 0$ outside. The set A of anti-coordination games inside the hypercube is defined by the inequalities

$$0 \leq a < c \leq 1$$
$$0 \leq d < b \leq 1$$

By symmetry and independence of (a, b, c, d), the probability of this set is $P(A) = 1/4$. Hence, the conditional probability density at (a, b, c, d), given that it is an anti-coordination game, is

$$p(a, b, c, d \mid A) = \begin{cases} 4 & (a, b, c, d) \in A \\ 0 & \text{otherwise} \end{cases}$$

We now approximate the conditional probability that $G = (a, b, c, d)$ is strongly informative, given that it is an anti-coordination game. To this purpose, define a coordinate transformation T by

$$X = \frac{b-a}{c-a+b-d} \tag{4}$$
$$Y = \frac{c-a}{c-a+b-d} \tag{5}$$
$$V = \frac{a}{c-a+b-d} \tag{6}$$
$$Z = d \tag{7}$$

Under T, the region $A \subseteq \mathbb{R}^4$ is mapped onto a region $T(A) \subseteq \mathbb{R}^4$ defined by

$$0 \le V < Y + V \le \frac{X+Y+V-1}{Z}$$
$$0 \le X+Y+V-1 < X+V \le \frac{X+Y+V-1}{Z}$$

These joint inequalities on X, Y, V and Z imply the following conditional inequalities: given a fixed value of (X, Y), we have $v^*(x, y) \le V$, where

$$v^*(x, y) = \max\{0, 1 - X - Y\},$$

and given a fixed value for (X, Y, V), we have $0 \le Z \le z^*(x, y, v)$, where

$$z^*(x, y, v) = \min\left\{\frac{X+Y+V-1}{Y+V}, \frac{X+Y+V-1}{X+V}\right\}.$$

If we know the probability density of (X, Y, V, Z), we can use these conditional ranges to integrate out the nuisance variables V and Z. The transformation T is almost everywhere invertible, with the inverse relation T^{-1} given by

$$a = \frac{VZ}{X+Y+V-1}$$
$$b = \frac{Z(X+V)}{X+Y+Z-1}$$
$$c = \frac{Z(Y+V)}{X+Y+V-1}$$
$$d = Z$$

At $(X, Y, V, Z) = (x, y, v, z)$, the Jacobian determinant of T^{-1} satisfies

$$\det(\nabla T^{-1}) = \frac{z^3}{(x+y+v-1)^4}.$$

This determinant measures the amount of volume deformation introduced by the transformation T. Since $p(a, b, c, d \mid A) = 4$ everywhere on A, it follows that the density of (X, Y, V, Z) is given by

$$q(x, y, v, z \mid A) = \begin{cases} 4z^3/(x+y+v-1)^4 & (x, y, v, z) \in T(A) \\ 0 & \text{otherwise.} \end{cases}$$

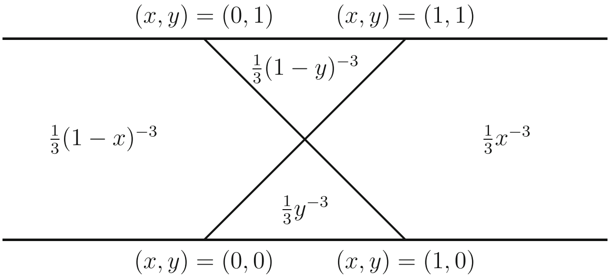

Fig. 3 A map of the four regions of (X, Y)-space on which the conditional probability density function $q(x, y)$ is nonzero. For $y < 0$ or $y > 1$, $q(x, y) = 0$

We can find the probability density function of (X, Y) by integrating out V and Z from this density function, using the ranges computed above:

$$\begin{aligned} q(x, y \mid A) &= \iint_{T(A)} q(x, y, v, z \mid A)\, dz\, dv \\ &= \int_{v=v^*(x,y)}^{v=\infty} \int_{z=0}^{z=z^*(x,y,v)} q(x, y, v, z \mid A)\, dz\, dv. \end{aligned}$$

The value of the upper bound z^* depends on whether $x \leq y$ or $x > y$, while the lower bound v^* depends on whether $x + y \geq 1$ or $x + y < 1$. Hence, this double integral can be split up into four subintegrals covering each of these four cases. We compute each of these four double integrals separately, finding

$$q(x, y \mid A) = \begin{cases} 1/(3x^3) & x > y,\ x + y \geq 1 \\ 1/(3(1-x)^3) & x \leq y,\ x + y < 1 \\ 1/(3y^3) & x > y,\ x + y < 1 \\ 1/(3(1-y)^3) & x \leq y,\ x + y \geq 1 \\ 0 & y < 0 \text{ or } y > 1 \end{cases}$$

Equivalently, this density function can be written as

$$q(x, y \mid A) = \frac{1}{3}\left(\frac{1}{1/2 + \max\{|x - 1/2|, |y - 1/2|\}}\right)^3 \tag{8}$$

where (x, y) lies on the horizontal band defined by $0 < y < 1$. Note that the four different cases are separated by the diagonals of the unit square, as shown in Fig. 3.

5.2 The Probability of Informativity in the Uniform Case

Having now found the exact conditional distribution of (X, Y), we can approximate the conditional probability that the informative policy profile $(\sigma^\circ, \sigma^\circ)$ is the sole symmetric equilibrium of the game. As observed in Sect. 3.3, this event coincides with the set

$$\varphi^\circ = \left\{x < 0,\ \frac{\overline{\lambda} + \underline{\lambda}}{2} < y < \overline{\lambda}\right\} \cup \left\{0 \leq x < \frac{1}{2},\ \frac{\overline{\lambda} + \underline{\lambda}}{2} < y < \overline{\lambda} - (\overline{\lambda} - \underline{\lambda})x\right\}. \tag{9}$$

As previously discussed, a similar set $\varphi^\bullet$ exists for the profile $(\sigma^\bullet, \sigma^\bullet)$. We can then compute the conditional probability of φ° given that (a, b, c, d) is an anti-coordination game as two separate integrals:

$$P(\varphi^\circ \mid A) = \int_{x=-\infty}^{x=0} \int_{y=\frac{\underline{\lambda}+\overline{\lambda}}{2}}^{y=\overline{\lambda}} q(x, y \mid A)\, dy\, dx + \int_{x=0}^{x=1/2} \int_{y=\frac{\underline{\lambda}+\overline{\lambda}}{2}}^{y=\overline{\lambda}-(\overline{\lambda}-\underline{\lambda})x} q(x, y \mid A)\, dy\, dx.$$

On each of these two regions, q behaves differently. Since $0 < y < 1$, when $x \leq 0$ we can reduce the density expression $q(x, y \mid A)$ in Eq. 8 to

$$q(x, y \mid A) = \frac{1}{3}\left(\frac{1}{1-x}\right)^3.$$

The first term of $P(\varphi^\circ \mid A)$ thus has the value

$$P\left\{x < 0,\ \frac{\overline{\lambda}+\underline{\lambda}}{2} < y < \overline{\lambda} \,\middle|\, A\right\} = \int_{x=-\infty}^{x=0} \int_{y=\frac{\underline{\lambda}+\overline{\lambda}}{2}}^{y=\overline{\lambda}} q(x, y \mid A)\, dy\, dx = \frac{\overline{\lambda}-\underline{\lambda}}{12}.$$

For $0 \leq x < 1/2$, we simplify computations by means of the sandwich bounds

$$\frac{1}{3} \leq q(x, y \mid A) \leq \frac{1}{3}\left(\frac{1}{1-x}\right)^3.$$

By integrating all three components, we find that

$$\frac{\overline{\lambda}-\underline{\lambda}}{24} \leq P\left\{0 \leq x < \frac{1}{2},\ \frac{\overline{\lambda}+\underline{\lambda}}{2} < y < \overline{\lambda} - (\overline{\lambda}-\underline{\lambda})x \,\middle|\, A\right\} \leq \frac{\overline{\lambda}-\underline{\lambda}}{12}.$$

Adding up the two terms, we thus have the conditional probability

$$\frac{\overline{\lambda}-\underline{\lambda}}{8} \leq P(\varphi^\circ \mid A) \leq \frac{\overline{\lambda}-\underline{\lambda}}{6}. \tag{10}$$

Since $P(A) = 1/4$, this entails that the unconditional probability satisfies

$$\frac{\overline{\lambda}-\underline{\lambda}}{32} \leq P(\varphi^\circ) \leq \frac{\overline{\lambda}-\underline{\lambda}}{24}. \tag{11}$$

Note that this agrees with the result in Sect. 4.1 for the case of $\overline{\lambda} - \underline{\lambda} = 1$. Finally, the total probability of sampling an informative game is

$$P(\varphi) = P(\varphi^\circ \cup \varphi^\bullet) = P(\varphi^\circ) + P(\varphi^\bullet) = 2P(\varphi^\circ) \tag{12}$$

since the sets φ° and $\varphi^\bullet$ are symmetric and disjoint.

5.3 The Asymptotic Speed of Learning

In the previous subsection, we assumed that the uncertainty interval was fixed and arbitrary and computed some bounds on the probability that a randomly generated game is strongly informative. In this section, we will assume such randomly generated games are played repeatedly between randomly selected pairs of players from a mixed population of maximinimizers and regret minimizers, and that the behavior of the selected players is visible to all members of the population.

Under some reasonable assumptions about the learning rule employed by the members of the population, this causes the probability of observing a new strongly informative game to tend to zero. However, as we also will argue, this convergence is slow enough to allow the members of the population to exactly determine the type proportions in the limit.

Stipulation of learning rule Every time two randomly selected players are plucked from the population and play a strongly informative game, they both reveal their type. Upon observing a sequence of strongly informative games, one can therefore count the number of times the players revealed themselves as maximinimizers or regret minimizers and use this tally to estimate the probability that a randomly drawn player will be of a certain type.

Suppose that k is the number of times the players have revealed themselves to be regret minimizers, and that m is the total number of strongly informative games played. For reasons that we will justify below, we assume that any rational observer of this sequence of events forms the belief that λ, the true proportion of regret minimizers, lies within an uncertainty interval with bounds

$$\underline{\lambda}_m = \frac{k}{m} - \frac{C}{\sqrt{m}}$$
$$\overline{\lambda}_m = \frac{k}{m} + \frac{C}{\sqrt{m}}$$

where $C > 0$ is a fixed but arbitrary constant independent of m. In other words, our assumption is that the width of the uncertainty interval is

$$\overline{\lambda}_m - \underline{\lambda}_m = \frac{2C}{\sqrt{m}} \tag{13}$$

after the observation of m strongly informative games.

The waiting-time distribution Between two strongly informative games, a number of non-informative games is played. After the first m strongly informative games have been played, one needs to play a certain number of games before the next strongly informative game appears. Since the games themselves are generated randomly, this waiting time $\Delta_m \geq 1$ is itself a random variable.

As we have seen in the previous section, the probability that a specific random game is strongly informative depends on the width of the uncertainty interval, but we have assumed that the uncertainty interval remains unchanged as long as no new information is revealed. The process of waiting for the next strongly informative game can therefore be modeled as the process of flipping a bent coin with a fixed success parameter p_m until it comes up heads. This means that the random variable Δ_m follows a geometric distribution with parameter p_m. This distribution has an expected value of

$$E[\Delta_m] = \frac{1}{p_m}$$

and a variance of $(1 - p_m)/p_m^2 \leq 1/p_m^2$.

Bounds on the waiting time As we have seen above, the probability p_m depends on the width of the uncertainty interval and satisfies

$$\frac{\overline{\lambda} - \underline{\lambda}}{16} \leq p_m \leq \frac{\overline{\lambda} - \underline{\lambda}}{12}. \tag{14}$$

With the assumptions above, this is equivalent to

$$\frac{C}{8\sqrt{m}} \leq p_m \leq \frac{C}{6\sqrt{m}}. \tag{15}$$

As we have seen, the expected waiting time before the $(m+1)$th strongly informative game arrives is $E[\Delta_m] = 1/p_m$. We thus have

$$\frac{6\sqrt{m}}{C} \;\leq\; E[\Delta_m] \;\leq\; \frac{8\sqrt{m}}{C} \tag{16}$$

Note also that $Var[\Delta_m] \leq 1/p_m \leq 64m/C$.

Total waiting time Since the waiting time from the mth to the $(m+1)$th strongly informative game is Δ_m, the waiting time until a total of m strongly informative games have been observed is a random variable

$$T \;=\; \Delta_1 + \Delta_1 + \cdots + \Delta_m$$

By the linearity of expectations,

$$\sum_{i=1}^{m} \frac{6\sqrt{i}}{C} \;\leq\; E[T] \;\leq \sum_{i=1}^{m} \frac{8\sqrt{i}}{C}$$

which can be weakened to

$$\frac{3}{C} m^{3/2} \;\leq\; E[T] \;\leq \frac{8}{C} m^{3/2}.$$

Since the waiting times $\Delta_1, \Delta_2, \ldots, \Delta_m$ are independent, we also have

$$Var[T] \;\leq\; \sum_{i=1}^{m} \frac{64i}{C} \;\;\leq\;\; \frac{64}{C} m^2.$$

This variance grows quadratically in m, so $(T - E[T])^2$ will tend to be larger when m is large. In relative terms, however, we have

$$Var\left[\frac{T - E[T]}{E[T]}\right] \;=\; E\left[\frac{Var[T]}{E[T]^2}\right] \;\leq\; \frac{64m^2/C}{(C/(3m^{3/2}))^2} \;=\; \frac{64C}{9m},$$

which goes to zero as $m \to \infty$. Hence $T/E[T]$ will converge to 1 as $m \to \infty$ by, for instance, Chebyshev's inequality.

Limit-learnability We can now briefly summarize in informal terms what we have seen so far. We have shown that the total number of games required in order to observe m strongly informative games is on the order of $T \sim m^{3/2}$. Conversely, when one has observed a total of T games, one can expect the number of strongly informative games among them to be on the order of $m \sim T^{2/3}$. Since we have assumed that the width of the uncertainty interval is inversely proportional to $\sqrt{m}$, it will be on the order of

$$\left|\overline{\lambda}_m - \underline{\lambda}_m\right| \;\sim\; (T^{2/3})^{-1/2} \;=\; T^{-1/3} \tag{17}$$

These results obtain because we have assumed that the agents in the population update their beliefs when new information is revealed, so that the waiting time before the next informative event gradually increases. By contrast, if the agents did not update their beliefs, then the waiting time between strongly informative games would remain constant over time, and m and T would be of the same order of magnitude.

Since $T^{-1/3} \to 0$ for $t \to \infty$, it follows that the players will ultimately learn the proportions of the two agent types in the sense that

$$\left|\overline{\lambda}_m - \underline{\lambda}_m\right| \to 0$$

for $m \to \infty$. This is a somewhat surprising result since it is also the case that $p_m \to 0$ for $m \to \infty$, so that the strongly informative games become more infrequent as the agents become more informed.

Learning rule justification Going back to the stipulated learning rule, we still have to give a justification for the assumption that the width of the uncertainty interval is proportional to $1/\sqrt{m}$ after the observation of m strongly informative games. Our argument for this choice has a positive and a negative aspect. The positive part of the argument shows that it is possible to estimate the parameter of a Bernoulli distribution from m observations with an expected error of $1/\sqrt{m}$, whereas the negative part shows that no substantially lower error is possible.

The positive part of our argument consists of the law of large numbers as formulated, for instance, by Chebyshev:

Theorem 17 *Let $X_1, X_2, X_3, \ldots$ be a series of independent and identically distributed random variables with a shared mean $E[X_i] = \lambda$ and a finite variance, and let*

$$\hat{\lambda}_m = (X_1 + \cdots + X_m)/m$$

be the empirical average of the first m observations. Then for any $\alpha > 0$ there is a constant $\delta > 0$ such that for all $m > 0$,

$$P(|\lambda - \hat{\lambda}_m| < \delta/\sqrt{m}) \ \geq \ 1 - \alpha.$$

This theorem ensures that the mean of a distribution is estimated by the empirical mean of a sample with an accuracy proportional to $1/\sqrt{m}$. A proof can be found in [15], ch. IX and X.

The negative part of our argument relies on much more recent results about the limits on the speed of learning. For the purposes of stating this theorem, let $\lambda_1, \lambda_2 \in (0, 1)$ be the parameters of two coin flipping distributions, and let m observations be drawn from one of these two distributions. A hypothesis test is then a function that maps a data set to one of the two parameter values. We say that the two distributions are (m, α)-distinguishable if there is a hypothesis test that returns the correct parameter value with probability $1 - \alpha$ for a data set of size m. We then have:

Theorem 18 *There are $\delta > 0$ and $\alpha > 0$ such that for all m, if two parameters $\lambda_1, \lambda_2 \in (0, 1)$ satisfy the proximity condition $|\lambda_1 - \lambda_2| < \delta/\sqrt{m}$, then the corresponding coin flipping distributions are not (m, α)-distinguishable.*

This theorem allows us to conclude that the confidence bounds provided by Chebyshev's inequality are the best we can hope for: the width of the confidence interval cannot shrink at a rate faster than $1/\sqrt{m}$. This result can be proven by using the Hellinger distance between two binomial distributions to lower-bound the minimax risk for the hypothesis test, as discussed extensively elsewhere [36, 37].

Together, these two results show that a confidence interval of the shape

$$\underline{\lambda}_m = \hat{\lambda}_m - \frac{C}{\sqrt{m}} \tag{18}$$

$$\overline{\lambda}_m = \hat{\lambda}_m + \frac{C}{\sqrt{m}} \tag{19}$$

will contain the true parameter λ with a probability that neither converges to 0 nor to 1 as $m \to \infty$. By contrast, if

$$\frac{|\overline{\lambda}_m - \underline{\lambda}_m|}{\sqrt{m}} \to \infty$$

for $m \to \infty$, the probability of error would tend to 0, and if we had

$$\frac{|\overline{\lambda}_m - \underline{\lambda}_m|}{\sqrt{m}} \to 0$$

for $m \to \infty$, the probability of error would tend to 1. This hence justifies our assumption that any rational agent must use confidence intervals of width proportional to $1/\sqrt{m}$ when estimating the parameter of a Bernoulli distribution from m observations.

6 Strong Informativity for Beta and Dirichlet Distributions

In this section, we expand the analysis from the previous section to games with more than two actions and different utility-matrix distributions. We empirically estimate the probability of encountering a strongly informative game both in the case where each cell in the utility matrix is sampled independently from the beta distribution and in the case where the entire utility matrix is a sample from a Dirichlet distribution (and hence utility values are no longer independent).

The beta distribution The beta distribution is a distribution over the unit interval parameterized by two parameters α and β. We focus exclusively on the case where the two parameters are identical, $\alpha = \beta = r$. In our first set of experiments, we use such beta distributions to sample each value of the utility matrix independently.

The probability density functions of the beta distribution for some of these parameters are shown in Fig. 4. As the figure illustrates, the beta distribution is identical to the uniform distribution when $\alpha = \beta = 1$, while it comes to resemble a Bernoulli distribution for $\alpha, \beta \to 0$ and an increasingly narrow normal distribution as $\alpha, \beta \to \infty$.

The Dirichlet distribution The Dirichlet distribution of order D is a probability distribution over the $(D-1)$-dimensional probability simplex (i.e., over $(D-1)$-vectors v with $0 \leq v_d \leq 1$ and $\sum_d v_d = 1$). The Dirichlet distribution of order D is parameterized by a vector of D positive parameters $\alpha_1, ..., \alpha_D$.

We once again focus on the symmetric parameter vectors, $\alpha_d = r$ for all d. Dirichlet distributions with such parameter vectors are equal to the uniform distribution when $r = 1$. They become increasingly concentrated around the corners of the simplex as $r \to 0$ and increasingly concentrated around the center point of the simplex for $r \to \infty$.

Beta-distribution experiments In our first set of experiments, we consider symmetric $n \times n$ game matrices whose values are drawn from symmetric beta distributions. We can thus empirically estimate the probability that such randomly generated games are strongly informative.

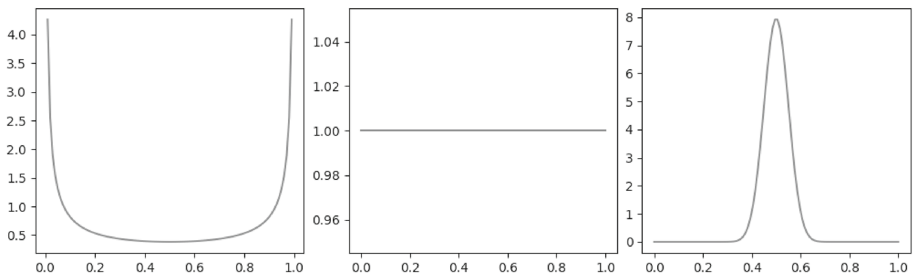

Fig. 4 Examples of beta distributions. Left: $\alpha = \beta = 0.25$. Center: $\alpha = \beta = 1$. Right: $\alpha = \beta = 50$

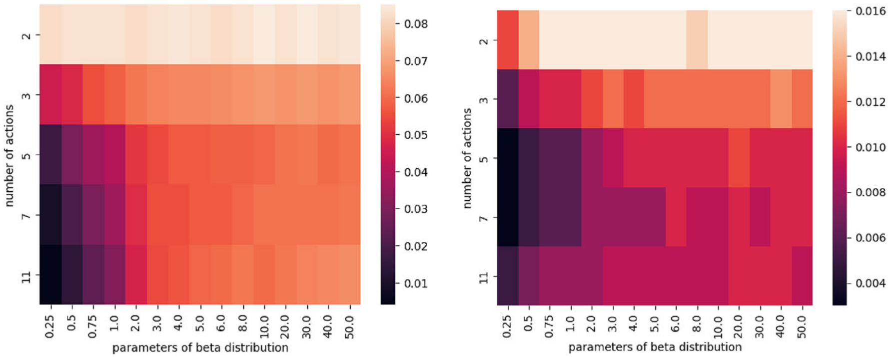

Fig. 5 Frequencies of strongly informative games for independently beta-distributed utility values in the case of maximum uncertainty $[\underline{\lambda}, \overline{\lambda}] = [0, 1]$ on the left, and in the case of reduced uncertainty $[\underline{\lambda}, \overline{\lambda}] = [0.6, 0.8]$ on the right

We set the number of actions equal to

$$n = 2, 3, 5, 7, 11$$

and the two (identical) beta parameters equal to

$$r = 0.25, 0.5, 0.75, 1, 2, 3, 4, 5, 6, 8, 10, 20, 30, 40, 50.$$

We additionally try two different uncertainty intervals, $[\underline{\lambda}, \overline{\lambda}] = [0, 1]$ and $[\underline{\lambda}, \overline{\lambda}] = [0.6, 0.8]$. Figure 5 tabulates the estimated probability that a random game will be informative for every possible combination of these parameter choices. The estimated probabilities are based on a sample of 10^5 games.

A few observations are in order. First, the observed frequency of strongly informative games for the case of maximum uncertainty $[\underline{\lambda}, \overline{\lambda}] = [0, 1]$ with two actions and beta parameters $\alpha = \beta = 1$ is precisely 0.08296, which is within the theoretical bounds found in Sect. 5, $0.0625 \leq 0.08296 \leq 0.0833$. This is also true for the case of reduced uncertainty $[\underline{\lambda}, \overline{\lambda}] = [0.6, 0.8]$, where we have $0.0125 \leq 0.0159 \leq 0.0166$.

Second, both the case of maximum uncertainty and the case of reduced uncertainty display a similar pattern, with lower frequency of strongly informative games in the bottom left corner, as the number of actions increases and the beta parameters decrease. Overall, 2×2 games turn out to be the most informative for all the chosen parameters of the beta distribution.

Figure 6 illustrates the effect of shrinking the uncertainty interval in two cases, one in which the number of actions is held fixed at $n = 2$, and one in which the beta parameters are held fixed at $\alpha = \beta = 1$. As expected, both tables show that the probability of encountering a strongly informative game tends to 0 as the uncertainty interval shrinks. Perhaps more interestingly, the right-hand inset shows that the probability of encountering a strongly informative game also depends negatively on the number of actions available to the players.

Dirichlet-distribution experiments We now turn to the case of Dirichlet-distributed utility values. As mentioned above, we now sample the entire $n \times n$ symmetric game matrix from a Dirichlet distribution of order $D = n^2$ with identical parameters $\alpha_1 = \cdots = \alpha_D = r$. We again choose the following number of actions

$$n = 2, 3, 5, 7, 11$$

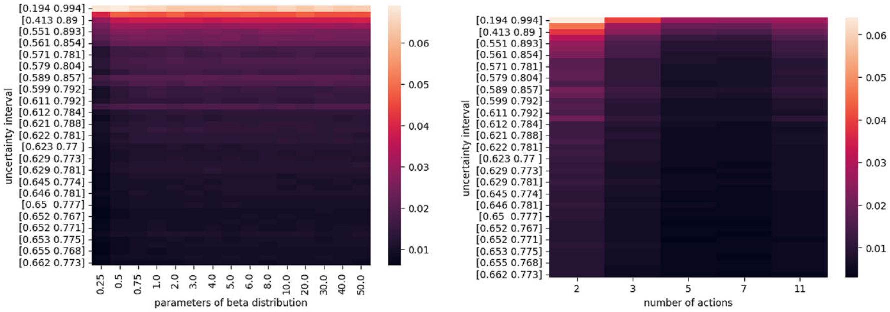

Fig. 6 Frequencies of strongly informative games during learning. On the left: frequencies for different parameters of the beta distribution during learning in 2×2 games. On the right: frequencies for different numbers of actions during learning for beta parameters $\alpha = \beta = 1$

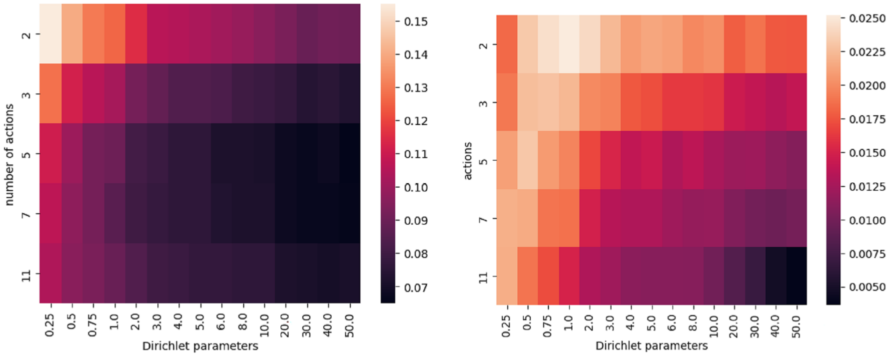

Fig. 7 Frequencies of strongly informative games for Dirichlet-distributed utility values in the case of maximum uncertainty $[\underline{\lambda}, \overline{\lambda}] = [0, 1]$ on the left, and in the case of reduced uncertainty $[\underline{\lambda}, \overline{\lambda}] = [0.6, 0.8]$ on the right

in all combinations with the parameter values

$$r = 0.25, 0.5, 0.75, 1, 2, 3, 4, 5, 6, 8, 10, 20, 30, 40, 50$$

and in all combinations with two choices of uncertainty interval. The results are shown in Fig. 7.

There are some noticeable differences between Figs. 5 and 7. First, the highest frequencies of strongly informative games are roughly twice as high when utility values are Dirichlet-distributed as when utility values are beta-distributed. This holds both in the case of maximum uncertainty (0.15 vs 0.08) and in the case of reduced uncertainty (0.016 vs 0.025). Second, the frequency of strongly informative games for 2×2 games is nearly independent of the distribution parameters when the utility values are sampled independently from the beta distribution, whereas it depends heavily on the distribution parameters when the game matrix is sampled in its entirety from a Dirichlet distribution.

Figure 8 instead shows the frequencies of strongly informative games as the uncertainty interval shrinks. Except for the differences just observed, the two graphs of Fig. 8 look similar to those of Fig. 6.

Learning dynamics The probability of encountering a strongly informative games decreases as the uncertainty interval shrinks. As we have seen in Sect. 5, this has the consequence that

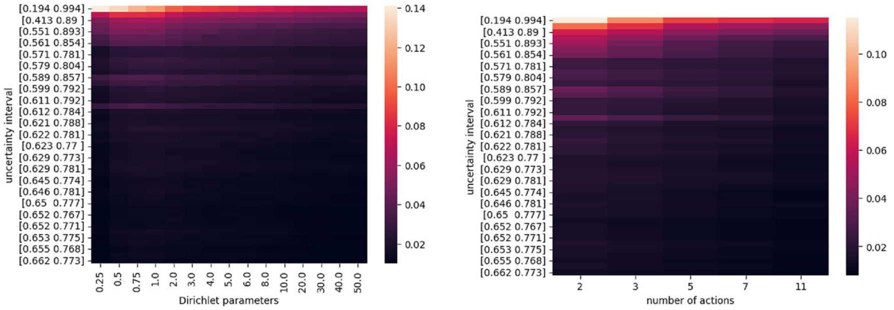

Fig. 8 Frequencies of strongly informative games during learning. On the left: frequencies for different parameters of the Dirichlet distribution during learning for two actions. On the right: frequencies for different numbers of actions during learning for Dirichlet parameters $(\alpha_1, ..., \alpha_D) = (1, ..., 1)$

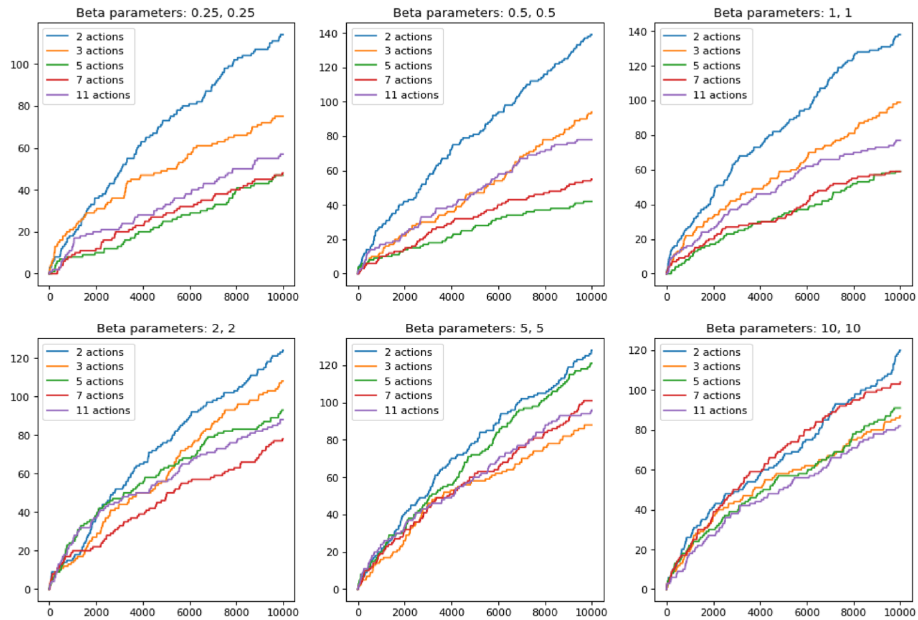

Fig. 9 Cumulative number of strongly informative games (y-axis) over 10000 games randomly generated by drawing i.i.d. utility values from a symmetric beta distribution (x-axis). The number of games between two subsequent steps in the curves corresponds to the waiting time between a strongly informative game and the next strongly informative game

the total number of strongly informative games over time grows slower than linearly when the agents learn as they play. Figures 9 and 10 illustrate this effect by plotting the cumulative number of strongly informative games under different distributional assumptions.

All the curves plotted are concave, illustrating the fact that the waiting time before the next strongly informative game increases as the total number of strongly informative games increases. In particular, the growth rate for the 2×2 games with uniformly distributed utility values ($\alpha = \beta = 1$) is consistent with a growth rate of $T^{2/3}$ strongly informative games after a total of T games have been played (Fig. 9, top right inset).

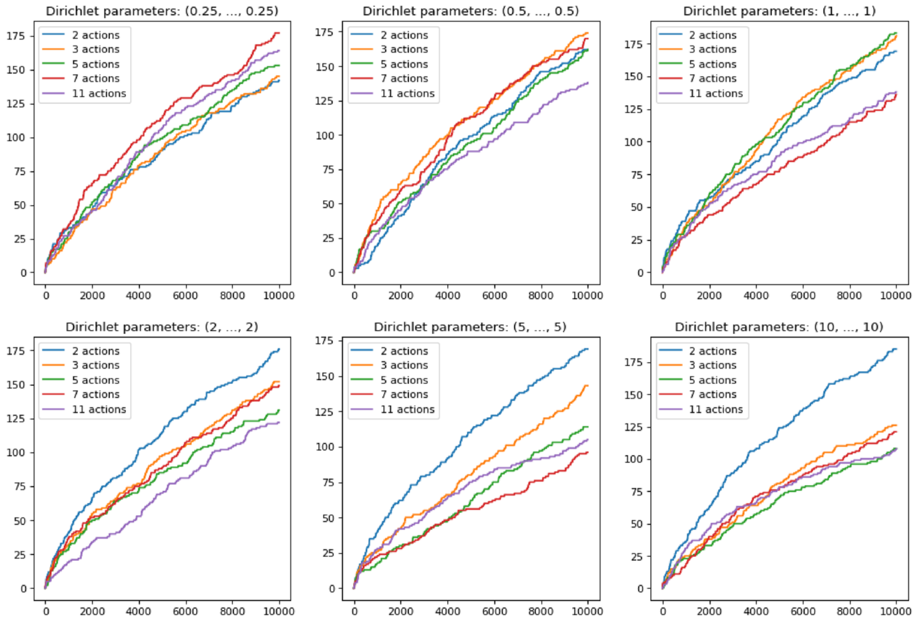

Fig. 10 Cumulative number of strongly informative games (y-axis) over 10000 games randomly generated by drawing i.i.d. matrices of utility values from a Dirichlet distribution (x-axis). The number of games between two subsequent steps in the curves corresponds to the waiting time between a strongly informative game and the next strongly informative game

Figure 9 also shows that random 2×2 games are more likely to be strongly informative than games with $n > 2$ when the utilities are drawn independently from a beta distribution. This contrast is strongest when α and β are closer to 0 and is barely detectable when $\alpha = \beta = 10$.

As Fig. 10 shows, random 2×2 games are also more likely to be strongly informative in the Dirichlet-distributed case when the parameters $\alpha_1 = \cdots = \alpha_D$ are larger. When the parameters $\alpha_1 = \cdots = \alpha_D$ are close to 0 instead, the informativity with $n = 2$ is lower than the informativity with $n > 2$.

7 Discussion

One of the most crucial steps in the development of modern psychology was the realization that an exclusive focus on expressed behavior had started to weigh down the discipline: in order to explain the behaviors, the concept of a private and unobservable mental process would have to be reintroduced. Recent works in biology and ethology have started to suggest that the same might hold for animal behavior in general [14, 21, 23]. In evolutionary game theory, ideas of this kind have given rise to studies on the evolution of preferences, which elevate the level of explanation from expressed behaviors to subjective utilities (e.g., [1, 9, 11]).

The present paper too can be read as an attempt to explain observable behaviors in terms of more general processes, in this case different rules for decision making under uncertainty. Rather than considering various decision criteria as competing philosophical theories, they can be interpreted as high-level strategies that may coexist or play off against each other.

From this point of view, the multigame model introduced here allows to study the various criteria of rational choice from a population standpoint, as advocated in [23] and [21].

The focus here was on learning: we have investigated a population consisting of distinct subpopulations using different decision criteria, asking whether the members of that population could ultimately come to learn its composition just from observing the actions of their peers. Since the only source of ambiguity was the population composition, and the difference between the decision criteria vanishes when ambiguity does, it was not immediately obvious that this could ever be the case. As it turned out, such limit-learnability is guaranteed in a large class of important cases.

The general research question thus pertained to the relationship between directly observable and not directly observable features of the population. The methods employed here, however, are very general in nature and could also be used to address many other issues surrounding decision making and rational choice in the presence of uncertainty: for instance, research addressing the merits of different criteria from an evolutionary rather than philosophical perspective could prove profitable to the theory of rational choice. In allowing the various decision criteria to coexist and interact in the model, we may come to better understand the advantages and disadvantages of each of them.

A Proofs

A.1 Proofs of Sect. 3

A.1.1 Proof of Proposition 4

Proof When $a_i \in br(a_i)$ for some action a_i, then all the types of all the players playing a_i is an equilibrium of the game. Both coordination games and mixed games therefore cannot be strongly informative. □

A.1.2 Proof of Theorem 5

Proof To prove the first part of this theorem, we show that σ° is a strictly best reply to σ° by considering player i's two types separately: we show that action I is strictly better for type M if and only if $a > d$, and that action II is strictly better for type R if and only if $c - a > b - d$.

For $i \in \{1, 2\}$, the expected utilities for a maximinimizing type from actions I and II when faced with the policy $\sigma^\circ = (I, II)$ are

$$E_\lambda[u_i \mid I, M, \sigma^\circ] = (1-\lambda)a + \lambda b$$
$$E_\lambda[u_i \mid II, M, \sigma^\circ] = (1-\lambda)c + \lambda d$$

Action I is thus a best reply for the maximinimizer if and only if

$$\min_{\lambda\in[0,1]} \{(1-\lambda)a + \lambda b\} > \min_{\lambda\in[0,1]} \{(1-\lambda)c + \lambda d\}.$$

Since the minimum of a linear function over a compact interval is attained at one of the endpoints, this is equivalent to

$$\min\{a, b\} > \min\{c, d\}.$$

Since we assume that (a, b, c, d) is an anti-coordination game, we know that $a < c$ and $d < b$. This implies that:

- $a < d$ if and only if $\min\{a, b\} < \min\{c, d\}$;
- $a > d$ if and only if $\min\{a, b\} > \min\{c, d\}$;
- $a = d$ if and only if $\min\{a, b\} = \min\{c, d\}$.

Hence, the maximinimizing type strictly prefers action I if and only if $a > d$.

Given a fixed $\lambda \in [0, 1]$, the regrets of player i against the policy $\sigma^\circ = (I, II)$ are

$$\max_{a_i} E_\lambda[u_i \mid a_i, R, \sigma^\circ] - E_\lambda[u_i \mid I, R, \sigma^\circ] = \max\{0, (1-\lambda)(c-a) + \lambda(d-b)\}$$
$$\max_{a_i} E_\lambda[u_i \mid a_i, R, \sigma^\circ] - E_\lambda[u_i \mid II, R, \sigma^\circ] = \max\{0, (1-\lambda)(a-c) + \lambda(b-d)\}$$

When maximizing these regrets over $\lambda \in [0, 1]$, we obtain

$$\max_{\lambda \in [0,1]} \{\max_{a_i} E_\lambda[u_i \mid a_i, R, \sigma^\circ] - E_\lambda[u_i \mid I, R, \sigma^\circ]\} = \max\{c-a, d-b\}$$
$$\max_{\lambda \in [0,1]} \{\max_{a_i} E_\lambda[u_i \mid a_i, R, \sigma^\circ] - E_\lambda[u_i \mid II, R, \sigma^\circ]\} = \max\{a-c, b-d\}$$

Therefore, the regret minimizer strictly prefers II over I if and only if

$$\max\{c-a, d-b\} > \max\{a-c, b-d\}.$$

Assuming that (a, b, c, d) is an anti-coordination game, the regret minimizer has a strict preference for action II if and only if

$$c - a > b - d.$$

The second half of the theorem for $\sigma^\bullet$ is proven analogously. □

A.1.3 Proof of Corollary 6

Proof The conditions

$$a > d \quad \text{and} \quad c - a > b - d$$

for $(\sigma^\circ, \sigma^\circ)$ to be an equilibrium are trivially incompatible with the conditions

$$a < d \quad \text{and} \quad c - a < b - d$$

for $(\sigma^\bullet, \sigma^\bullet)$ to be an equilibrium. Any anti-coordination game (a, b, c, d) can thus satisfy at most one of the two pairs of inequalities. □

A.1.4 Proof of Theorem 9

To prove Theorem 9, we need to prove some auxiliary propositions first.

Proposition 19 *A policy σ_i is a best reply to $\sigma^\circ = (I, II)$ in $\langle (a, b, c, d), [\underline{\lambda}, \overline{\lambda}] \rangle$ if and only if it is a best reply to σ° in $\langle (a^\circ, b^\circ, c^\circ, d^\circ), [0, 1] \rangle$. Similarly, a policy σ_i is a best reply to $\sigma^\bullet = (I, II)$ in $\langle (a, b, c, d), [\underline{\lambda}, \overline{\lambda}] \rangle$ if and only if it is a best reply to $\sigma^\bullet$ in $\langle (a^\bullet, b^\bullet, c^\bullet, d^\bullet), [0, 1] \rangle$.*

Proof Suppose player $3-i$ uses the policy $\sigma^\circ = (I, II)$. For type M of player i, the expected utilities associated with actions I and II are then, for a fixed value of λ,

$$E_\lambda[u_i \mid I, M, \sigma^\circ] = (1-\lambda)a + \lambda b \;=\; a + \lambda(b-a)$$
$$E_\lambda[u_i \mid II, M, \sigma^\circ] = (1-\lambda)c + \lambda d \;=\; c + \lambda(d-c)$$

These two are functions of λ, and the maximinimizing type of player i prefers the action with the highest minimum. Action I is thus a strictly best reply for this type if and only if

$$\min_{\lambda\in[\underline{\lambda},\overline{\lambda}]} \{a + \lambda(b-a)\} \;>\; \min_{\lambda\in[\underline{\lambda},\overline{\lambda}]} \{c + \lambda(d-c)\}\,.$$

Given that linear functions on compact intervals reach their minima at the endpoints, this is equivalent to

$$\min\left\{a + \underline{\lambda}(b-a), a + \overline{\lambda}(b-a)\right\} \;>\; \min\left\{c + \underline{\lambda}(d-c), c + \overline{\lambda}(d-c)\right\},$$

that is,

$$\min\left\{a^\circ, b^\circ\right\} \;>\; \min\left\{c^\circ, d^\circ\right\}.$$

By Theorem 5, this inequality is satisfied if and only if action I is the maximinimizer's strictly best reply to σ° in $\langle(a^\circ, b^\circ, c^\circ, d^\circ), [0, 1]\rangle$. By inverting the inequalities, we similarly find that action II is the strictly best reply to σ° for the maximinimizer in $\left\langle(a, b, c, d), [\underline{\lambda}, \overline{\lambda}]\right\rangle$ if

$$\min\left\{a^\circ, b^\circ\right\} \;<\; \min\left\{c^\circ, d^\circ\right\}.$$

Now consider the regret-minimizing type of player i, faced with an opponent using the policy σ°. For a fixed value of λ, the regrets associated with actions I and II are then

$$\max_{a_i} E_\lambda[u_i \mid a_i, R, \sigma^\circ] - E_\lambda[u_i \mid I, R, \sigma^\circ] = \max\{0, (1-\lambda)(c-a) + \lambda(d-b)\}$$
$$\max_{a_i} E_\lambda[u_i \mid a_i, R, \sigma^\circ] - E_\lambda[u_i \mid II, R, \sigma^\circ] = \max\{0, (1-\lambda)(a-c) + \lambda(b-d)\}$$

The regret-minimizing type of player i prefers whichever of these expression has the lowest maximum. Action II is thus a strictly best reply if and only if

$$\max_{\lambda\in[\underline{\lambda},\overline{\lambda}]} \{0, (1-\lambda)(a-c) + \lambda(b-d)\} \;<\; \max_{\lambda\in[\underline{\lambda},\overline{\lambda}]} \{0, (1-\lambda)(c-a) + \lambda(d-b)\}\,.$$

This inequality has the form

$$\max_{\lambda\in[\underline{\lambda},\overline{\lambda}]} \{0, f(\lambda)\} \;<\; \max_{\lambda\in[\underline{\lambda},\overline{\lambda}]} \{0, -f(\lambda)\}$$

where f is the function

$$f(\lambda) \;=\; (1-\lambda)(a-c) + \lambda(b-d).$$

Since either $f(\lambda)$ or $-f(\lambda)$ is nonpositive, the inequality is equivalent to

$$\max_{\lambda\in[\underline{\lambda},\overline{\lambda}]} \{f(\lambda)\} \;<\; \max_{\lambda\in[\underline{\lambda},\overline{\lambda}]} \{-f(\lambda)\}.$$

The function $f(\lambda)$ is linear in λ and ranges between the values

$$f(\underline{\lambda}) = (1-\underline{\lambda})(a-c) + \underline{\lambda}(b-d) = a^\circ - c^\circ$$
$$f(\overline{\lambda}) = (1-\overline{\lambda})(a-c) + \overline{\lambda}(b-d) = b^\circ - d^\circ$$

and of course inversely for $-f$. Maximizing over these pairs of values, we thus find that action II is a strictly best reply to σ° if and only if

$$\max\left\{a^\circ - c^\circ, b^\circ - d^\circ\right\} \;<\; \max\left\{c^\circ - a^\circ, d^\circ - b^\circ\right\}.$$

By Theorem 5 this is satisfied if and only if action II is the regret minimizer's strictly best reply to σ° in $\langle(a^\circ, b^\circ, c^\circ, d^\circ), [0, 1]\rangle$. Inverting the inequalities, we similarly find that action I is a strictly best reply if and only if

$$\max\left\{a^\circ - c^\circ, b^\circ - d^\circ\right\} \;>\; \max\left\{c^\circ - a^\circ, d^\circ - b^\circ\right\}.$$

Replacing σ° with $\sigma^\bullet$, the second part of the theorem is proven analogously. □

Proposition 20 *For the policy profile $(\sigma^\circ, \sigma^\circ)$ to be the unique strict equilibrium of the population game with ambiguity $\left\langle(a, b, c, d), [\underline{\lambda}, \overline{\lambda}]\right\rangle$, it is necessary that $(a^\circ, b^\circ, c^\circ, d^\circ)$ be an anti-coordination game, i.e., $c^\circ > a^\circ$ and $b^\circ > d^\circ$. Similarly, for $(\sigma^\bullet, \sigma^\bullet)$ to be the unique strict equilibrium of $\left\langle(a, b, c, d), [\underline{\lambda}, \overline{\lambda}]\right\rangle$, $(a^\bullet, b^\bullet, c^\bullet, d^\bullet)$ must be an anti-coordination game.*

Proof Consider policy σ° and the corresponding inner game $(a^\circ, b^\circ, c^\circ, d^\circ)$. If this inner game has a weakly or strictly dominant action a_i^* such that, for all $\lambda \in [\underline{\lambda}, \overline{\lambda}]$, and all $a_i \in A_i$,

$$E_\lambda[u_i \mid a_i^*, t_i, \sigma^\circ] \;\geq\; E_\lambda[u_i \mid a_i, t_i, \sigma^\circ],$$

then a_i^* is the best reply to policy σ° under both decision criteria $t_i \in \{M, R\}$ in the inner game. Hence, $(\sigma^\circ, \sigma^\circ)$ is not an equilibrium. If $(a^\circ, b^\circ, c^\circ, d^\circ)$ is a coordination game, then (a, b, c, d) too is a coordination game. Hence, the policy profiles $((I, I), (I, I))$ and $((II, II), (II, II))$ are both symmetric equilibria regardless of the uncertainty interval. Lastly, if in the game $(a^\circ, b^\circ, c^\circ, d^\circ)$ all actions are equivalent (i.e., $a^\circ = c^\circ$ and $b^\circ = d^\circ$), then any profile is an equilibrium, regardless of the uncertainty interval. The same is true also if we consider the other revealing policy function $\sigma^\bullet$. It follows that unless the inner game is an anti-coordination game, neither $(\sigma^\circ, \sigma^\circ)$ nor $(\sigma^\bullet, \sigma^\bullet)$ can be the unique strict equilibrium. □

Proposition 21 *Suppose that (a, b, c, d) is an anti-coordination game. Then the inner game $(a^\circ, b^\circ, c^\circ, d^\circ)$ is an anti-coordination game (i.e., $c^\circ > a^\circ$ and $b^\circ > d^\circ$) if and only if*

$$\underline{\lambda} \;<\; \lambda^\circ \;<\; \overline{\lambda}.$$

The inner game $(a^\bullet, b^\bullet, c^\bullet, d^\bullet)$ is an anti-coordination game (i.e., $c^\bullet > a^\bullet$ and $b^\bullet > d^\bullet$) if and only if

$$\underline{\lambda} \;<\; \lambda^\bullet \;<\; \overline{\lambda}.$$

Proof By expanding the inequality $\underline{\lambda} < \lambda^\circ$ we get

$$\underline{\lambda} < \frac{c - a}{c - a + b - d},$$

which simplifies to

$$a + \underline{\lambda}(b - a) < c + \underline{\lambda}(d - c),$$

that is,

$$a^\circ < c^\circ.$$

Similarly, by expanding and simplifying the inequality $\lambda^\circ < \overline{\lambda}$ we obtain

$$a + \overline{\lambda}(b - a) > c + \overline{\lambda}(d - c),$$

that is,

$$b^\circ > d^\circ.$$

Substituting σ° with $\sigma^\bullet$, the second part of the proof proceeds analogously. □

Finally, we are in the position to prove Theorem 9.

Proof By Proposition 21, the game $(a^\circ, b^\circ, c^\circ, d^\circ)$ is an anti-coordination game if and only if $\underline{\lambda} < \lambda^\circ < \overline{\lambda}$, and the game $(a^\bullet, b^\bullet, c^\bullet, d^\bullet)$ is an anti-coordination game if and only if $\underline{\lambda} < \lambda^\bullet < \overline{\lambda}$. By Proposition 19, the policy profile $(\sigma^\circ, \sigma^\circ)$ is an equilibrium of the game $\big\langle (a, b, c, d), [\underline{\lambda}, \overline{\lambda}]\big\rangle$ if and only if it is an equilibrium of the game $\langle (a^\circ, b^\circ, c^\circ, d^\circ), [0, 1]\rangle$, and $(\sigma^\bullet, \sigma^\bullet)$ is an equilibrium of $\big\langle (a, b, c, d), [\underline{\lambda}, \overline{\lambda}]\big\rangle$ if and only if it is an equilibrium of $\langle (a^\bullet, b^\bullet, c^\bullet, d^\bullet), [0, 1]\rangle$. By Theorem 5, these two equilibrium conditions reduce to the two conjunctions ($a^\circ > d^\circ$ and $c^\circ - a^\circ > b^\circ - d^\circ$), and ($a^\bullet < d^\bullet$ and $c^\bullet - a^\bullet < b^\bullet - d^\bullet$), respectively. Finally, at most one of the two profiles $(\sigma^\circ, \sigma^\circ)$ and $(\sigma^\bullet, \sigma^\bullet)$ can be a strict equilibrium by Corollary 6. □

A.1.5 Proof of Corollary 10

Proof Suppose the game (a, b, c, d) is strongly informative for the interval $[\underline{\lambda}, \overline{\lambda}]$. Any transformation $T : \mathbb{R}^4 \to \mathbb{R}^4$ such that

$$(a, b, c, d) \;\mapsto\; (sa + r, sb + r', sc + r, sd + r') \quad r, r', s \in \mathbb{R}, s > 0$$

preserves the quantities λ° and $\lambda^\bullet$ as well as all the relevant inequalities involving the variables $a^\circ, b^\circ, c^\circ, d^\circ$ and $a^\bullet, b^\bullet, c^\bullet, d^\bullet$. □

A.2 Proofs of Sect. 4

A.2.1 Proof of Proposition 12

Proof By the assumption of absolute continuity it holds that

$$P(L < a < H) > 0$$
$$P(L < d < a \,|\, a) > 0$$
$$P(a < c < H \,|\, a) > 0$$

Note moreover that $d < a$ and $c < H$ collectively imply that $d + c - a < H$, while $a < c$ implies that $d < d + c - a$. It follows that $(d, d + c - a) \subseteq (L, H)$, and that

$$P(d < b < d + c - a \,|\, a, c, d) \;>\; 0.$$

Hence, the conjunction of these four events has positive probability. By Theorem 5, the second, third, and fourth of these events also ensures that $(\sigma^\circ, \sigma^\circ)$ is the only symmetric equilibrium of the game, and therefore that the game is strongly informative with respect to the uncertainty interval [0, 1]. (A similar argument can be made for $(\sigma^\bullet, \sigma^\bullet)$.) □

A.2.2 Proof of Proposition 14

Proof We give the proof for the case of $(a^\circ, b^\circ, c^\circ, d^\circ)$. The case of $(a^\bullet, b^\bullet, c^\bullet, d^\bullet)$ is analogous. The inner game $(a^\circ, b^\circ, c^\circ, d^\circ)$ consists of the two rows, (a°, b°) and (c°, d°), which are linear transformations of the corresponding outer rows (a, b) and (c, d), respectively. The linear transformation can be represented by the matrix

$$T = \begin{pmatrix} 1 - \underline{\lambda} & \underline{\lambda} \\ 1 - \overline{\lambda} & \overline{\lambda} \end{pmatrix}.$$

By construction, T has the same effect on both outer rows, and the distribution of both rows is absolutely continuous over $(L, H)^2$ by assumption. It thus follows that if we can inscribe a square $(L^\circ, H^\circ)^2$ in the image of $(L, H)^2$ under T, then the hypercube $(L^\circ, H^\circ)^2 \times (L^\circ, H^\circ)^2$ will be contained in the image of $(L, H)^2 \times (L, H)^2$ under T. We want to argue that there is a square $(L^\circ, H^\circ)^2$ which is contained in the image of $(L, H)^2$.

The mapping defined by T leaves two of the corners of $(L, H)^2$ unchanged,

$$T \begin{pmatrix} L \\ L \end{pmatrix} = \begin{pmatrix} L \\ L \end{pmatrix}, \qquad T \begin{pmatrix} H \\ H \end{pmatrix} = \begin{pmatrix} H \\ H \end{pmatrix},$$

and it leaves all points on the main diagonal unchanged as well. Moreover, T maps the square $(L, H)^2$ onto a parallelogram of area

$$\det(T) \cdot (H - L)^2 = (\overline{\lambda} - \underline{\lambda}) \cdot (H - L)^2 > 0.$$

This parallelogram degenerates to a line on the main diagonal of the square when $\overline{\lambda} - \underline{\lambda} = 0$, but it otherwise has positive area. Putting these two observations together, we thus conclude that all points on the diagonal of $(L, H)^2$ are inner points of the parallelogram $T((L, H)^2)$ when $\overline{\lambda} - \underline{\lambda} > 0$. Hence, it is possible to inscribe an open set and therefore a small square $(L^\circ, H^\circ)^2$ into the parallelogram $T((L, H)^2)$ along the diagonal of $(L, H)^2$, as shown in Fig. 11.

It thus follows that if the distribution of (a, b, c, d) is absolutely continuous on a non-empty hypercube $(L, H)^4$, then the distribution of $(a^\circ, b^\circ, c^\circ, d^\circ)$ is absolutely continuous on a non-empty hypercube $(L^\circ, H^\circ)^4$. □

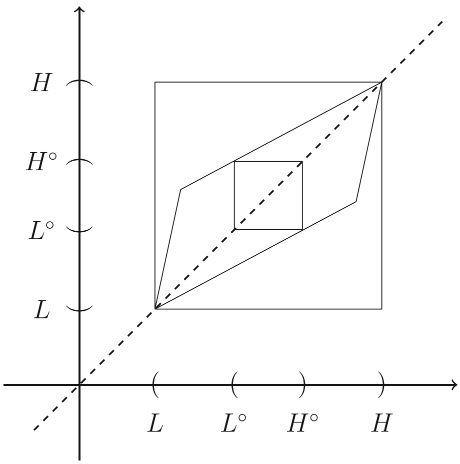

Fig. 11 The square $(L, H)^2$ contains a parallelogram $T((L, H)^2)$ which in turn contains a smaller square $(L^\circ, H^\circ)^2$. All three figures contain a segment of the diagonal

A.2.3 Proof of Theorem 15

Proof Fix $[\underline{\lambda}, \overline{\lambda}]$ with $\underline{\lambda} < \overline{\lambda}$. By Proposition 14, we can find a non-empty hypercube $(L^\circ, H^\circ)^4$ where the distribution of $(a^\circ, b^\circ, c^\circ, d^\circ)$ is absolutely continuous. Hence, by repeating the argument given in Proposition 12 within the smaller hypercube $(L^\circ, H^\circ)^4$ we get

$$\begin{aligned} P(L^\circ < a^\circ < H^\circ) &> 0 \\ P(L^\circ < d^\circ < a^\circ \,|\, a^\circ) &> 0 \\ P(a^\circ < c^\circ < H^\circ \,|\, a^\circ) &> 0 \\ P(d^\circ < b^\circ < d^\circ + c^\circ - a^\circ \,|\, a^\circ, c^\circ, d^\circ) &> 0 \end{aligned}$$

By Theorem 9, this ensures the positive probability of the profile $(\sigma^\circ, \sigma^\circ)$ being the sole symmetric equilibrium for $[\underline{\lambda}, \overline{\lambda}]$. The case of $(a^\bullet, b^\bullet, c^\bullet, d^\bullet)$ is analogous. □

A.2.4 Proof of Theorem 16

Proof Fix $[\underline{\lambda}, \overline{\lambda}]$ with $\underline{\lambda} \neq \overline{\lambda}$. By Theorem 15, there is positive probability of sampling $u_{1,1}, u_{1,2}, u_{2,1}, u_{2,2}$ that satisfy the anti-coordination inequalities $u_{1,1} < u_{2,1}$ and $u_{2,2} < u_{1,2}$, and the inequalities

$$u^\circ_{1,1} > u^\circ_{2,2}, \quad u^\circ_{2,1} - u^\circ_{1,1} > u^\circ_{1,2} - u^\circ_{2,2}, \quad \text{and} \quad \underline{\lambda} < \lambda^\circ < \overline{\lambda},$$

where $u^\circ_{1,1}, u^\circ_{1,2}, u^\circ_{2,1}, u^\circ_{2,2}$ are the equivalent of $a^\circ, b^\circ, c^\circ, d^\circ$ for $u_{1,1}, u_{1,2}, u_{2,1}, u_{2,2} = a, b, c, d$. Absolute continuity implies that $n \times n$ games where

$$u_{i,j} < \min\{u_{1,j}, u_{2,j}\} \quad \text{for } 2 < i \leq n \text{ and } 1 \leq j \leq n$$

have positive probability too. In such games, all actions different from I and II turn out to be strictly dominated and will not be chosen by either player type in any equilibrium. The only possible equilibria are thus profiles where only actions I and II are chosen. But then notice that from

$$u^\circ_{1,1} > u^\circ_{2,2}, \quad u^\circ_{2,1} - u^\circ_{1,1} > u^\circ_{1,2} - u^\circ_{2,2}, \quad \text{and} \quad \underline{\lambda} < \lambda^\circ < \overline{\lambda},$$

it follows that the only possible equilibrium is when type M plays I and tyspe R plays II. Hence, the only equilibrium of the $n \times n$ game is $(\sigma^\circ, \sigma^\circ)$, which proves that strongly informative $n \times n$ games have positive probability. The argument for $(\sigma^\bullet, \sigma^\bullet)$ is analogous. □

Author Contributions The two authors contributed equally to the development of the results and to the writing of the paper.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data Availability The Python script that generated the results of Sect. 6 is available here.

Declarations

Conflict of interest There are no conflict of interest to be disclosed.

Ethical approval Not applicable.

References

1. Alger I, Weibull JW (2013) Homo moralis: preference evolution under incomplete information and assortative matching. Econometrica 81(6):2269–2302
2. Azrieli Y, Teper R (2011) Uncertainty aversion and equilibrium existence in games with incomplete information. Games Econ Behav 73(2):310–317
3. Battigalli P, Cerreia-Vioglio S, Maccheroni F, Marinacci M (2015) Self-confirming equilibrium and model uncertainty. Am Econ Rev 105(2):646–677
4. Bednar J, Page S (2007) Can game(s) theory explain culture?: The emergence of cultural behavior within multiple games. Ration Soc 19(1):65–97
5. Blau RA (1974) Random-payoff two-person zero-sum games. Oper Res 22(6):1243–1251
6. Cassidy RG, Field CA, Kirby MJL (1972) Solution of a satisficing model for random payoff games. Manag Sci 19(3):266–271
7. Charnes A, Kirby MJL, Raike WM (1968) Zero-zero chance-constrained games. Theory Probab Its Appl 13(4):628–646
8. Cheng J, Leung J, Lisser A (2016) Random-payoff two-person zero-sum game with joint chance constraints. Eur J Oper Res 252(1):213–219
9. Dekel E, Ely JC, Ylankaya O (2007) Evolution of preferences. Rev Econ Stud 74(3):685–704
10. Di Tillio A (2008) Subjective expected utility in games. Theor Econ 3(3):287–323
11. Ely JC, Ylankaya O (2001) Nash equilibrium and the evolution of preferences. J Econ Theory 97:255–272
12. Epstein L (1997) Preference, rationalizability and equilibrium. J Econ Theory 73(1):1–29
13. Epstein L, Wang T (1996) "Beliefs about Beliefs" without probabilities. Econometrica 64(6):1343–73
14. Fawcett T, Hamblin S, Giraldeau L-A (2012) Exposing the behavioral gambit: the evolution of learning and decision rules. Behav Ecol 24:2–11
15. Feller W (1968) An introduction to probability theory and its applications, vol I, 3rd edn. Wiley, New York
16. Galeazzi P, Franke M (2017) Smart representations: rationality and evolution in a richer environment. Philos Sci 84(3):544–573
17. Galeazzi P, Galeazzi A (2021) The ecological rationality of decision criteria. Synthese 198:11241–11264
18. Galeazzi P, Marti J (2023) Choice structures in games. Games Econ Behav 140(C):431–455
19. Gigerenzer G (2008) Why heuristics work. Perspect Psychol Sci 3(1):20–29
20. Gigerenzer G, Goldstein D (1996) Reasoning the fast and frugal way: models of bounded rationality. Psychol Rev 103:650–669
21. Hagen EH, Chater N, Gallistel CR, Houston A, Kacelnik A, Kalenscher T, Nettle D, Oppenheimer D, Stephens DW (2012) 97Decision making: what can evolution do for us? In: Evolution and the mechanisms of decision making. The MIT Press
22. Halpern JY, Pass R (2012) Iterated regret minimization: a new solution concept. Games Econ Behav 74(1):194–207
23. Hammerstein P, Stevens JR (2012) Six reasons for invoking evolution in decision theory. In: Hammerstein P, Stevens JR (eds) Evolution and the mechanisms of decision making. MIT Press, Cambridge
24. Kajii A, Ui T (2005) Incomplete information games with multiple priors. Jpn Econ Rev 56:332–351
25. Klibanoff P (1996) Uncertainty, decision, and normal form games. Manuscript
26. Linhart PB, Radner R (1989) Minimax-regret strategies for bargaining over several variables. J Econ Theory 48(1):152–178
27. Lo KC (1996) Equilibrium in beliefs under uncertainty. J Econ Theory 71(2):443–484
28. Marinacci M (2000) Ambiguous games. Games Econ Behav 31(2):191–219
29. Osborne MJ, Rubinstein A (1994) A course in game theory. MIT Press, Cambridge
30. Renou L, Schlag KH (2010) Minimax regret and strategic uncertainty. J Econ Theory 145(1):264–286
31. Robalino N, Robson A (2016) The evolution of strategic sophistication. Am Econ Rev 106(4):1046–72

32. Rubinstein A (1991) Comments on the interpretation of game theory. Econometrica 59(4):909–24
33. Schlag KH, Zapechelnyuk A (2024) Compromise, donâ€™t optimize: generalizing perfect Bayesian equilibrium to allow for ambiguity. J Polit Econ Microecon 2(1):77–128
34. Solan E (2022) A course in stochastic game theory. London mathematical society student texts. Cambridge University Press, Cambridge
35. Song T (1992) On random payoff matrix games. Springer, Boston, pp 291–308
36. Tsybakov AB (2009) Introduction to nonparametric estimation. Springer, New York
37. Yu B (1997) Assouad, Fano, and Le Cam. In: Pollard D, Torgersen E, Yang GL (eds) Festschrift for Lucien Le Cam, chapter 29. Springer, Cham, pp 423–435
38. Zollman KJS (2008) Explaining fairness in complex environments. Polit Philos Econ 7(1):81–97