A Map of Ecologically Rational Heuristics for Uncertain Strategic Worlds

Leonidas Spiliopoulos and Ralph Hertwig
Max Planck Institute for Human Development, Berlin, Germany

The ecological rationality of heuristics has been extensively investigated in the domain of individual decision making. In strategic decision making, however, the focus has been on repeated games, and there is a lack of research on 1-shot games, where opponents and the game itself can vary from one interaction to another. Mapping the performance of simple versus more complex decision policies (or strategies) from the experimental game theory literature is an important first step in this direction. We investigate how 10 policies fare conditional on strategic properties of the games and 2 classes of uncertainty. The strategic properties are the complexity (number of actions) and the degree of harmony (competitiveness) of the games. The first class of uncertainty is environmental (or payoff) uncertainty, arising from missing payoff values. The second class is strategic uncertainty about the type of opponent a player is facing. Policies' performance was measured by 3 criteria: a mean criterion averaging over the whole set of opponent policies, a maxmin criterion capturing the worst-case scenario and another criterion measuring robustness to different distributions of opponent policies. Heuristics performed well and were more robust than complex policies such as pure-strategy Nash equilibria, while simultaneously requiring significantly less information and fewer computational resources. Our ranking of the decision policies' performance was closely aligned to their prevalence in experimental studies of games. In particular, the Level-1 policy, which completely ignores an opponent's payoffs and uses equal weighting to determine the expected payoffs of different actions, exhibited a robust beauty.

Keywords: game theory, strategic uncertainty, fast and frugal heuristics, ecological rationality

Supplemental materials: http://dx.doi.org/10.1037/rev0000171.supp

In 1974, Tversky and Kahneman wrote a foundational article in which they argued that "people rely on a limited number of heuristic principles which reduce the complex task of assessing probabilities and predicting values to simpler judgmental operations. In general, these heuristics are quite useful, but sometimes they lead to severe and systematic errors" (Tversky & Kahneman,

1974, p. 1124). A series of innovative experiments that Kahneman and Tversky published in the 1970s and 1980s drew psychology's attention to the importance of cognitive heuristics in human reasoning and helped to create new fields such as behavioral economics and behavioral law. Their programmatic work focused research attention for years to come on the existence of severe and systematic errors in people's judgments under uncertainty; at the same time, however, the positive side of the equation, the efficacy of heuristics, received little to no attention. Bolstered by the discovery of ever more biases as well as sweeping conclusions such as "mental illusions should be considered the rule rather than the exception" (Thaler, 1991, p. 4), the predominant view in psychology and behavioral economics has been that heuristics are fallible shortcuts that people, burdened by cognitive limitations, fall back on in situations where they would be better off using more sophisticated reasoning and decision-making strategies. Against the background of this emphasis on the fallibility of heuristics, surprising results began to emerge from another research program on heuristics. In the 1990s, the research program on ecologically rational heuristics (sometimes also called fast and frugal heuristics; Gigerenzer, Todd, & the ABC Research Group, 1999) began to demonstrate, often on the basis of systematic model competitions (e.g., Gigerenzer & Brighton, 2009; Gigerenzer & Goldstein, 1996; Katsikopoulos, Schooler, & Hertwig, 2010), that inferential heuristics that use less information and computation than more computationally complex strategies (e.g., logistic regression, CART, and nearest neighbor classifier) can actually offer good and sometimes even superior inferential accuracy. In addition, this research em-

phasized the ecological rationality of heuristics, asserting that heuristics are not good or bad per se; instead, the key question is: Given a specific heuristic, in what environments does it succeed or fail, and given an environment, what heuristics will succeed or fail in it (Gigerenzer, Hertwig, & Pachur, 2011; Hertwig, Hoffrage, & the ABC Research Group, 2013; Todd, Gigerenzer, & ABC Research Group, 2012)?

One important objection to these findings has been that all or most model competitions involved heuristics that concern individual decision making; thus, leaving open the question of whether their success would generalize to strategic situations in the social world. There are certainly grounds for doubt. The philosopher Sterelny (2003), for instance, argued that simple heuristics—despite reaching surprisingly high levels of accuracy in nonsocial worlds—will fail in interactions with other intelligent (and especially competitive) agents. "*For it is precisely in such situations that simple rules of thumb will go wrong*. . . . Catching a ball is one problem; catching a liar is another" (p. 53, emphasis added). More generally, his point was that the inferential paradigms in which heuristics' good performance was observed "rarely involve competitive, interacting, responsive aspects of the environment" (p. 208). Those paradigms fail to represent the complex need to strategize in the social world. Other scholars shared Sterelny's portrayal of social environments as especially complex and challenging (e.g., Byrne & Whiten, 1988; Humphrey, 1988; Whiten & Byrne, 1988; Whiten & Byrne, 1997, see Hertwig et al., 2013). Indeed, trying to catch a liar invokes very different cognitive faculties (that may be unique to humans) than catching a ball, and competitive social environments may very well be more complex than nonsocial environments—but does this mean that simple heuristics are doomed to fail? Empirical investigations of behavior in strategic interaction have regularly concluded that many people rely on strategies that are significantly less complex than the Nash equilibrium strategy—evidence comes from not only choice data but also process data (Crawford, Costa-Gomes, & Iriberri, 2013). What does this mean for how well people fare in strategic interactions? As we argue later in this section, in the context of one-shot games little is actually known about the cost of using simple decision policies. Furthermore, bounded cognitive and computational capacities are not the only possible reason people use heuristics so frequently in strategic interactions; it could be that heuristics promise success and robustness under conditions of profound uncertainty about others' strategic intent and incomplete knowledge. Investigating this possibility is our goal in this article.

In what follows, we examine decision making in strategic environments where the outcomes are determined by the interaction of at least two players, that is, we are in the realm of game theory. Games introduce a new type of uncertainty regarding how an opponent is likely to behave: strategic uncertainty. We are interested in systematically mapping how different decision policies perform in the presence of both strategic and payoff uncertainty, with the latter implemented as missing payoff information. Both types of uncertainty arise in the set of games referred to as games of incomplete information in the game theory literature. The interaction of strategic uncertainty and incomplete information makes for a particularly unpredictable game, especially if there is no relevant historical prior incident to refer to. In repeated games, there are various ways to reduce the degree of strategic uncertainty: Repeated behavioral observations allow for establishing a

reputation (Mailath & Samuelson, 2006), learning about the behavioral tendencies of other players (Cheung & Friedman, 1997; Erev & Roth, 1998; Spiliopoulos, 2012b, 2013a, 2013b), imitating other players (Duersch, Oechssler, & Schipper, 2012; Rendell et al., 2010), and strategically using historical decisions to teach the opponent to play differently in the future (Camerer, Ho, & Chong, 2002), to name but a few. The work in the iterated prisoner's dilemma game is of particular interest, because simple decision policies such as tit-for-tat in early work (Axelrod, 1984; Axelrod & Hamilton, 1981) and the win–stay, lose–shift heuristic in later work (Imhof, Fudenberg, & Nowak, 2007; Nowak & Sigmund, 1993) performed remarkably well (but see Rapoport, Seale, & Colman, 2015, for a more nuanced examination) However, one-shot games render it impossible to reduce strategic uncertainty. Players have no recourse but to face extreme uncertainty in a situation fitting of Savage's (1954) "large world." An intriguing but until now relatively underexplored (at least, systematically) question is whether the finding that simple strategies can be effective in repeated games carries over to one-shot games. Many important real-world decisions are either one-shot games or strategic situations that are so rare that very little prior information exists to reduce uncertainty. For example, important life decisions around one's career choices and workplace behavior, education, real estate investments, or personal relationships are made relatively rarely and carry large consequences. Communication, especially because of the shift toward online, shorter interactions with relatively unknown participants of different backgrounds and knowledge, carries a greater deal of uncertainty than in the past. One-shot games may have a disproportionately high degree of importance because of the ecological, negative correlation between probabilities and the magnitude of payoffs (Pleskac & Hertwig, 2014). Therefore, not infrequently very important decisions tend to be the ones with maximal strategic and payoff uncertainty because of the decision maker's lack of prior experience. These are the types of strategic interactions that we are interested in.

Why the Normative Solution Is Beyond Mere Mortals: The Unrelenting Complexity of the Nash Equilibrium

How can players resolve, reduce, or otherwise handle the various kinds of uncertainty involved in strategic one-shot interactions, which do not afford the opportunity for learning? Homo economicus resolves this strategic (and payoff) uncertainty by deductively, rather than inductively, solving for the Nash equilibrium—but this requires numerous strong epistemic assumptions, which in many cases are unlikely to be met even in games of complete information.[1] A set of strategies constitutes a Nash equilibrium if no player can gain by unilaterally changing their strategy (see Appendix A). These assumptions allow the Nash equilibrium to circumvent the problem of strategic uncertainty, but at the cost of not specifying a process by which the equilibrium is reached. The strong epistemic assumptions and reliance on deductive reasoning

---

[1] These assumptions include that each player's beliefs about their opponent's action are correct, that players are fully rational, (i.e., they best-respond to their beliefs), and that each player is aware of the other players' strategies (common knowledge assumption)—but see Aumann and Brandenburger (1995) for a loosening of these requirements for certain games.

on the part of players constitute the first limitation of the Nash equilibrium as a descriptive, rather than normative, model of behavior.[2] Our use of the term "normative" refers to the implications of both rational beliefs about players' behavior and optimization given those beliefs as applied to all players in a game. The second is the computational complexity of calculating a Nash equilibrium (Daskalakis, Goldberg, & Papadimitriou, 2009), particularly in mixed strategies (Savani & Von Stengel, 2006). The third is that multiple Nash equilibria often exist in a game (McLennan, 2005; McLennan & Berg, 2005) and, therefore, strategic uncertainty is not resolved in these cases. Players would still need consistent beliefs of how the opponent will choose to play from the set of equilibria (i.e., a theory of equilibrium selection). While many equilibrium selection refinements have been proposed (e.g., Harsanyi & Selten, 1988; Selten, 1975), they require ever more complex rationalizations and beliefs about the equilibrium selection behavior of their opponent. Therefore, they are unlikely to be good descriptive models of human behavior. While other less demanding and less strict normative criteria exist, they also suffer from the problem that the set of admissible actions is often even larger than the set of Nash equilibria.[3] One exception is the theory of focal points and how players can use them to coordinate on a specific equilibrium out of a large set of equilibria (Schelling, 1960). While simple in execution, focal points and social norms still require a shared culture or knowledge, and are relevant for games of coordination but not generally applicable to any type of strategic interaction.

Dealing with games of incomplete information requires even stronger assumptions about knowledge of the unknown information and computational requirements—the normative solution in this case is the Perfect Bayesian equilibrium (Harsanyi, 1967). Its existence is predicated on the common prior assumption, which states that players must hold common beliefs about all the unobservable information in the game.[4] Furthermore, in equilibrium all beliefs must have been updated using Bayes rule to ensure their consistency with the players' strategies. The indiscriminate application of Bayesian foundations to strategic decision making under uncertainty is not without its critics, even among prominent game theorists (e.g., Binmore, 2007).

Can Homo heuristicus navigate such difficult environments more effectively than Homo economicus? Using computer simulations, we compared the performance of random behavior (a baseline), eight heuristics, and the Nash equilibrium across environments with different levels of strategic and payoff uncertainty—we refer to these collectively as decision policies.

Why Heuristics?

The economics literature on bounded rationality in game theory in one-shot simultaneous move games has primarily been concerned with inferring the decision policies players use. The findings of the literature clearly support the conclusion that the majority of participants do not play according to the Nash equilibrium in one-shot games.[5] Instead, they use a variety of what are commonly referred to as boundedly rational decision policies, or heuristics—we return to these shortly. These experimental studies used games with strategic uncertainty but no payoff uncertainty.[6] In the face of both types of uncertainty, it is reasonable to expect that heuristic use would be even more prevalent, as any benefits

accruing from sophisticated reasoning are dependent on the available information. We propose an explicit and systematic exploration of whether using these heuristics leads to a significant performance loss compared with normative solutions for a broad range of environments. A few studies have touched upon related issues, albeit in a narrower, yet still instructive, scope—we present these below.

It is well documented that the Nash equilibrium often does not perform well either as a descriptive model or as a strategy against the empirically observed behavior in experiments (for a review, see Camerer, 2003). For example, Stahl and Wilson (1995) considered different Level-k policies' performance based only on a single experimental study with 12 different games. They concluded that the average performance of the policies over the set of games did not vary much, despite the existence of more variability

---

[2] While the Nash equilibrium may be an adequate predictor of the limiting behavior in the long run for certain types of repeated games, this is typically the consequence of simpler inductive learning rules that converge to the Nash equilibrium. In one-shot games, such inductive techniques are not possible; therefore, the Nash equilibrium—even as a predictive model of behavior—must be implemented by players using deduction.

[3] For example, an alternative normative solution, rationalizability weakens the requirement that beliefs are correct, and only requires that players best-respond to some belief (Bernheim, 1984; Pearce, 1984). As a generalization of the Nash equilibrium, the rationalizable set of actions is necessarily larger than the set of Nash equilibria of a game. Consequently, in many games the Nash equilibrium (and other weaker solutions, e.g., the correlated equilibrium) suffers from the problem of multiple equilibria; therefore, players must still find a way to coordinate on the same equilibrium.

[4] Thus, strategic uncertainty arising from not knowing an opponent's type or preferences is essentially treated as a form of uncertainty about an opponent's subjective payoffs. Players must then hold beliefs about the probability distribution of their opponent's type. Similarly, the Perfect Bayesian equilibrium can also deal with payoff uncertainty, but it requires consistent beliefs about the distribution of the payoffs. Note that in general, beliefs are typically modeled as point estimates, that is, the uncertainty about the incomplete information is assumed to be reducible to risk. Unfortunately, extending the Bayesian Nash equilibrium concept to uncertainty (or ambiguity) results in increasingly complex characterizations of the equilibrium, leading to intractable solutions.

[5] In repeated games of pure conflict, where participants accumulate experience with the same game (and against the same partner), it is possible that players inductively learn and converge to the Nash equilibrium (Binmore, Swierzbinski, & Proulx, 2001; Erev & Roth, 1998; Ochs, 1995; Roth & Erev, 1995). This hinges on the type of game played and the characteristics of the learning algorithm involved.

[6] There is another important limitation in the game theory literature concerning payoff uncertainty. While many theoretical proposals based on normative axioms exist, only a few experimental studies manipulate strategic uncertainty or examine behavior in the laboratory for games with payoff uncertainty. Because experimental studies are practically constrained, our knowledge of behavior is restricted to a set of games that can be implemented in the lab, typically with a small number of available actions and little payoff uncertainty. An interesting anecdote on the importance of strategic uncertainty is found in the aptly named manuscript "Granny versus Game Theorist: Ambiguity in Experimental Games" (Eichberger et al., 2008). Participants in this experiment knew that their opponent was a granny, a game theorist, or another participant. Players reported greater uncertainty (often referred to as ambiguity in economics) against the granny than the theorist, and stated beliefs about the opponent's likely behavior that were coarser (and less certain) against a granny than a theorist. These beliefs affected participants' behavior: They were more likely to choose an ambiguity-averse action when playing against a granny than against a theorist.

within games.[7] Consequently, significant incentives for players to use more sophisticated reasoning did not exist. Relatedly, Camerer, Ho, and Chong (2004) assessed how well the Nash policy and Level-k policies performed against the actual distribution of choices (and implicitly, policies) used by participants in experiments. Their measure of "economic value" of decision policies quantified how much more players would have earned had they used a study's proposed theory to predict their opponent's behavior (and best-responded to this prediction). The Nash equilibrium generally added less value than a Cognitive Hierarchy model, which assumes a mixture of levels of boundedly rational players, and in some cases even exhibited negative economic value. However, the arguments put forth in these studies justify deviations from Nash behavior *conditional* on the existence of other non-Nash players. As such, they are related to, but do not exactly answer the question that we are more interested in: Why are players *collectively* engaged in behavior that is incompatible with the normative solution?[8] Closer to our perspective, Camerer and Fehr (2006) related the theoretical payoff incentives for more sophisticated reasoning to the strategic structure of games. Games with strategic substitutes reward higher-level thinking (i.e., outsmarting an opponent), whereas games with strategic complements lead to higher payoffs when players cooperate or coordinate (i.e., reason similarly to an opponent).[9]

These findings are certainly supportive of the notion that more sophisticated reasoning—as exemplified in the limit by the infinitely recursive Nash equilibrium solution—is not necessarily beneficial. However, the studies involving bimatrix games were generally restricted to comparing the performances of a small set of heuristics or a limited number of highly stylized games (but see Camerer, Ho, & Chong, 2015, for a summary of heuristic use in a diverse range of other game classes). For example, the games typically had very few actions and none of these studies explored, simultaneously to strategic uncertainty, the impact of payoff uncertainty on performance. Consequently, little is known about the wider environmental conditions under which heuristics in one-shot games perform well or are likely to fail—that is, about their ecological rationality. Apart from exploring the robustness of these findings for a wide range of psychologically realistic environments, we are also concerned with *why* we find a preponderance of heuristic use in experimental studies. While the literature has not explicitly investigated why heuristics are often used in place of the Nash equilibrium, three arguments have been invoked implicitly.

The first is that players are not sophisticated enough to compute a normative solution. Because it is psychologically implausible to compute the Nash equilibrium, they turn to simpler decision policies. They are boundedly rational: They should use the Nash equilibrium but are not equipped to do so. The use of the term "boundedly rational" here is different from that espoused by Simon (1955, 1956), where bounded rationality is not viewed as a detrimental quality per se, but rather may be advantageous conditional on the decision environment. The second argument is that calculating the Nash equilibrium is associated with high decision costs and, therefore, a perfectly rational player should take these into account and optimize subject to constraints ("constrained optimization"; see, e.g., Arrow, 2004; Sargent, 1993). If the decision costs are high enough, then it may be rational to use a simpler decision policy; however, this ignores the additional computational cost associated with selecting the optimal simplicity of a policy. A

third explanation is based on the beliefs of players and violations of common knowledge about the rationality of all players. The Nash equilibrium is a rational response only if one believes that the opponent is also rational. However, rational players may have an incentive to deviate from the Nash equilibrium if they believe that their opponent is also likely to do so. Clearly, this conundrum is driven by the strategic uncertainty of the interaction.

We are explicitly interested in yet another explanation, which is not necessarily incompatible with the three mentioned above. This is that simple heuristics can be as accurate as, and sometimes even more accurate than, decision policies that take all the available information into account, including optimization models (see Gigerenzer et al., 2011). Our objective is to challenge the view that strategic complexity thwarts the success of simple heuristics by answering the following related questions. Are heuristics ecologically rational? Do heuristics fail in the face of strategic uncertainty and incomplete information or can a reduction in decision complexity lead to better and/or more robust behavior? In which environments can heuristics outperform optimizing policies, and in which will they lag behind?

We argue that the widespread use of heuristics relative to more complex policies (i.e., the Nash equilibrium strategy but also other relatively sophisticated policies such the L3 policy) may be a result of heuristics often performing on par with more complex policies and exhibiting greater robustness to both strategic uncertainty and payoff uncertainty. This provides an ecologically rational foundation (Smith, 2003) for the well-documented descriptive prevalence of heuristics in one-shot games, significantly extending the scope of similar arguments made for a limited set of games and/or heuristics (e.g., Camerer et al., 2004; Stahl & Wilson, 1995). We first address what we view as an important missing component of the literature concerned with strategic one-shot interactions, namely the lack of rigorous investigation—especially for games with incomplete information—into whether the use of boundedly rational heuristics leads to significant performance losses (in terms of payoffs) vis–à–vis the proposed normative solutions. Our work parallels investigations into cue-based inference, which often find that simpler heuristics perform better than normative approaches such as linear regression. Taking a cue from Hogarth and Karelaia (2006), we propose a first map of ecological rationality in games:

---

[7] For related arguments, consider the criticism of experimental studies purporting to show deviations from rational behavior based on the "flat maximum" critique (Harrison, 1989; Rapoport, 1975; von Winterfeldt & Edwards, 1973). The argument is that if the payoff function is relatively flat relative to different behaviors, then the experimental study has not provided adequate incentives to test for rationality. Note, however, that this critique is aimed at the specific games used in a study, not at whether this holds in general over a large number of games drawn from different environments. If this were the case, then simple heuristics would not entail a large payoff loss and, therefore, could not be said to be irrational.

[8] An example of an ecologically rational analysis, that is in a different domain (continuous double auctions) from our investigation is presented in Gode and Sunder (1993). They found that markets with agents using simple rules could achieve near-perfect allocative efficiency and arrive at the rational or equilibrium outcome. They concluded that this arose from the interaction between the simple rules of the agents and the specific market structure.

[9] Intuitively, for games with strategic substitutes (complements), an opponent's action induces a player to choose the opposite (same) action based on the incentives of the game. Coordination games are examples of strategic complements whereas games with incentives to free-ride are examples of strategic substitutes.

We assign the performance of simple decision policies to environmental niches, such as the properties of the games being played and the degree of payoff and strategic uncertainty.

The Setup

In the spirit of Simon's (1990) scissors analogy, a map of ecological rationality requires both a characterization of the environment (external to the agent) and the decision policies (internal to the agent). Simon (1955, 1956) saw bounded rationality in terms of two interlocking components: the limitations of the human mind and the structure of the environment. The implications of this conceptualization are twofold. First, models of simple heuristics need to reflect the mind's actual capacities rather than unbounded resources and capacities. Second, the structure of the environment may be the key to a heuristic's performance to the extent that the heuristic's architecture successfully maps onto it (or parts of it). For this reason, we now describe our computer simulations in terms of the game environment and the competing heuristics.

The Strategic Games

The games were all one-shot normal-form games with simultaneous moves, meaning that players made their decisions at the same time, without first being able to observe the opponent's choice. Table 1 gives an example of a normal-form game with two players: Player 1 can choose one of the strategies in the rows (Up, Middle, or Down); Player 2 can choose one of the strategies in the columns (Left, Center, or Right). The number of possible actions for each player is denoted by *n*—we assume that it is the same for both players and refer to this as the size of the game. The game in Table 1 is, therefore, a 3×3 game, with each player having three strategies (or actions) at their disposal. The numbers in the table represent the payoffs that each player will receive for every possible combination of strategies. The first number represents Player 1's payoff; the second, Player 2's payoff. For example, if Player 1 chooses Up and Player 2 chooses Center, they receive 32 and 89 units, respectively.

We distinguish between the true game, which is the one that nature has drawn, and the inferred game, which is the game a player believes to be playing. Individual payoffs of the true game can be categorized as observed (revealed to the player) or missing (not revealed to the player). Therefore, the true game differs from the inferred game only if there are missing payoffs. In inferred games, the player infers what the missing payoffs are to construct the game they believe they are playing around the observed payoffs. The actual payoffs of a decision policy are equal to the payoffs of the true game, not the inferred game.

Next, we describe how the true games were generated. The payoffs were generated by randomly sampling each payoff from the joint bivariate normal distribution:

$$[\pi, \pi'] \sim N(\mu, \Sigma) \quad \mu = [0, 0] \quad \Sigma = \sigma^2 \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}. \quad (1)$$

We chose the normal distribution rather than a uniform distribution to capture the typically negative correlation between magnitudes in payoff (whether negative or positive) and the probability of their occurrence in the real world (Pleskac & Hertwig, 2014). However, because the decision policies are all invariant to positive monotone transformations of the game payoffs, the qualitative findings or ordinal ranking of the performance of the decision policies will be identical for payoffs drawn from the uniform distribution. For the same reason, the exact values of μ and σ are irrelevant, as they simply scale the payoffs without affecting the actual choices of the decision policies. We fixed these to $\mu = [0, 0]$ and $\sigma = 100$ for simplicity. The value of ρ, however, plays an important role: It mediates the degree of correlation between the payoffs of the two players referred to as the *game harmony* (Zizzo & Tan, 2007, 2011). The more positive the degree of harmony, the more the players' interests are aligned in an environment, resulting in less conflict of interest. At the extreme where $\rho = 1$, there is no conflict and a self-interested player will choose the action that also allows a self-interested opponent to achieve their preferred outcome. The more negative the degree of harmony, the more the players' interests are opposed or in conflict. At the extreme where $\rho = -1$, players' interests are diametrically opposed and one player's gain necessarily comes at the other's expense; they are engaged in a constant-sum game. At the interior values of $-1 < \rho < 1$, the degree of harmony varies smoothly with a value of $\rho = 0$ implying an environment with no tendency for either conflict or harmony. Also, note that ρ is the expected degree of game harmony in an environment, but the realized game harmony of a single game is stochastic (Equation 1). For example, there will still be some games with significant conflicts of interest between players even in environments with $\rho > 0$, although the likelihood of these games arising will be lower than in environments with $\rho < 0$.

In environments with payoff uncertainty (more on this in the next subsection), the true and inferred games diverge. The payoffs of the inferred game consist of the observed payoffs of the true game and the inferred payoffs, which are imputed using the following mechanism. Relying on the principle of vicarious functioning (Brunswik, 1952; Dhami, Hertwig, & Hoffrage, 2004), players can use available cues (here, the observed payoffs) to infer unreliable or unavailable cues (the missing payoffs). For games with payoff uncertainty, we assumed that players impute the missing payoffs based on their estimates of the set of the parameters $\{\hat{\mu}, \hat{\sigma}, \hat{\rho}\}$ from the observed payoffs of the current game.[10] That is, they use these estimates to generate imputed values for the payoffs using Equation 1. For normal distributions, $E[\hat{\rho}] = \rho - \rho(1 - \rho^2)/2s$, where *s* is the number of samples the estimate is based on, which is equal to n^2 in our application. Consequently, for a small number of samples *s*, the value of the correlation coefficient is underestimated on average (Fisher, 1915). Despite this, the prob-

---

[10] The calculation of the exact statistical definitions of the variance and correlation are quite complex. There is, however, significant evidence that decision makers are aware of and relatively well calibrated to the statistical properties of an environment—that is, that they act as intuitive statisticians (Peterson & Beach, 1967). Because the exact modeling of these estimates is beyond the scope of our manuscript, we use the statistical definitions as a proxy, although in practice the actual processes used to compute these estimates may differ from their statistical definition.

Table 1

An Example of a 3×3 Normal-Form Game

Player 1	Player 2		
	Left	Center	Right
Up	58, 57	32, 89	94, 41
Middle	34, 46	23, 31	37, 16
Down	70, 74	41, 12	23, 53

ability of overestimating the correlation coefficient is higher than the probability of underestimating it because of the skew in the sample correlation coefficient, with important implications for the ability to detect correlations (Kareev, 1995, 2000; Kareev, Lieberman, & Lev, 1997).

Our random payoff-generating mechanism implies that every possible true game has a nonzero probability of being generated. Therefore, we did not restrict our investigation to widely researched 2×2 games such as the prisoner's dilemma game, the chicken game, and the stag hunt game. In fact, our analysis included all 78 types of 2×2 (ordinal) games taxonomized by Rapoport, Guyer, and Gordon (1976). Similarly, for $n \times n$ games where the size of the game is greater than 2 (and the possible types of games too many to enumerate), we did not restrict our attention to a particular type of game. To the best of our knowledge, this is the most comprehensive analysis of bounded heuristics in games.

The Environments

An environment e is defined by the properties of the games that nature randomly (exogenously) determines are to be played. Each environment e is defined by three properties: (a) the size of the action space of the generated games, n, (b) the probability that each payoff in the game is missing or unknown to a player, m (presented as a %), and (c) the degree of harmony ρ.

Size of the games.   This is perhaps the most straightforward property of the environment. It is simply the number of actions available to each player in a game. Larger games can be considered to be more complex, or cognitively demanding, as they offer players a larger number of possible contingent outcomes to consider.

Degree of payoff uncertainty.   In an environment with $m\%$ missing observations, each payoff of the game is randomly chosen with probability m to be missing, that is, unobservable to a player. The missing information induces payoff uncertainty about the true values of the missing payoffs. We take this opportunity to emphasize that a Bayesian Nash equilibrium would require that players impute the whole probability distribution of possible values for the missing payoffs. In other words, they would need to calculate a decision policy repeatedly (a very large number of times) to approximate the distribution of each possible realization of the missing payoffs.

Degree of harmony.   This third environmental property captures the extent to which players' interests are aligned or misaligned. To summarize, the set of environments E we investigated varied across the action space size, the likelihood of missing payoff information, and the degree of correlation between own and opponent payoffs: $\{n, m, \rho\}$. This set consists of all possible combinations of the following values: $n \in N = \{2, 3, \ldots, 19, 20\}$, $m \in M = \{0, 5, 10, \ldots, 75, 80\%\}$ and $\rho \in P = \{-0.5, 0, 0.5\}$. For example, a single environment $e(n, m, \rho)$ consists of a set of $n \times n$ games, sampled using the procedures noted in the previous subsection, where the probability that payoff values are missing is m and the correlation (degree of harmony) is ρ. For the sake of exposition, we refer to environments where $\rho = 0.5, 0, -0.5$ as harmonious, neutral, and discordant environments, respectively. Examining the performance of the decision policies across environments defined by these three dimensions (n, m, ρ) will produce maps of the ecological rationality of each decision policy.

The above constitute the environments and assumptions regarding the core simulation S reported in the main text. We also ran a number of additional simulations, varying other key parameters to test the robustness of our findings.[11] These simulations implement the following changes to the environments:

1. S': Each player knows their own payoffs perfectly, but does not know $m\%$ of their opponent's payoffs. This assumption is better suited to the possibility of an asymmetry in payoff information.

2. S'': Each action's quality can vary in expectation in terms of the average performance across an opponent's actions. A game is constructed by drawing payoffs for each action from a distribution with an action-specific mean; in the core simulation, the mean is the same (zero) for all actions. The mean associated with each action is drawn from a normal distribution with mean zero and variance equal to 100—we refer to the latter as between-action variance. This introduces within-action positive correlation in payoffs, that is, it models the possibility that some actions are inherently better than others on average (over possible opponent moves). Finally, within-action variance across payoffs is still equal to 100 as in the core simulation and for correlated environments payoffs are drawn from the same joint bivariate normal distribution, albeit with different means for each combination of players' actions.

3. S''': A simpler mechanism is used for inferring missing payoffs. Only $\hat{\mu}$ is imputed in each game, that is, players do not use information about the variance or correlation of payoffs, as in the core simulation. Missing values are replaced by $\hat{\mu}$ plus a very small random error $\epsilon \sim N(0, 10^{-6})$ for each payoff to avoid complications with ties.

Where relevant we refer to the results of these additional simulations or point the reader to the online supplemental materials where more details can be found.

The Competing Decision Policies

Our analysis compared 10 decision policies (playing against each other): the Nash equilibrium, a baseline random policy, and a set of eight heuristics that have been identified as commonly used by actual players in real-life situations (Costa-Gomes, Crawford, & Broseta, 2001; Costa-Gomes & Weizsäcker, 2008; Devetag, Di Guida, & Polonio, 2016; Polonio, Di Guida, & Coricelli, 2015; Spiliopoulos, Ortmann, & Zhang, 2018; Stahl & Wilson, 1994, 1995; Sutter, Czermak, & Feri, 2013). The decision policies are defined in Table 2; more detailed examples of how to compute them can be found in Appendix B.

The policies range from the very complex Nash equilibrium to moderately complex boundedly rational solutions (Level-k policies) to relatively simple and naive solutions such as choosing the

---

[11] We thank Colin Camerer and an anonymous referee for suggesting these robustness checks.

Table 2
The Set of Decision Policies

Decision policy	Abbreviation	Ω	OwnP	EQW	NW	SR	Dominance		Description
							Self	Opp.	
Maxmax	MaxMax	6th	Y	—	Y	Y	Y	—	Choose the action(s) offering the highest payoff for the player
Maxmin	MaxMin	6th	Y	—	Y	Y	Y	—	Choose the action(s) offering the highest worst-case payoff for the player
Level-1	L1	6th	Y	Y	—	—	Y	—	Choose the action(s) offering the best response to the assumption that an opponent is choosing randomly
Social maximum	SocMax	3rd	—	—	—	—	—	—	Choose the action(s) maximizing the sum of the player's own payoff and the opponent's payoff
Equality	Eq	3rd	—	—	—	—	—	—	Choose the action(s) minimizing the difference between the player's own payoff and the opponent's payoff
Dominance-1	D1	2nd	—	Y	—	—	Y	Y	Choose the action(s) offering the best response to the assumption that an opponent is choosing randomly over their nondominated actions
Level-2	L2	5th	—	—	—	—	Y	Y	Choose the action(s) offering the best response to the assumption that an opponent is applying L1
Level-3	L3	4th	—	—	—	—	Y	Y	Choose the action(s) offering the best response to the assumption that an opponent is applying L2
Nash equilibrium	NE	1st	—	—	—	—	Y	Y	Choose the action(s) consistent with the Nash equilibrium

Note. OwnP = the policy considers only the player's own payoffs and ignores the opponent's payoffs; EQW = equal weighting of own payoffs; NW = no weighting; SR = consider a single reason (payoff); Dominance Self/Opp. = obeys the principle of strategic self-/opponent-dominance (see Appendix A). Y = policy has the property; — = policy does not have the property. Ω refers to the rank order of the decision policies according to their complexity (as analyzed in Spiliopoulos, Ortmann, & Zhang, 2018), with 1st referring to the most complex one and 6th to the least complex one.

action that contains the maximum own payoff. We address the evaluation of complexity shortly. Level-k reasoning (and the related cognitive hierarchy theory) has been particularly successful in modeling players' behavior in one-shot games (Camerer et al., 2004; Costa-Gomes et al., 2001; Stahl & Wilson, 1995). The levels of reasoning are anchored on the least strategic player, who uses the Random policy (often referred to as Level-0). A Level-1 (L1) player best-responds to the belief that their opponent is Level-0. In general, a Level-k player expects an opponent to be a Level-(k-1) player, and best-responds to this belief. Thus, the Level-k (Lk) reasoning model loosens the assumptions of the Nash equilibrium (NE), namely that all players are fully rational and aware of their rationality (common knowledge). Beyond L1, L2, L3, and NE, the set of policies includes maxmax (MaxMax), maxmin (MaxMin), social maximum (SocMax), equality (Eq), and dominance-1 (D1).

In terms of benchmark strategies, we focused only on the pure-strategy NE for two reasons. First, as discussed earlier, the calculation of a mixed strategy NE is computationally demanding.[12] Therefore, we consider only the subset of NE that could plausibly also be implemented as a descriptive model of behavior in one-shot games, as there is some evidence of pure-strategy NE use in simple games (Rey-Biel, 2009; Spiliopoulos et al., 2018; Stahl & Wilson, 1995). Second, the epistemic interpretation of a mixed strategy, particularly in a one-shot game, is controversial:[13] "We are reluctant to believe that our decisions are made at random. We prefer to be able to point to a reason for each action we take. Outside of Las Vegas we do not spin roulettes" (Rubinstein, 1991, p. 913). Consequently, we regard the NE not only as a normative strategy, which also serves as a benchmark in our simulation, but also as a candidate for a descriptive policy.

Table 3 lists experimental studies comparing the observed, descriptive prevalence of decision policies. Our selection of implemented decision policies was guided by these empirical studies suggesting that they are, to varying degrees, actually used by players. The L1 heuristic is the most prevalent policy in six cases and ranks second in another three. D1 is also very common, ranking first in two cases and second in another three. L2 ranks first in three cases and third in another three. The NE is never the most common policy in any study. While it is ranked second in two studies and third in a further two, in six other cases it does not figure into the three most prevalent policies at all. Fudenberg and Liang (2017) found that L1 is the modal action in 72% of previously investigated one-shot games in the laboratory and in 87% of randomly generated games (similar to ours, but without payoff uncertainty). Finally, Spiliopoulos et al. (2018) concluded that under time pressure there is a shift to less complex decision policies such as L1 and L2, but also to SocMax, which was ranked in the top three for only another two cases without time pressure.

Each of the policies studied, with the exception of the (pure-strategy) NE, always suggest a unique action for any normal-form game (with unique payoffs). To resolve the coordination problem in the case of multiple NE, we assumed that players choose the equilibrium maximizing the joint payoffs to both players. This assumption is the most favorable to NE as any other equilibrium choice would necessarily lead to poorer performance. However, some games may not have a pure-strategy NE; in such cases we assumed that the NE strategy chooses an action randomly. The

frequency of games without pure-strategy NE in some environments will be nonnegligible—we address the impact of this in the results section.

Classification of Decision Policies

The decision policies can be classified according to at least two dimensions: the (quantifiable) level of complexity and the chosen path toward simplification.

Level of complexity. Spiliopoulos et al. (2018) calculated the complexity of this set of decision policies in terms of the number of elementary information-processing units (Payne, Bettman, & Johnson, 1993) required to execute them. According to a ranking based on elementary information-processing units, the NE is the most computationally complex policy—a rank ordering of the policies can be found in Table 2, denoted by Ω. The heuristics L3, L2, D1, SocMax, and Eq are of intermediate complexity. Finally, the L1, MaxMax, and MaxMin heuristics are the least complex policies, requiring significantly fewer computational resources than the rest. The appropriateness of measuring complexity along these lines is corroborated by process-tracing studies using eye-tracking or Mouselab, which find a strong relationship between the order of comparison and fixation of game payoffs and the types of decision policies participants used. For example, the complexity of Level-k policies increases in k as a larger number of fixations and comparisons must be performed. Representative studies using process-tracing in strategic games include Brocas, Carrillo, Wang, and Camerer (2014); Camerer, Cachon, and Johnson (2004); Camerer, Johnson, Rymon, and Sen (1993); Chen, Huang, and Wang (2018); Devetag et al. (2016); Johnson, Camerer, Sen, and Rymon (2002); Knoepfle, Wang, and Camerer (2009); Polonio et al. (2015); Wang, Spezio, and Camerer (2010).

Path to simplification. The heuristics can be also be categorized qualitatively by the paths to reduce computational complexity they embody. These include reducing the amount of information required, rendering the process of integrating information less complex, or simplifying the assumptions made about an opponent or their beliefs. The paths taken to reduce complexity can be broadly categorized as payoff-based or probability-based simplification. Here we identify four paths: using only own payoff

---

[12] To enumerate all equilibria (pure and mixed) in a 20×20 game using the algorithm in Avis (2000) requires approximately 8.36 hr per game (calculated using the implementation in the Game Theory Explorer http://gte.csc.liv.ac.uk/index/). For 15×15 games, it is approximately 2 min per game.

[13] One interpretation of a mixed strategy Nash equilibrium, is to specify it over a *population* of players individually choosing a pure strategy. This is consistent with our assumption that players use a decision policy recommending a pure strategy; therefore, if players use different decision policies then at the population level aggregate behavior will appear to be mixed. Another alternative, which is applicable at the individual level, is the purification interpretation (R. J. Aumann, Katznelson, Radner, Rosenthal, & Weiss, 1983; Harsanyi, 1973). A player chooses deterministically based on a random variable that only he or she is privy to—an external observer would view this player's behavior as stochastic.

Table 3
Empirical Prevalence of Decision Policies

	Prevalence (ranking)		
Studies	First	Second	Third
No time pressure			
Stahl and Wilson (1994)	L2	NE	L1
Stahl and Wilson (1995)	L1	Random	NE
Costa-Gomes, Crawford, and Broseta (2001)	L2	L1	D1
Costa-Gomes et al. (2001)[a]	D1	L1	L2
Costa-Gomes and Weizsäcker (2008)	L1	D1	L2
Rey-Biel (2009)	D1	NE	L2/L3
Sutter, Czermak, and Feri (2013)	L1	D1	SocMax
Polonio and Coricelli (2019)[a]	L2	L1	Comp.
Polonio, Di Guida, and Coricelli (2015)[a]	L1/SocMax		L2
Spiliopoulos et al. (2018)[a]	L1	D1	NE
Time pressure			
Spiliopoulos et al. (2018)[a]	L1	SocMax	L2

Note. NE = Nash equilibrium; L1 = Level-1; L2 = Level-2; L3 = Level-3; SocMax = social maximum; D1 = dominance-1. Comp. is a competitive policy maximizing the difference between own and opponent payoffs.
[a] Marks analyses including information search (process-tracing).

information, which is payoff-based; and equal weighting, no weighting, and using a single reason, which are probability-based.

Payoff-based simplification.

Path 1: Ignore the strategic element of games by using only own payoff information and ignoring the opponent's payoffs (OwnP). In terms of information search, players may choose to search for both own and opponent payoffs, or reduce the game to a nonstrategic decision task by completely ignoring an opponent's payoffs. The set of decision policies that require only information about own payoffs are MaxMax, MaxMin, and L1. The remaining decision policies require information about both own and opponent payoffs.

Probability-based simplification. In contrast to individual decision making involving risky gambles (decisions from description; Hertwig & Erev, 2009), no explicit probabilities of outcomes are provided to players in the games. A common approach in the game theory literature is for a player to form beliefs about an opponent's behavior. Based on these beliefs an expected payoff calculation can be made for each action, and the player can choose the action affording the highest expected payoff. At the NE, beliefs about the opponent's behavior must be consistent. Abolishing this requirement opens the door to nonequilibrium beliefs: Players are free to form beliefs as they see fit; thereby, allowing for at least the following three probability-based simplification strategies.

Path 2: Do not estimate the probability of each action but apply the principle of indifference (equal weighting; EQW). One simple assumption is that the probabilities of each action of the opponent are equal, resulting in an equal weighting of the player's own payoffs when inferring the expected payoffs to an action. This can be viewed as a strategic variant of the equal-weighting principle proposed for prediction (Dawes, 1979) and of Laplace's principle of indifference (Stigler, 1986). The heuristics that make this assumption are L1 and D1. D1 first removes the opponent's dominated actions, and only then assigns equal weights to the remaining actions. L1 is of particular

interest because it is the most frequently recruited by real players in strategic games.

Path 3: Do not form beliefs about an opponent's behavior (no weighting; NW). This path to simplification does away with the notion of beliefs completely so that payoffs are not weighted at all by the likelihood of an opponent choosing a particular action. SocMax and Eq take this path; they do not weight payoffs according to the likelihood of the relevant outcomes being obtained. Instead they require either the addition or subtraction of own and opponent payoffs for every possible outcome, and then perform only ordinal comparisons on the results.

Path 4: Base the decision on a single reason only (SR). An even more spartan path to simplification is to not integrate any payoffs at all, but instead make a choice based only on a single reason (or payoff, in games). This approach avoids both weighting and addition (completely forgoing any belief formation), and is the most computationally simple class of decision policies.

Combining payoff- and probability-based simplification. The decision policies that combine both a payoff- and probability-based path to simplification are L1, MaxMax, and MaxMin. L1 combines Paths 1 and 2 as it both ignores the opponent's payoffs and uses equal weighting. MaxMax and MaxMin combine Paths 1 and 4. MaxMax is predicated on choosing the action with the highest own payoff, MaxMin on choosing the action that guarantees the highest own payoff assuming the worst-case scenario. Neither MaxMax nor MaxMin require weighting or addition, only ordinal comparisons among own payoffs based on maximum and minimum operations.

Other more nuanced paths to simplification also exist. For example, we have already hinted at the fact that some decision policies make only ordinal comparisons—that is, they ignore the cardinal information given in the absolute magnitude of the game's payoffs. L1, for instance, requires cardinal payoff information, whereas MaxMax and MaxMin use ordinal information only.

Strategic dominance. Decision policies can also be classified along yet another dimension, namely, whether they obey the

principle of strategic dominance, and their assumptions about whether or not the opponent adheres to this principle. If one of a player's (pure) strategies is better than another strategy—independent of the strategy chosen by the opponent—then that strategy dominates the other strategy (the latter is the dominated strategy).[14] If a strategy dominates all other strategies in a game, it is referred to as a dominant strategy. We classify a policy as obeying self-dominance if it never chooses a dominated strategy. MaxMax, MaxMin, all Level-k policies (including D1) and NE obey self-dominance. We classify a policy as obeying opponent-dominance if it assumes that the opponent policy obeys its own self-dominance. Accordingly, D1, L2, L3, and NE also obey opponent-dominance.

The advantage of obeying dominance principles in a particular environment depends on the probability that dominated actions exist in each of the environment's games. The proportion of games that have at least one dominated action falls quickly as the size of the game increases (as $n \rightarrow \infty$, this probability approaches zero).[15] Therefore, it is reasonable to expect that the influence of dominance plays an increasing role in environments with small games; thus, creating an ecological niche for policies that uphold strategic dominance. Recall that L1 and D1 differ only in the first step: D1 removes the opponent's dominated strategies. Consequently, L1 and D1 will converge in their recommended actions as the number of actions increases.[16] However, in terms of processes and information needs, D1 would still require more steps than would L1 to arrive at the same recommendation.

The Performance Metrics

For a given environment, we performed simulations using 25,000 randomly drawn games where each target policy played against all other policies, including itself. Because of the symmetrical nature of the games, by averaging over both row and column players, our findings are based on 50,000 strategic interactions per target policy-competitor pair. Imagine each player choosing in advance which decision policy to use across all games in a specific environment. Now consider how strategic uncertainty affects the choice of policy. If a player knows the opponent's policy, they can easily figure out the policy offering the best response. If they do not know the opponent's policy, they may form beliefs about the distribution of decision policies in the population and, at least in theory, calculate the expected payoffs for each policy conditional on those beliefs. This is a probabilistic quantity; it requires weighting the expected payoff against each decision policy by the probability of being matched with a player using that particular policy. However, it is very difficult to learn the distribution of policies in these environments. Consequently, players face profound strategic uncertainty and are most likely unable to assign probabilities to the distribution of policies in the population. Luce and Raiffa (1957) argued that decisions in such large worlds (Savage, 1954) may be enabled by one of two principles: the principle of indifference (also known as the principle of insufficient reasoning) or the maxmin principle. These two principles—which both assume that players are interested only in their own payoffs, not those of their opponents—inform our performance criteria. While this is a limitation of our analyses, it is not an unreasonable first-order approximation given the environments that we are examining for two reasons. First, some kinds of other-regarding preferences (such as

reciprocity and social image concerns) are typically more important in repeated rather than one-shot interactions with random rematching. Also, as-if prosocial behavior is more likely to be found by selfish individuals engaged in repeated interactions because of concerns about future play, for example, fear of punishment or reputational concerns. Second, because our environments include significant payoff uncertainty, players with social preferences would be obliged to use noisy inferred payoffs and, therefore, would not be able to reliably make decisions that are also in the interest of their opponents.

According to the principle of indifference, each decision policy is equally likely to be used by the opponent. A player's expected payoff over the whole set of games in an environment is then simply an average of the expected payoffs against each decision policy. We define this as the Indifference criterion (see Appendix C for the mathematical definition). It is the expected payoff a decision policy will achieve under two conditions: first, if the policy plays against a population of policies that are uniformly distributed in the player population or, second, if it plays against a single decision policy but does not know which, and believes that the policy is drawn with equal probability from the set of decision policies.[17] Under the assumption that players' utility functions are linear in the payoffs (i.e., risk neutral), then this criterion is what a player with uniform beliefs would attempt to maximize.

To avoid confusing the maxmin principle with the MaxMin heuristic, we refer to it here as the Wald criterion (Wald, 1945).[18] According to the Wald criterion, in contrast, the performance of a policy is measured in terms of the worst-case scenario with respect to its opponent's decision policy. In other words, in the face of strategic uncertainty about an opponent's policy, the player assumes the worst: That given their chosen policy the opponent will be using the policy that imposes the worst possible average payoff for that environment. Consequently, the player chooses the policy that maximizes the expected payoff from this worst-case scenario. Policies that are robust to severe strategic uncertainty should

---

[14] Our concept of strategic dominance is restricted only to pure strategies, not mixed strategies, to match our decision to consider only pure-strategy Nash equilibria.

[15] Indicatively, for a true game of size n = 2, 3, 4, 5, 10, 20 it is 0.5, 0.58, 0.55, 0.48, 0.08, 0.0004, respectively.

[16] Considering the existence of dominated or dominant strategies in a game can reveal further relationships between the decision policies. As observed by Costa-Gomes et al. (2001), a Level-k heuristic's proposed action is identical to that of the Nash equilibrium for games that are solvable by k rounds of iterated dominance (in pure or mixed strategies). Level-k heuristics (that first removed dominated strategies of their opponent from the consideration set, e.g., D1) are identical to the Nash equilibrium in games that can be solved by k + 1 rounds of pure-strategy dominance. Consequently, in many 2 × 2 games with complete information there is significant overlap between decision policies, leading many simpler heuristics to emulate the Nash equilibrium.

[17] Recall that L1 and D1 converge in their prescribed actions, especially for $n > 10$, and that L2 is by definition a best response to L1. Therefore, to avoid giving L2 an advantage over other policies (that are at best, a best response only to one other policy) we have halved the weighting of L1 and D1 in this criterion. Without this adjustment, qualitative results are similar but L2 does receive a small boost as expected. See Appendix C for more details.

[18] The MaxMin heuristic operates over the possible actions chosen in a single game, whereas the Wald criterion operates over the policy chosen by an opponent in an environment, that is, a set of games.

perform well on the Wald criterion. This criterion could also be interpreted as one associated more closely with a satisficing rather than optimizing player, as maximizing the worst-case scenario guarantees a particular level of payoffs, which can be interpreted as the minimum aspiration level.

In a nutshell, the Indifference criterion, $\bar{\pi}$, captures policies' average performance against the whole set of possible opponent policies, assuming each is equally likely. If players choose the best policy according to the Indifference criterion, this would imply uncertainty neutrality. In contrast, the Wald criterion, π_w, captures the minimum performance that can be guaranteed regardless of the opponent's policy. If players choose the best decision policy according to the Wald criterion, as suggested by Gilboa and Schmeidler (1989), this would imply (extreme) uncertainty aversion.

Last but not least, we propose a third alternative measure of performance, the Composition-Robustness criterion. In contrast to the Indifference criterion, where opponents are implicitly assumed to be encountered in the population with equal probability, the Composition-Robustness criterion allows for every possible mixture (composition) of opponent policies in the population. Therefore, this measures the robustness of a decision policy with respect to strategic uncertainty for a given environment—that is, whether it performs well regardless of the composition of opponent policies in the population. The criterion is calculated using the percentile ranking of a decision policy's performance in an environment, averaged across all possible population compositions. A ranking of 1 (0) indicates that a policy was always the best (worst) performing policy in an environment, regardless of the composition of the opponent population. A detailed definition of this criterion and a discussion of the findings can be found in Appendix D. Finally, in Appendix E (Tables E1–E3) we also examine pairwise comparisons of the policies, thereby breaking down policies' performance according to the Indifference criterion into its constituent components conditional on each possible opponent policy.

Having defined these performance criteria, we now turn to the results of the competition to examine how well the heuristics fared when facing both strategic and payoff uncertainty. Note that the online supplemental materials contains reproductions of the same analyses as below for the additional simulations—we will report the most importance differences between the core and additional simulations at relevant places in the main text. Furthermore, we will also assess policy performance over two different subsets of the core simulation:

1. In contrast to the results in the main text, which explore the whole space of environments $n \leq 20$ and $m \leq 80\%$, we also examine a subset of environments that are more in line with the those explored in the experimental literature, where game sizes are typically smaller and missing payoffs are less likely. This subset consists of the environments where $n \leq 10$ and $m \leq 50\%$.

2. As pointed out, the NE decision policy chooses the pure-strategy NE when at least one exists, but otherwise chooses randomly. To explore the degree to which this policy is affected by a lack of pure-strategy NE, we restrict the performance criteria to only those inferred

games (in each environment) for which at least one pure-strategy NE exists.

Performance on the Indifference Criterion, $\bar{\pi}$

Figures 1, 2, and 3 show the performance of the competing policies across the neutral, discordant, and harmonious environments, respectively, measured in terms of the Indifference criterion. Each figure consists of 10 subgraphs (heatmaps) showing the performance of each decision policy conditional on the size of the game (n) along the y-axis and payoff uncertainty (m, the percentage of missing information) along the x-axis for a given value of ρ. As a benchmark, the top–left and top–center panels show the performance of random choice and the NE. The darker the shading, the better the performance. For each environment, we will refer to the best-performing policy and all other policies within 5% of its performance as the "set of top-performing" policies; these policies can be considered as not being significantly economically different.[19] Alternatively, these could be considered as the set of satisficing policies: Heuristics need not necessarily outperform more complex policies to be considered useful, as this ignores the tradeoff in terms of computational complexity and performance. This set of policies is marked by a white dot overlaid on the heatmaps. The robustness of a policy for given ρ is a function of how large the area of dark shading is—the larger the area, the more robust the policy's performance over a range of game sizes and degree of payoff uncertainty. A precursor to parts of the analysis in this section has appeared as a book chapter (Spiliopoulos & Hertwig, 2019).

Neutral Environment ($\rho = 0$)

We begin with the neutral environment, in which the player's and opponent's interests are neither in harmony nor in conflict. The following comparisons discussed here can be observed in Figure 1.

The top-performing policies obey the self-dominance principle and use equal weighting. L1 and D1 are the two best performing heuristics across the map. They perform well across all game sizes, and exhibit considerable robustness to high payoff uncertainty. The next best policies, L2 and L3, are similarly robust, but perform on average significantly worse than L1 and D1. Note that L1 obeys the self-dominance principle; D1, L2, and L3 are consistent with both the opponent- and self-dominance principles.

Assuming higher levels of rationality leads to poorer performance. Increasing the level of rationality attributed to the opponent beyond those assumed by L1 and D1 does not offer any advantage. L2 and L3 not only perform significantly worse, but also impose a significant increase in the computational complexity and informational requirements. Similarly, the NE, which makes the strongest assumptions about the opponent's rationality, is not as robust as L1 and D1. The NE achieves relatively high payoffs

---

[19] The notion of statistically (rather than economically) significant differences between decision policies is not informative in these simulations, because for a large enough set of games an arbitrarily small difference will always be significantly different in the statistical sense. The findings are very similar if the criterion for economic significance is a difference greater than 10% of mean payoffs to all decision policies in an environment.
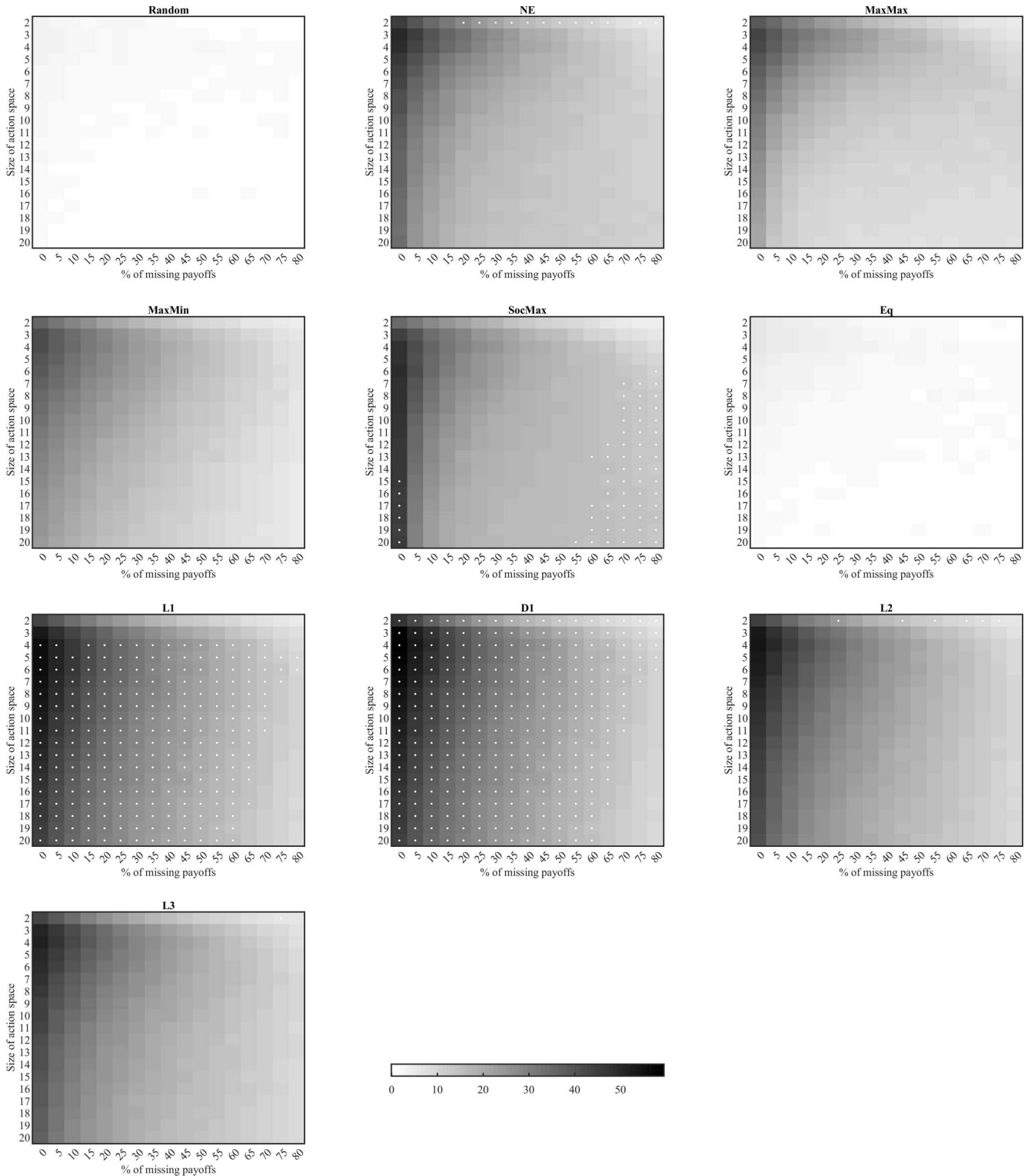
Figure 1. Decision policy performance according to the Indifference criterion ($\overline{\pi}$) for neutral environments ($\rho = 0$). The set of top-performing policies for each environment is marked by an overlaid white dot. NE = Nash equilibrium; L1 = Level-1; L2 = Level-2; L3 = Level-3; SocMax = social maximum; D1 = dominance-1; Eq = equality.

Figure 2. Decision policy performance according to the Indifference criterion ($\overline{\pi}$) for discordant environments ($\rho = -0.5$). The set of top-performing policies for each environment is marked by an overlaid white dot. NE = Nash equilibrium; L1 = Level-1; L2 = Level-2; L3 = Level-3; SocMax = social maximum; D1 = dominance-1; Eq = equality.
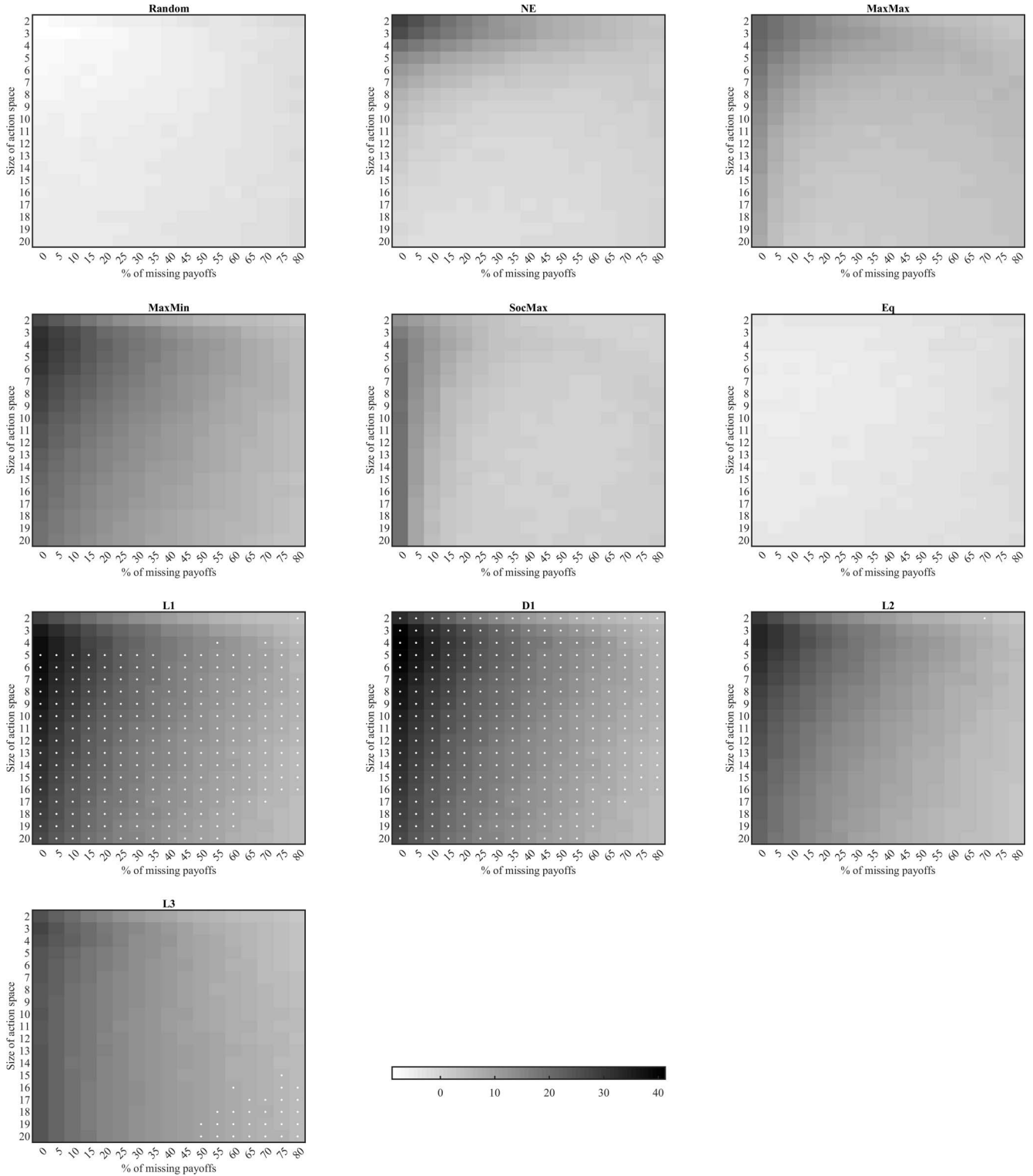
Figure 3. Decision policy performance according to the Indifference criterion ($\overline{\pi}$) for harmonious environments ($\rho = 0.5$). The set of top-performing policies for each environment is marked by an overlaid white dot. NE = Nash equilibrium; L1 = Level-1; L2 = Level-2; L3 = Level-3; SocMax = social maximum; D1 = dominance-1; Eq = equality.
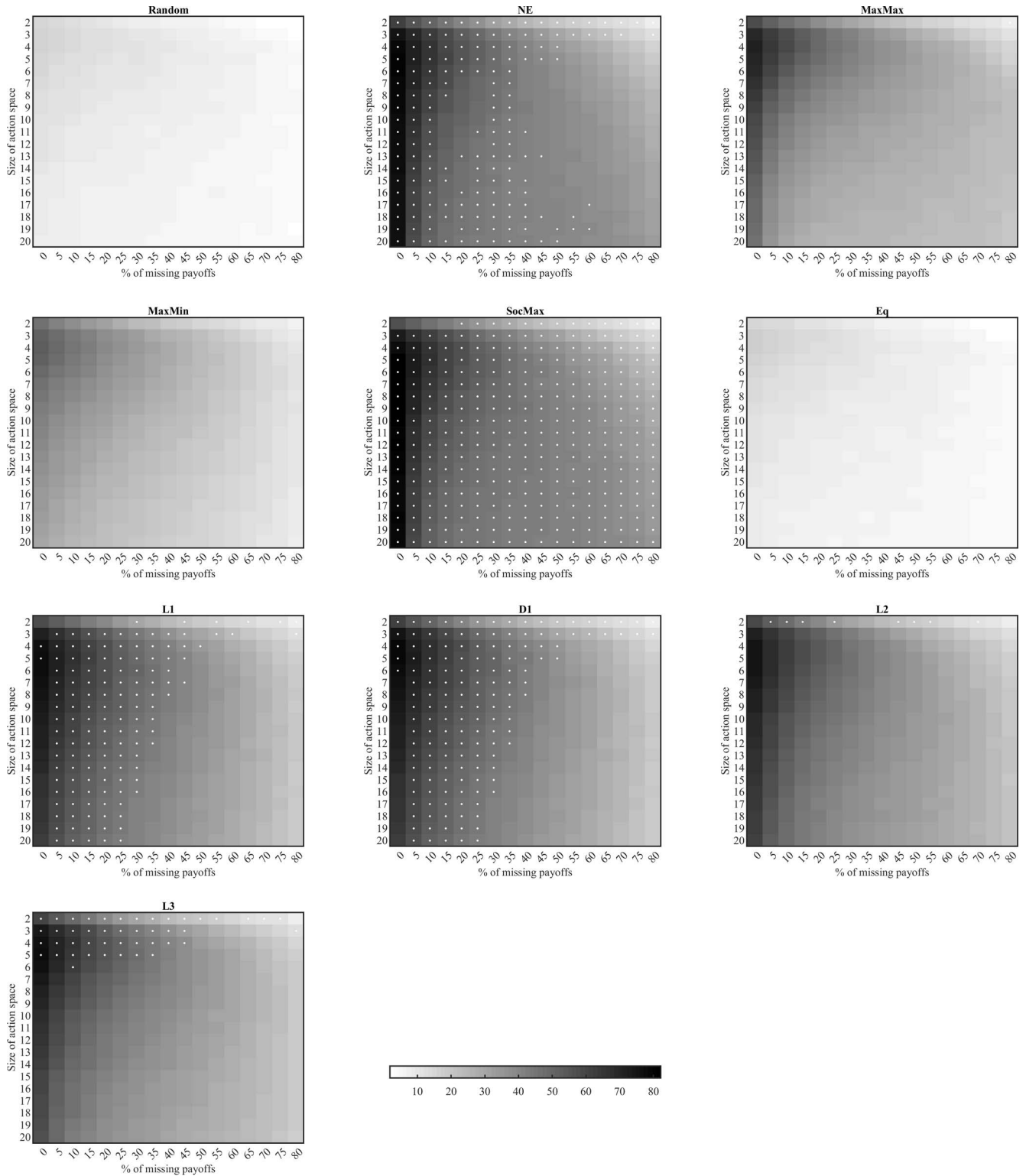
only in very small games and is in the set of best performers only for games of size 2 and moderate to high payoff uncertainty.

Policies using a single reason oversimplify. Two of the computationally simplest heuristics, MaxMax and MaxMin, achieve moderate performance in just a small subset of environments involving low payoff uncertainty. The extreme simplification embodied in these heuristics seems to overshoot the mark, leading to poor performance. However, SocMax, which is computationally simple but requires information about both own and opponent payoffs, is in the set of top performers for two niches: in large games if there is no missing information and in games where $n \geq 7$ combined with very high uncertainty. Overall, SocMax is the fifth highest performing policy (averaged over all n and m).

Discordant Environment ($\rho = -0.5$)

To what extent do the aforementioned findings generalize to an environment in which a player's and an opponent's interests diverge? The answer can be found in Figure 2. For the sake of conciseness, we spell out the main differences in the results compared with the neutral environment. The best performing heuristics overall and in terms of robustness to game size and payoff uncertainty are still L1 and D1. However, their closest competitors in neutral environments, NE, L2, L3, and SocMax, perform significantly worse and are less robust in discordant environments. The demise of the NE is particularly striking as its performance in medium to large games collapses across all degrees of payoff uncertainty. This is primarily because of the fact these environments include many games without a pure-strategy NE—we return to this point later. With the exception of a small niche of large games with high payoff uncertainty where L3 is the best performer, no policies other than the top performers in the neutral environment belong to the set of top performers in the discordant environment.

Harmonious Environment ($\rho = 0.5$)

Finally, what happens in an environment in which interests are aligned? Results in Figure 3 suggest substantial differences. Averaging over the whole map, SocMax is the best performing policy, followed closely by the NE; the latter lags behind SocMax if payoff uncertainty is high ($m \geq 40\%$). Computational complexity is also in favor of SocMax, which is significantly less complex than the NE. Furthermore, the NE implementation assumes that the equilibrium with the highest joint payoffs is chosen; therefore, the NE policy would perform worse for an alternative equilibrium selection assumption. L1 and D1 are the next best policies and they belong to the set of top performers for approximately half of the environments (low to moderate payoff uncertainty, largely independent of the game size). They are followed by L2 and L3, which belong to the set of top performers only for small games. Very simple strategies such as MaxMax and MaxMin never belong to the set of top performers.

Performance on the Wald Criterion, π_w

The second performance criterion focuses on the robustness of a policy under the assumption that the player is matched with an opponent who chooses the policy with the worst possible outcome

for the player. Figures 4, 5, and 6 plot the competing policies' performance, using the same logic as the previous figures. A high proportion of dark shading implies robustness to both strategic and payoff uncertainty for the majority of game sizes. The key results are summarized below.

Neutral Environment ($\rho = 0$)

Using equal weighting and obeying the self-dominance principle pays off. The top-performing heuristics across all environments are L1 and D1; no other policy belongs to the set of top performers in any of the environments. Their performance is particularly high relative to the other policies in games of all sizes with low to moderate payoff uncertainty. Both top-performing heuristics implement equal weighting: They assume that the opponent is equally likely to play any of their actions (for D1, this holds for nondominated actions). Thus, L1 and D1 do not make specific bets about the behavior of their opponent. Furthermore, they both adhere to the principle of self-dominance and, therefore, never choose a clearly inferior action.

Single reason simplification pays off. MaxMax and MaxMin, which use only a single reason, are the next-best policies and are fairly robust to environmental properties. While their performance lags significantly compared with L1 and D1, they are significantly simpler than, yet still outperform, more complex strategies such as L2, L3, and NE.

Assuming higher levels of rationality leads to poorer performance. The payoffs for NE, L2, and L3 are very low across the majority of environments; their performance is only moderately better for small games with low payoff uncertainty. This is a consequence of the implicit assumptions of these policies that other players also exhibit relatively high levels of rationality. Therefore, if matched with simple policies that are not based upon strategic principles, such as MaxMax, this mismatch will lead to underperformance.

SocMax is not robust to the worst-case scenario. Despite its strong performance according to the Indifference criterion—particularly in harmonious environments—SocMax proves extremely fragile in terms of the Wald criterion for all game sizes and degrees of payoff uncertainty. It is outperformed by single-reason policies, which also dominate it in terms of computational simplicity.

Discordant Environment ($\rho = -0.5$)

Discordant environments change the map of ecological rationality significantly for the Wald criterion. The most robust policy is MaxMin, which is the best performer for all environments except for games with size 2, where NE dominates. Averaging over the whole map, the best performing policies after MaxMin are L3 and MaxMax, although the performance difference with respect to MaxMin is greater than 5%. The L1, D1, and L2 heuristics perform relatively well (but do not belong to the set of top performers) for small games of five actions or less across the whole range of payoff uncertainty, but are no longer robust in larger games. Random behavior and other simple heuristics such as MaxMax and SocMax outperform these heuristics in large games.
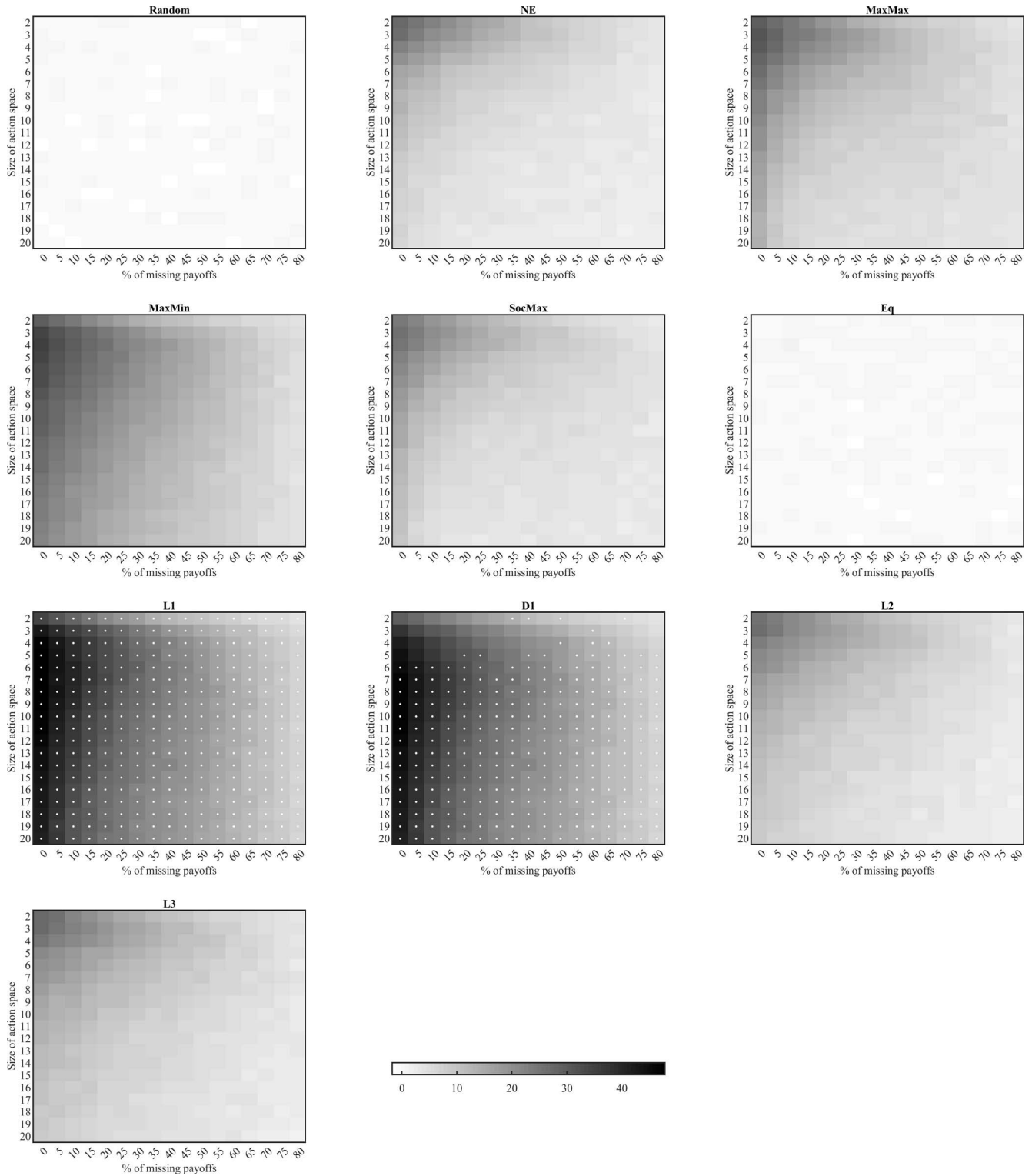
Figure 4. Decision policy performance according to the Wald criterion (π_w) for neutral environments ($\rho = 0$). The set of top-performing policies for each environment is marked by an overlaid white dot. NE = Nash equilibrium; L1 = Level-1; L2 = Level-2; L3 = Level-3; SocMax = social maximum; D1 = dominance-1; Eq = equality.
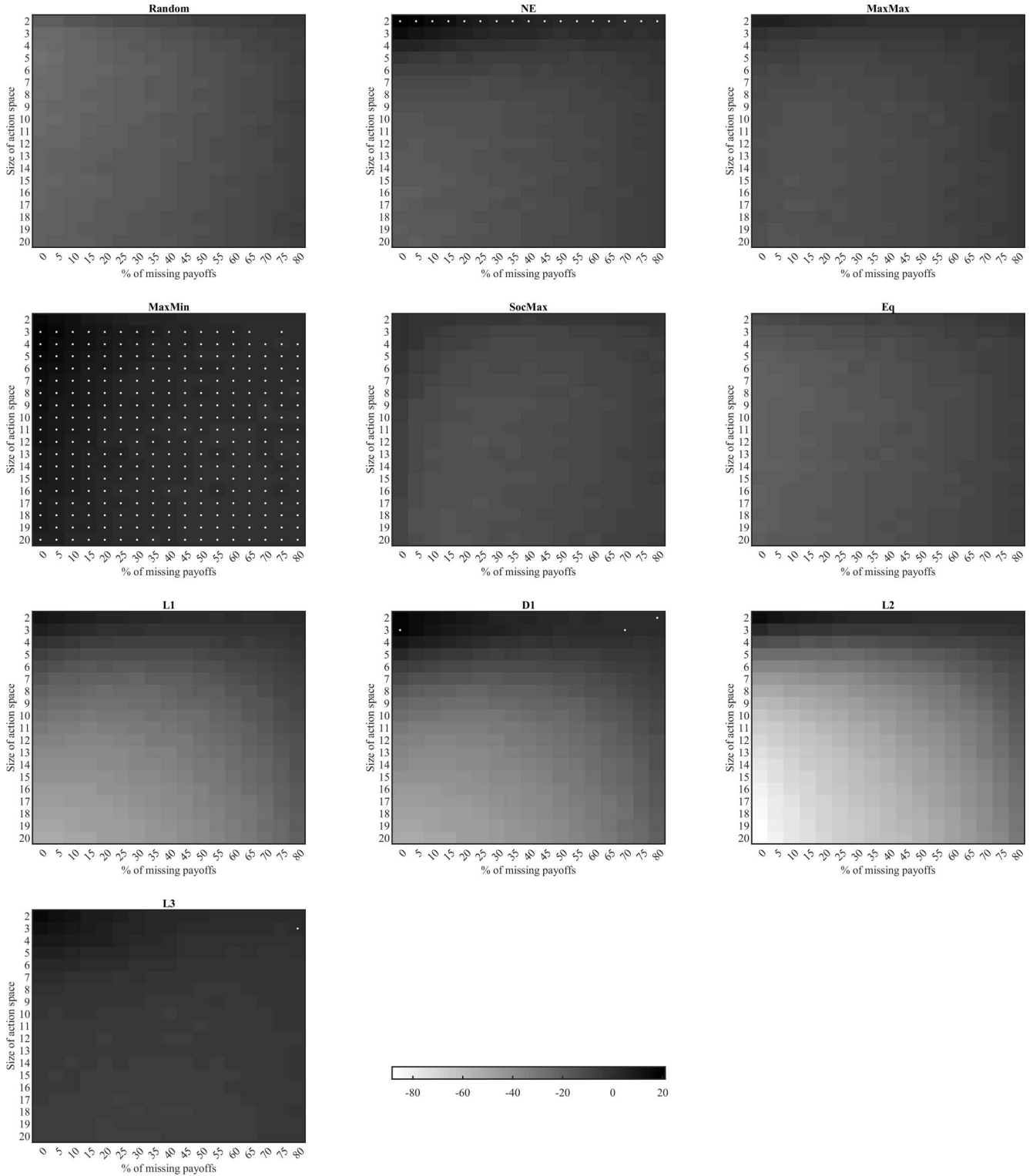
Figure 5. Decision policy performance according to the Wald criterion (π_w) for discordant environments ($\rho = -0.5$). The set of top-performing policies for each environment is marked by an overlaid white dot. NE = Nash equilibrium; L1 = Level-1; L2 = Level-2; L3 = Level-3; SocMax = social maximum; D1 = dominance-1; Eq = equality.

Figure 6. Decision policy performance according to the Wald criterion (π_w) for harmonious environments ($\rho =$ 0.5). The set of top-performing policies for each environment is marked by an overlaid white dot. NE = Nash equilibrium; L1 = Level-1; L2 = Level-2; L3 = Level-3; SocMax = social maximum; D1 = dominance-1; Eq = equality.
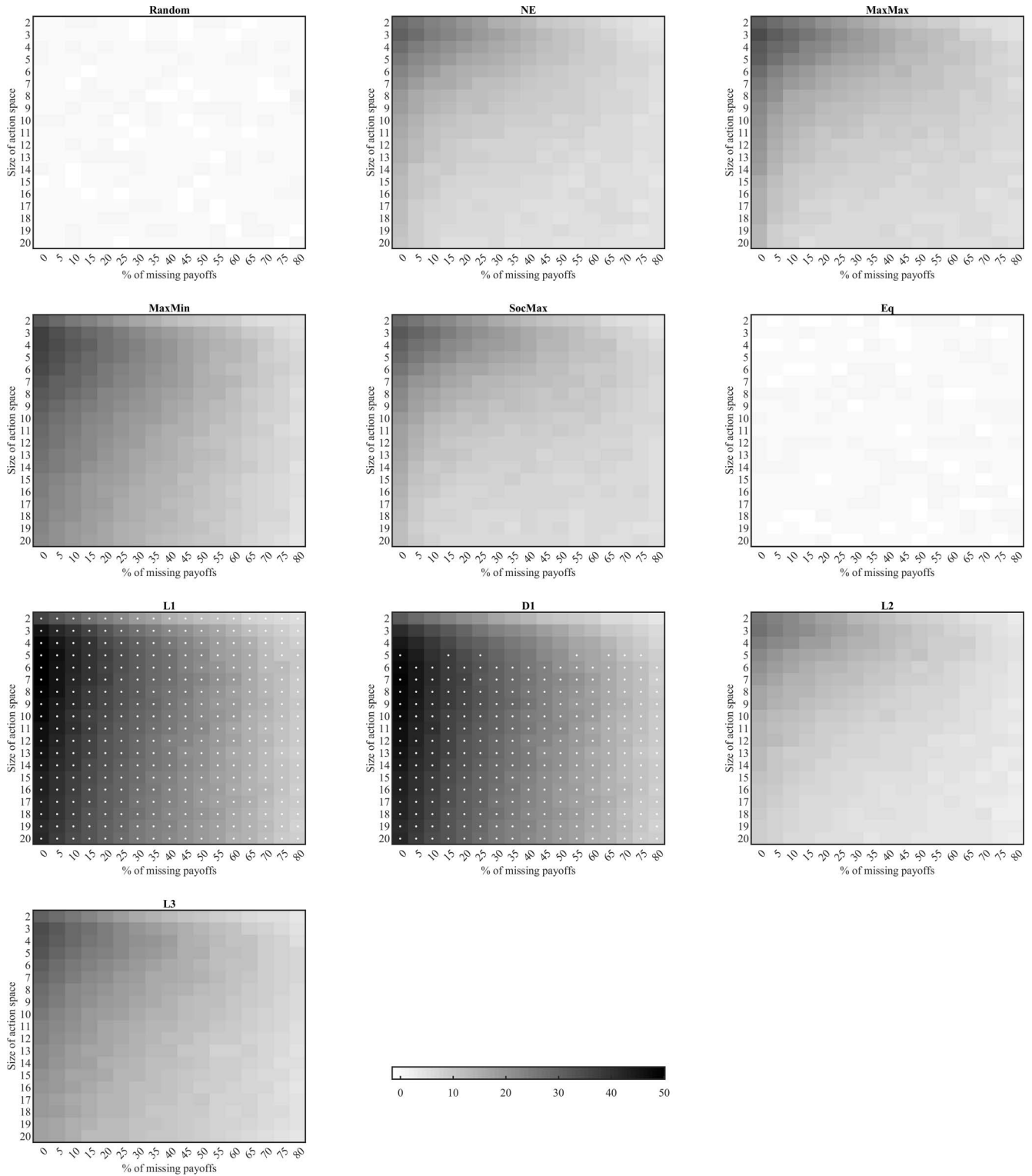
Harmonious Environment (ρ = 0.5)

As can be witnessed by comparing Figures 4 and 6, the performance of the decision policies is virtually identical to the neutral environment.

A Summary of Performance on the Composition-Robustness Criterion

We briefly summarize the findings according to this third criterion; details can be found in Appendix D. For neutral environments, L1 and D1 perform best, followed by the other Level-k policies. In harmonious environments, SocMax, L1 and D1 are the top performers, although SocMax dominates the niche associated with high payoff uncertainty ($m \geq 50\%$). While NE is in the set of top-performing policies only for $n = 2$, it still maintains a consistently high level of performance across the entire map. Finally, discordant environments significantly alter the performance landscape: D1 and L1 are the clear winners (with the former exhibiting greater robustness to game size), whereas NE performs very poorly. MaxMin and higher Level-k policies perform moderately well across the map, but exhibit different levels of robustness to strategic uncertainty conditional on the environmental niches.

General Discussion

Our analyses have shown that the surprising performance of heuristics observed in individual decision making does generalize to strategic choice as well. The conclusion, however, is not that heuristics are invariably good (or bad, for that matter); rather, judging the success of a heuristic necessitates asking under what environmental conditions they work and fail. We now consider implications of these results and outline extensions for further work.

Matching Heuristics and Environments

In the previous section we presented detailed maps outlining the ecological rationality of the 10 decision policies according to the performance criteria. What are the key findings from the simulations? Table 4 shows the performance, according to all three criteria, for every decision policy conditional on the degree of harmony of the environments, but averaged over all possible combinations of payoff uncertainty and game size. The columns marked "All ρ" present the performance averaged over all three environmental properties: n, m, ρ. According to the Indifference criterion, L1 and D1 are the top-performing policies except in harmonious environments, where they are clearly outperformed by SocMax and the NE; in the case of the latter, however, this is conditional on equilibrium selection based on the maximum joint payoffs. According to the Wald criterion, L1 and D1 are again the top-performing policies, except for discordant environments, where MaxMin is the best policy. The results for the Composition-Robustness criterion are very similar to those of the Indifference criterion. D1 is the best policy (followed by L1) in all environments with the exception of harmonious environments, where SocMax performs best. More detailed maps of ecological rationality across n and m for the Composition-Robustness criterion can be found in Appendix D. These findings are robust to the following three alternatives.

First, restricting the space of possible environments to a subset that consists of smaller games ($n \leq 10$) with less missing information ($m \leq 50\%$) yields similar conclusions—see Table 1 in the online Supplemental Material. If anything, for these environments L1 and D1 expand their dominance; the latter, in particular, belongs to the set of top performing rules for environments of all degrees of harmony and performance criteria, with the exception only of ρ = −0.5 for the Wald criterion.

Second, the general performance of L1 and D1 is further enhanced in simulation S' where players have perfect information about their own payoffs, but missing (incomplete) information about opponent payoffs—see Table 3 and Figures 1–9 in the online Supplemental Material. For the Indifference criterion, the NE policy belongs to the set of top performing rules only for ρ = 0.5 and for virtually zero missing opponent payoffs. For the Wald criterion, this occurs only for ρ = −0.5, $n = 2$ and very low missing information ($m \leq 15\%$). L1 and D1 are top-performers for almost all environments and performance criteria, with the exception of ρ = −0.5 and the Wald criterion. The intuition behind the improvement in L1 performance is that because it relies only on own payoffs, the heuristic is never led astray because it does not have to infer any missing payoffs in these environments. By

Table 4
Summary of Decision Policies' Performance According to the Performance Criteria

ρ =	Indifference				Wald				Comp.-Robustness			
	0	.5	−.5	All ρ	0	.5	−.5	All ρ	0	.5	−.5	All ρ
Random	.9	7.3	−4.3	1.3	−.6	−.3	−15.3	−5.4	5	8	5	6
NE	18.7	**43.1**	1.9	21.2	6.0	9.4	−8.7	2.2	54	74	33	54
MaxMax	14.2	32.2	5.4	17.3	9.2	10.8	−8.6	3.8	33	43	48	41
MaxMin	16.5	23.7	12.5	17.5	14.5	16.2	**3.8**	**11.5**	40	26	68	45
SocMax	20.4	**44.8**	3.4	22.8	7.1	10.7	−9.8	2.7	59	**80**	36	59
Eq	1.4	7.4	−3.6	1.7	−.4	−.3	−13.1	−4.6	13	10	15	12
L1	**25.7**	40.6	**16.0**	**27.4**	**22.2**	**25.0**	−23.5	7.9	79	71	79	76
D1	**26.1**	40.8	**16.4**	**27.7**	21.8	24.5	−21.6	8.2	**86**	71	**86**	**81**
L2	23.3	38.5	12.4	24.7	7.6	8.2	−39.1	−7.8	69	62	66	65
L3	21.6	37.3	12.1	23.6	8.0	13.9	−2.1	6.6	62	56	65	61

Note. NE = Nash equilibrium; L1 = Level-1; L2 = Level-2; L3 = Level-3; SocMax = social maximum; D1 = dominance-1; Eq = equality. The set of top-performing policies are in bold.

contrast, more sophisticated policies such as higher Level-k and NE will be more susceptible to the uncertainty in their opponent's payoffs.

Third, in simulation S''' where missing payoffs are inferred only by imputing the mean of the observed payoffs and not their variance or correlation, the qualitative results are virtually identical to those of the core simulation. Details are reported in the online supplemental materials, Table 5 and Figures 19–27.

Generally, the most complex decision policy, the NE, performed anywhere ranging from poorly to modestly with respect to all performance criteria, with the main exception occurring for the Indifference criterion in harmonious environments. In the few other cases where it was relatively competitive, this was typically for small games with minimal payoff uncertainty. This implies that it is not particularly robust to payoff uncertainty, strategic uncertainty, or the size of the games. We mentioned earlier that not all games have at least one pure-strategy NE and, therefore, the NE policy we have implemented reverts to random choice for these cases. That is, NE is not directly applicable for some games, given that we have ruled out mixed-strategy NE (in one-shot games) on the basis of their complexity, especially as the size of the game increases. How detrimental is this for the performance of NE? This depends on the proportion of games in an environment that will not have any pure-strategy NE. These statistics can be found in Table F1; the following numbers correspond to no payoff uncertainty, but the qualitative results are similar for varying payoff uncertainty. For neutral environments, this ranges from 0.125 ($n = 2$) to 0.35 ($n = 20$), for harmonious environments from 0.05 ($n = 2$) to 0.01 ($n = 20$), and for discordant environments 0.21 ($n = 2$) to 0.93 ($n = 20$).[20] For the latter, the probability of no pure-strategy NE rises quickly with the game size; it is already 0.55 for $n = 4$. This reveals another weakness of the NE policy, namely, that it cannot propose a specific course of action for a large proportion of the games in discordant environments and a moderate proportion of large games in the neutral environment. Therefore, the only alternative to random choice for these cases would be to resort to another—necessarily less complex—policy, constituting further evidence and an explanation for the frequent use of heuristics.[21] Consequently, part (but not all) of NE's poor performance can be attributed to this problem. For example, even in harmonious environments where this is not an issue, NE performs well according to the Indifference criterion but exhibits subpar performance according to the Wald criterion. In the online supplemental materials (see Table 2), we present the analogue of Table 4 for games with at least one pure-strategy NE, where our implementation of NE does not resort to random choice. Averaging over n and m, the NE policy still lags significantly behind L1 and D1 for the Wald criterion, but now belongs to the set of top performing decision rules for all degrees of harmony for the Indifference criterion. However, it is never the sole policy in the set of top performers. For $\rho = 0$ and -0.5, it shares this accolade with L1 and D1 and for $\rho = 0.5$ with SocMax. Consequently, it is always possible to achieve approximately the same level of performance as NE with significantly less complex heuristics. We note also that the games where the NE policy resorts to random choice provide a boost to the L1 and D1 policies because they perform well against random behavior. However, the above results based on games with pure-strategy NE confirm that the high performance of L1 and D1 is not crucially dependent on this—see also the high pairwise perfor-

mance of L1 and D1 against other policies independent of NE documented in Appendix E.

This brings us to two other disadvantages of NE related to its complexity that render it less desirable even in cases where its performance matches that of other heuristics in our simulations. The first is that if we account for the cognitive cost of implementing the NE, it would be at an even greater disadvantage compared with heuristics such as L1 and D1. It also has greater informational requirements than does L1, which requires only information about own payoffs (and not opponent payoffs). The second is that the complexity in implementing it will likely lead to more errors in applying this decision policy compared with simpler heuristics. In our simulations we have assumed that each of the 10 decision policies are implemented perfectly without error. Alternatively, the inclusion of the Random policy in our simulation can also be interpreted as a 10% likelihood that each of the other nine opponent policies are implemented with error resulting in a random choice. However, if implementation errors are more likely the more complex the decision policy is, this would put L3 and the NE at an even greater disadvantage compared with simple heuristics than our results suggest.

Ecological Rationality in the Environmental Niche of Laboratory Studies

The experimental literature from which we drew the decision policies predominantly used small games ($n \leq 5$) with no payoff uncertainty ($m = 0$) and nonnegative degrees of harmony ($\rho \geq 0$).[22] This corresponds to the upper-left edge in the heatmaps of Figures 1, 3, 4, and 6. If, as we propose, heuristic use in laboratory studies is at least partly driven by the ecological rationality of these heuristics, then there should be a correspondence between the simulated performance and the empirical prevalence of these policies in this niche. Recall from Table 3 that the most common policies were L1, D1, L2, and the NE. According to the Indifference criterion, there is a significant correspondence with the empirical prevalence, as L1, D1, and L2 exhibit similar high performance, followed by the NE. The latter's performance is typically at its highest in environments of small games with no to little uncertainty (i.e., the niche occupied by most experimental studies).

---

[20] The probability that no pure-strategy Nash equilibria exist in a randomly drawn game where $\rho = 0$ is given by $\sum_{k=0} (-1)^k \cdot \binom{n}{k}^2 \cdot k! \cdot (k)^{-k}$, where $k = n^2$ (Goldberg, Goldman, & Newman, 1968). Consequently, the probability of zero pure-strategy Nash equilibria for $n = 2, 3, 5, 10, 20$ is 0.125, 0.21, 0.28, 0.33, and 0.35, respectively. This tends to $1 - (1 - e^{-1}) = 0.368$ as $n \to \infty$ (Dresher, 1970; Goldberg et al., 1968). The expected number of pure-strategy Nash equilibria is 1, regardless of the size of the game (Powers, 1990).

[21] At least without resorting to mixed-strategy Nash equilibria, which as we have argued previously, are far too computationally complex to be descriptive models of one-shot behavior. The existence of a pure-strategy Nash equilibrium is less likely in neutral and discordant environments with large games. That is, the mixed-strategy Nash equilibrium's computational complexity is at a maximum exactly for the types of games where it would have to substitute for the lack of a pure-strategy Nash equilibrium; therefore, it is not a realistic alternative for these cases.

[22] Discordant environments are more likely to yield games with no pure-strategy Nash equilibria. Experiments using such games typically use repeated rather than one-shot interactions, which are outside of our domain of investigation as learning plays an important role.

Note also that many of the experimental studies use coordination games to examine equilibrium selection. These games are more likely to arise in harmonious environments, where the NE policy belongs to the set of top-performers. Our findings predict that the empirical prevalence of NE behavior should fall significantly if experiments draw games outside of this niche. According to the Wald criterion, L1 and D1 are the best performers for neutral and harmonious environments, consistent with the empirical prevalence of these heuristics.

Ultimately, which of the two criteria are more relevant for participants' policy choice depends on the player's degree of uncertainty aversion. The Wald criterion is an extreme case of uncertainty aversion, whereas the Indifference criterion would correspond to uncertainty neutrality. Empirical evidence points to players exhibiting uncertainty aversion (Camerer & Karjalainen, 1994; Eichberger, Kelsey, & Schipper, 2008; Ellsberg, 1961; Pulford & Colman, 2007), albeit not as extreme as that implied by the Wald criterion. Therefore, a convex combination of the Indifference and Wald criteria is a reasonable solution—a policy must perform relatively well according to both criteria. This rules out MaxMin, which performs well on the Wald criterion but is overly pessimistic with respect to the Indifference criterion. As expected, the most prevalent policies, L1 and D1, perform well according to both criteria. Similarly, the least prevalent policies, such as Eq and MaxMax, do not perform consistently well. From this one can cautiously conclude that a notable correspondence between policy performance and empirical prevalence exists, consistent with our argument that ecological—rather than constructivist—rationality (V. L. Smith, 2003) plays a role in policy selection. In the future, it would be desirable for more experimental work to be directed toward broadening the scope of empirical findings to other regions of the environmental map, particularly with respect to payoff uncertainty. Our findings make clear predictions about how prevalent policies should be in a broader range of environments based on our mapping of their ecological rationality.

Sources of Heuristics' Robustness

Next, let us briefly discuss what properties of heuristics foster their robustness across payoff and strategic uncertainty. In particular, we consider two properties: adherence to dominance principles and resilience (consistency) against payoff uncertainty.

Adherence to dominance principles. Adherence to dominance principles plays an important role in a policy's robustness to strategic uncertainty. Dominance principles do not in any way depend on the behavior of the opponent or a player's beliefs thereof. It is always beneficial to avoid a self-dominated strategy because such a strategy is by definition inferior to the dominating strategy for all possible actions available to the opponent. Therefore, this property is completely independent and robust to strategic uncertainty. Of the decision policies under investigation, L1 adheres to the principle of self-dominance but not opponent-dominance. D1, L2, L3, and NE adhere to both self- and opponent-dominance principles. Whether the assumption of opponent-dominance is appropriate depends on the policy used by the opponent. It is, for instance, inappropriate for opponent decision policies that do not obey self-dominance, such as SocMax or Eq.

Consistency. One source of a policy's robustness to payoff uncertainty is the effect of the latter on the action chosen. That is,

given the same game, is the choice prescribed by a policy the same when uncertainty is introduced and missing payoffs are imputed? We define consistency of a policy as the percentage of choices made across games that are identical to the choices that would have been made had there been no payoff uncertainty ($m = 0$).

Consistency is fostered by two paths of simplification. First, heuristics using Path 1 (ignoring the opponent's payoffs) benefit from their independence with respect to opponent payoffs; inconsistency arises only from own payoff uncertainty. Thus, L1, MaxMin, and MaxMax stand to benefit from this simplification path. Second, heuristics using Path 2 (equal weighting) exhibit high consistency as decisions are based on comparisons of averages of payoffs per action, which are less susceptible to the noise introduced by each individual imputed payoff. Consequently, we expected L1 to benefit more than heuristics that are based on a single reason (e.g., MaxMax). Similarly, the best-response profiles associated with the NE are also fragile with respect to even a single, significantly different, imputed payoff, whether that be an own or opponent payoff.

These considerations are borne out by simulation results, presented in full in Table F2. Figure 7 reviews the consistency of selected policies of interest (based on their performance: L1, MaxMin, SocMax, and NE) averaged over all possible game sizes, but for each level of missing payoffs. The NE's consistency is presented separately for harmonious, neutral, and discordant environments; the other policies' consistency does not significantly change with ρ, because they use only own payoffs. L1 (and by association D1) exhibits high consistency across different levels of payoff uncertainty, as does MaxMin, for all degrees of harmony ($\rho = 0, 0.5, -0.5$). By contrast, NE exhibits lower consistency over all levels of payoff uncertainty and degrees of harmony than L1 and MaxMin. NE has particularly low consistency for discordant environments, as it falls below 20% with only 20% missing payoffs. In general, as the game size increases, the NE exhibits significantly lower consistency than the majority of other decision policies.

The importance of consistency becomes even more evident in simulation S'' that allows for differences in expectation of the average quality of actions. Recall, that in this case the payoffs for each action are drawn from an action-specific distribution, whose mean is determined by a random draw; therefore, payoffs exhibit greater relative variability across actions than in the core simulation. L1 and D1 perform even better in relative terms compared with other heuristics and NE because of the introduction of positive correlation within each action's payoffs; see Table 4 and Figures 10–18 in the online supplemental materials. This renders specific beliefs about an opponent's behavior significantly less useful, because it is more likely that regardless of the opponent's choice, the payoff will be high anyway. Consequently, policies that ignore opponent payoffs and assume equal-weighting are losing very little information, and the probability that L1 (and D1) will make the same choices as NE increases. Another way of thinking about this is that in these new environments, the probability of a dominant and dominated actions increase significantly compared with the core simulation where actions' payoffs have the same expectation. Therefore, L1 and D1, which obey own-dominance, can exploit this even more efficiently than in the core simulation. Should we then expect L1, D1, and NE (and higher Level-k heuristics) to converge in performance across the whole map of environments? This is only the case for environments without payoff uncertainty. As missing information increases, then the more
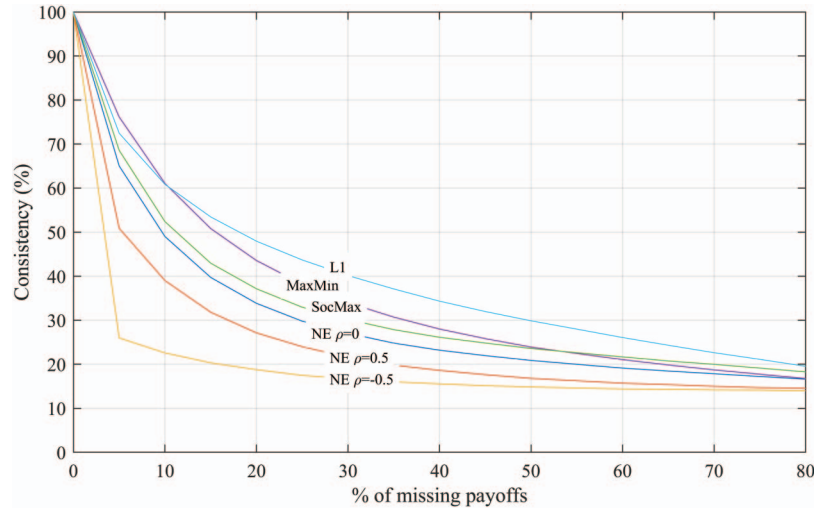
Figure 7. Consistency (%) for selected policies and environments. See the online article for the color version of this figure.

sophisticated policies NE, L2, and L3 will still fall prey to their relatively lower consistency compared with L1, leading them to choose less desirable actions with greater likelihood. Hence, their performance should degrade significantly more quickly with increases in *m*. This predicted divergence in performance is exactly what is found in simulation *S″*—details can be found in the online supplemental materials.

The Robust Beauty of the L1 Heuristic

The L1 heuristic deserves special mention, as it not only uses a probability-based simplification, equal weighting, but also a payoff-based simplification (i.e., completely ignoring the opponent's payoff information). Despite this, it still obeys self-dominance and exhibits high consistency. Because of these characteristics, L1 performed very well according to both the Indifference criterion and the Wald criterion over a large range of environments, with the exception of the Wald criterion in a discordant environment. It also outperformed other decision policies in harmonious and neutral environments according to the Composition-Robustness criterion, thereby exhibiting notable robustness to strategic uncertainty (see Appendix D). There is, indeed, a robust beauty (Dawes, 1979, p. 571) to this improper unit-weight decision policy not just in predictive inference but also in strategic choice. Finally, the closely related D1 also performed very well, sacrificing the payoff-based path to simplification to satisfy opponent-dominance. Opponent-dominance can improve the performance of D1 over the simpler L1 in environments of small games if there are enough opponents who do not play dominated strategies. As we showed above, increases in the variance of the expected quality of actions, render L1 and D1 even more robust because of their high consistency in the face of payoff uncertainty and adherence to self-dominance.

Learning and Dynamics

Our analysis has emphasized the one-shot nature of many interactions where players are randomly matched with opponents,

which ruled out repeated game strategies such as those in the Axelrod (1984) tournaments, and where the degree of strategic and environmental uncertainty made learning extremely difficult. Our belief is that this is a relatively underresearched class of environments. However, the existence of environments where enough experience may be accumulated to put decision policies under some form of selection pressure should not be ruled out. Would our findings carry over to a setting where decision policies are subject to some form of selection based on their historic performance? Dynamics are typically predicated on the simple assumption that better performing policies should become more prevalent over time either through an evolutionary (J. M. Smith, 1982) or learning process (Fudenberg & Levine, 1998). While the general principle of better performing policies becoming more prevalent transcends the different types of dynamics, the exact specification of these dynamics, the information that influences them, and auxiliary assumptions (such as the initial distribution of decision policies) can affect the outcomes. To precisely map this is a difficult and complex endeavor requiring significantly more space than we can devote to it here. However, below we lay out some broad intuition about the important questions and general principles regarding the dynamic behavior of the decision policies and implications of various dependencies. We conjecture that L1 and D1 will continue to play an important role in many, though not all, environments, even when dynamics are accounted for.

Consider a general case where policies compete over many rounds against a population of players in the set of games belonging to an environment—this is commensurate both with the replicator dynamics often used in the evolutionary game theory literature and also with reinforcement learning. Given the high performance of L1 and D1 according to the Indifference criterion—and more importantly, the Composition-Robustness criterion—we expect that these policies will be more likely to replicate in neutral and discordant environments for the majority of the possible trajectories of the composition of opponents over the rounds. There will still be some combinations of initial population distributions

that may push the dynamics into regions where these policies may be outperformed. However, policies that are highly robust to strategic uncertainty, especially with respect to the Composition-Robustness criterion (and to a lesser degree the Wald criterion), will be more likely to replicate strongly under the majority of circumstances.

Is there a possibility for policies that are less robust to strategic uncertainty to dominate the population? Yes, if the dynamics lead to a state where a policy is matched often enough with its ideal (or near ideal) opponent(s). In our analyses SocMax and NE are examples of policies whose dynamic behavior will be more strongly dependent on the evolution of the population. Recall that SocMax performs considerably well according to the Indifference criterion but not the Wald criterion in $\rho = 0$ and $\rho = 0.5$ environments (as does NE in the $\rho = 0.5$ environment). Examining this further reveals that SocMax performs extremely well when playing against itself—in fact, in most cases it is a best response to itself, as is also the case for NE—see Tables E1–E3 in Appendix E. Therefore, if at any stage of the dynamics (including the initial population) the prevalence of SocMax or NE reaches a critical mass, then it will lead to a self-sustaining takeover of the whole population.

Given that dynamics favor better performing policies, a crucial question is whether complex heuristics capable of higher levels of rationality or sophistication are more likely to thrive. Stahl (1993) is an instructive study on the evolution of players with different levels of sophistication based on their beliefs about opponents and degree of rationalizability. Players who were rational in the normative sense did not necessarily survive in the long run, especially if costs were associated with more sophisticated reasoning. Under many conditions, minimally sophisticated players were not driven out of the population entirely. Ultimately, this depends both on the population composition and on the type of game being played, as more sophisticated policies do not necessarily perform better against naive heuristics; recall the difference between games of strategic substitutes and strategic complements (Camerer & Fehr, 2006). Although Level-k policies by definition best-respond to Level-(k-1) policies, it is not self-evident that a player will receive higher payoffs than if they adopted a lower level of sophistication. On the one hand, in environments with a large proportion of games with strategic substitutes players would have an incentive to switch to higher-level reasoning; therefore, L2 will be more likely to emerge because of the fact that it is a best response to L1. Similarly, once the L2 population becomes significant this may encourage L3 and so on, ultimately encouraging even the NE policy in the long run; over time, one would expect to see successively higher levels of rationality in the population. However, this analysis ignores the computational complexity of the decision policies; the final outcome will depend on the strength of the effect of the strategic substitutes versus the degree that the computational complexity of a policy penalizes its performance. On the other hand, in games with strategic complements, an emergence of L1 is more likely to be self-sustaining; higher-level policies will not replace it because there are no payoff incentives for higher-order reasoning.

An alternative approach to modeling the dynamics of policy use is based on learning rather than evolution. One learning approach that has the advantage of not requiring the prior specification of players' policies models them as tabula rasa neural networks,

whose connection weights are adjusted in line with the myopic ex-post best response after each game (i.e., the best response to an opponent's realized play in the immediately prior round). Spiliopoulos (2011, 2012a, 2015) investigated the learning dynamics of such neural network populations in environments of size 2 and 3, with no missing payoffs and zero degree of harmony. This analysis concluded that despite not being prespecified, emergent behavior strongly approximating Level-k heuristics emerged over time, with the dynamics we predicted above. These findings match our analysis for small games. Another approach uses reinforcement learning over decision policies, or rule-based learning (Stahl, 1999, 2000). That is, estimates of the value of the decision policies arise from their use and resulting payoffs; the higher the valuation of a policy the greater the likelihood of its future use. Because players must explicitly use decision policies to gain information about their performance, there is an exploration—exploitation tradeoff. This would favor simple, robust policies that perform well against many different types of opponents in the exploration phase, such as L1 and D1.

In conclusion, based on these considerations we predict that L1 and D1 will retain their importance in many of the environments that we have studied even when dynamics are introduced, with the possibility that SocMax and NE will dominate some environments under specific conditions. However, detailed maps of rationality incorporating dynamics will be dependent on various conditions that we have touched upon briefly above. One of these is the inclusion of penalties for decision policies' complexity, which would be a significant advantage for D1 and, in particular L1, which uses only own payoff information and relatively simple computational processes. A thorough exploration of the dynamical evolution of our set of policies deserves to be high on the agenda.

Conclusion

Experimental studies of strategic interactions have found significant evidence for the frequent use of heuristics (see Table 3). Our results cast new light on this finding. The established narrative attributes heuristic use to humans' inability to reason according to the complex NE. Alternately, it has been argued that people still optimize their choice of policy, but subject to a constraint based on the decision costs, which leads to the use of simpler policies. This argument is based on the accuracy-effort (speed) trade-off, often seen as a general law of cognition: Those who invest less mental effort will pay a price in terms of lower accuracy (performance). From this perspective, heuristics are, by definition, always second-best; given unlimited resources, more computation and more time would always be better.

We conducted a systematic investigation of how well decision policies perform by varying the degree of payoff uncertainty, the strategic uncertainty, the complexity of a game (in terms of the number of actions available to each player), and the degree of common (or antithetical) player interests in environments' games. Our findings challenge arguments that simple rules will falter in competitive interactions (e.g., Sterelny, 2003). In parallel to work in inference or individual decision making (Gigerenzer et al., 2011), we found that less complex decision policies—heuristics—were more robust to strategic and payoff uncertainty than complex decision policies. Moreover, the heuristics that are more commonly used by participants in experimental studies of games (such

as L1 and D1) were also found to be ecologically rational policies for the environments typically implemented in these experiments.

Crucially, we have shown that in large-world strategic environments, abandoning normative axioms and Bayesian principles does not necessarily imply irrational behavior. In this respect, our findings are congruent with the beliefs of many prominent game theorists who are, perhaps surprisingly, partial to heuristics. Reinhard Selten had a keen sense of the differences between normative game theoretic models and descriptive models of behavior and was a strong advocate for heuristics (Gigerenzer & Selten, 2001, see also a discussion of his work in Nagel, Bayona, Kheirandish, & Mousavi, 2016). Similarly, R. Aumann (1997) wrote, "ordinary people do not behave in a consciously rational way in their day-to-day activities. Rather, they evolve 'rules of thumb' that work in general, by an evolutionary process like that discussed above (Section 1a), or a learning process with similar properties" (pp. 7–8). Hart (2005) made similar arguments for repeated games: "From a more general viewpoint, the results show how simple and far-from-rational behavior [adaptive heuristics based on the avoidance of regret] in the short run may well lead to fully rational outcomes in the long run" (p. 1415). Simon (1957) shared similar views: "He [Administrative man] makes his choices using a simple picture of the situation that takes into account just a few of the factors that he regards as most relevant and crucial" (p. xxv).

Two paths of simplification were particularly successfully: equal weighting and, surprisingly, ignoring an opponent's payoff (that constitutes the strategic component of the game). Equal weighting ensures that a heuristic will not choose a dominated strategy, which would be suboptimal regardless of the type of opponent. Ignoring an opponent's payoffs constrained a policy from forming overly specific beliefs regarding an opponent's likely behavior. Such beliefs would be detrimental insofar as their formation does not acknowledge the inherent strategic uncertainty—in short, less can be more. Simple heuristics such as L1, which adopts both paths of simplification, do not necessarily sacrifice performance compared with the NE. Furthermore, we showed that simple heuristics are more consistent than complex policies in the face of environmental uncertainty such as missing payoff information. That is, they are more likely to recommend the same course of action as they would have done if the game payoffs were perfectly known.

Map of Ecological Rationality

Undertaking a systematic investigation of the performance of decision policies for different environments has revealed a map of ecologically rational behavior, as defined by Gigerenzer et al. (1999) and V. L. Smith (2003). The equal-weighting heuristics L1 and D1, which attach the same likelihood to each of an opponent's possible actions (the latter after removing the opponent's dominated actions), showed consistently high performance across performance criteria and robustness to payoff uncertainty. In terms of the Indifference criterion, L1 and D1 were the top performers in neutral and discordant environments. By contrast, the (pure-strategy) NE policy was solely relatively competitive for harmonious environments, according to the Indifference criterion only. However, in these environments the NE policy suffers from the problem of multiple pure-strategy equilibria, whereas the set of heuristics all prescribe a single course of action. For neutral and discordant environments, the NE policy performed well only for the kind of games typically used in experimental investigations of strategic behavior: complete knowledge of payoff information and game sizes of roughly two to five actions. The subpar performance of the NE policy, particularly for these two environments is partially driven by the fact that a high proportion of games in these environments admits no pure-strategy NE. Focusing only on the subset of games exhibiting at least one pure-strategy NE, the NE policy often performed on par with L1 and D1 according to the Indifference, but not the Wald, criterion. However, this is achieved at the cost of significantly more complexity than the L1 and D1 heuristics. Furthermore, its inapplicability to games without pure-strategy NE implies that players would have to resort to heuristics or mixed-strategy NE—the latter are very unlikely in one-shot games given their high computational complexity. The relatively simple SocMax heuristic was the best performer in harmonious environments (followed by NE), but its relatively precise assumption that an opponent will also seek to maximize the joint payoffs to both players leaves it vulnerable to exploitation by certain policies. In terms of the Wald criterion, which measures a policy's robustness to extreme strategic uncertainty, L1 and D1 clearly dominated all others by a wide margin in neutral and harmonious environments, while the even simpler MaxMin heuristic dominated in discordant environments.

Processing Errors and the Risk of Oversimplification

In many respects, our analysis may have actually underestimated the performance and applicability of simple decision policies versus more complex policies. First, we have assumed perfect execution of each decision policy; in practice, however, this would be dependent on the complexity of executing a decision policy. Therefore, the NE and Level-k policies where k ≥ 2 would be more susceptible to processing errors. Second, a complex decision policy would require more time to execute: A decision maker using a heuristic would suffer a smaller opportunity cost of time and would handle time pressure more gracefully than one using a complex or normative policy. However, let us also emphasize the limits of oversimplification. Decision policies that base decisions on a single piece of information without any weighting, such as MaxMax, typically performed worse than other heuristics, with the exception of MaxMin in terms of the Wald criterion. Decision policies that implicitly bet on an opponent choosing a specific action with certainty typically assumed too little strategic uncertainty, and often paid a high price. The NE, which assumes both the rationality of the opponent and common knowledge of this rationality, is a case in point.

Our results challenge the common wisdom that social complexity necessitates cognitive complexity (e.g., Humphrey, 1976, 1988; Whiten & Byrne, 1988; but see Hertwig & Herzog, 2009) because of the additional uncertainty introduced by strategic intentions. We extended similar findings from the repeated games literature, where simple strategies such as tit-for-tat or win–stay, lose–shift perform remarkably well, to the one-shot game literature. Simpler decision policies such as Level-k heuristics—in particular, L1 and D1—do not inevitably sacrifice performance relative to a policy choosing a pure-strategy NE. In fact, they can achieve high performance and robustness in the face of substantial environmental and strategic uncertainty. However, as we have pointed out, this does not imply that there are no limits to the benefits of simplicity. Effective paths to simplification must match the environmental properties to be ecologically rational. We argue that the heuristics that people predominantly use—instead of (pure-strategy) NE—in one-shot games are indeed ecologically

rational for realistic environments, where significant environmental and strategic uncertainty coexist.

References

Arrow, K. J. (2004). Is bounded rationality unboundedly rational? Some ruminations. In M. Augier & J. G. March (Eds.), *Models of a man: Essays in memory of Herbert A. Simon* (pp. 47–55). Cambridge, MA: MIT Press.

Aumann, R. (1997). Rationality and bounded rationality. *Games and Economic Behavior, 21,* 2–14. http://dx.doi.org/10.1006/game.1997.0585

Aumann, R., & Brandenburger, A. (1995). Epistemic conditions for Nash equilibrium. *Econometrica, 63,* 1161–1180. http://dx.doi.org/10.2307/2171725

Aumann, R. J., Katznelson, Y., Radner, R., Rosenthal, R. W., & Weiss, B. (1983). Approximate purification of mixed strategies. *Mathematics of Operations Research, 8,* 327–341.

Avis, D. (2000). A revised implementation of the reverse search vertex enumeration algorithm. In G. Kalai & G. Ziegler (Eds.), *Polytopes—Combinatorics and computation* (pp. 177–198). Basel: Birkhäuser. http://dx.doi.org/10.1007/978-3-0348-8438-9_9

Axelrod, R. (1984). *The evolution of cooperation.* New York, NY: Basic Books.

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science, 211,* 1390–1396. http://dx.doi.org/10.1126/science.7466396

Bernheim, B. D. (1984). Rationalizable strategic behavior. *Econometrica, 52,* 1007–1028. http://dx.doi.org/10.2307/1911196

Binmore, K. G. (2007). *Does game theory work? The bargaining challenge.* Cambridge, MA: MIT Press.

Binmore, K. G., Swierzbinski, J., & Proulx, C. (2001). Does minimax work? An experimental study. *The Economic Journal, 111,* 445–464. http://dx.doi.org/10.1111/1468-0297.00645

Brocas, I., Carrillo, J. D., Wang, S. W., & Camerer, C. F. (2014). Imperfect choice or imperfect attention? Understanding strategic thinking in private information games. *The Review of Economic Studies, 81,* 944–970. http://dx.doi.org/10.1093/restud/rdu001

Brunswik, E. (1952). The conceptual framework of psychology. In *International encyclopedia of unified science* (pp. 656–760). Oxford: Chicago Press.

Byrne, R. W., & Whiten, A. (Eds.). (1988). *Machiavellian intelligence. Social expertise and the evolution of intellect in monkeys, apes, and humans.* New York, NY: Clarendon Press.

Camerer, C. F. (2003). *Behavioral game theory. Experiments in strategic interaction.* Princeton, NJ: Princeton University Press.

Camerer, C. F., Cachon, G., & Johnson, E. J. (2004). Thinking about attention in games: Backward and forward induction. In I. Brocas & J. Castillo (Eds.), *The psychology of economic decisions.* Oxford: Oxford University Press.

Camerer, C. F., & Fehr, E. (2006). When does "economic man" dominate social behavior? *Science, 311,* 47–52. http://dx.doi.org/10.1126/science.1110600

Camerer, C. F., Ho, T.-H., & Chong, J.-K. (2002). Sophisticated experience-weighted attraction learning and strategic teaching in repeated games. *Journal of Economic Theory, 104,* 137–188. http://dx.doi.org/10.1006/jeth.2002.2927

Camerer, C. F., Ho, T.-H., & Chong, J.-K. (2004). A cognitive hierarchy model of games. *The Quarterly Journal of Economics, 119,* 861–898. http://dx.doi.org/10.1162/0033553041502225

Camerer, C. F., Ho, T.-H., & Chong, J. K. (2015). A psychological approach to strategic thinking in games. *Current Opinion in Behavioral Sciences, 3,* 157–162. http://dx.doi.org/10.1016/j.cobeha.2015.04.005

Camerer, C. F., Johnson, E. J., Rymon, T., & Sen, S. (1993). Cognition and

framing in sequential bargaining for gains and losses. In A. Kirman & P. Tani (Eds.), *Frontiers of game theory* (pp. 27–47). Cambridge, MA: MIT Press.

Camerer, C. F., & Karjalainen, R. (1994). Ambiguity-aversion and non-additive beliefs in non-cooperative games: Experimental evidence. In B. Munier & M. J. Machina (Eds.), *Models and experiments in risk and rationality.* Dordrecht, Netherlands: Springer.

Chen, C.-T., Huang, C.-Y., & Wang, J. T.-y. (2018). A window of cognition: Eyetracking the reasoning process in spatial beauty contest games. *Games and Economic Behavior, 111,* 143–158. http://dx.doi.org/10.1016/j.geb.2018.05.007

Cheung, Y.-W., & Friedman, D. (1997). Individual learning in normal form games: Some laboratory results. *Games and Economic Behavior, 19,* 46–76. http://dx.doi.org/10.1006/game.1997.0544

Costa-Gomes, M. A., Crawford, V. P., & Broseta, B. (2001). Cognition and behavior in normal-form games: An experimental study. *Econometrica, 69,* 1193–1235. http://dx.doi.org/10.1111/1468-0262.00239

Costa-Gomes, M. A., & Weizsäcker, G. (2008). Stated beliefs and play in normal-form games. *The Review of Economic Studies, 75,* 729–762. http://dx.doi.org/10.1111/j.1467-937X.2008.00498.x

Crawford, V. P., Costa-Gomes, M. A., & Iriberri, N. (2013). Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications. *Journal of Economic Literature, 51,* 5–62. http://dx.doi.org/10.1257/jel.51.1.5

Daskalakis, C., Goldberg, P. W., & Papadimitriou, C. H. (2009). The complexity of computing a Nash equilibrium. *SIAM Journal on Computing, 39,* 195–259. http://dx.doi.org/10.1137/070699652

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist, 34,* 571–582. http://dx.doi.org/10.1037/0003-066X.34.7.571

Devetag, G. M., Di Guida, S., & Polonio, L. (2016). An eye-tracking study of feature-based choice in one-shot games. *Experimental Economics, 19,* 177–201. http://dx.doi.org/10.1007/s10683-015-9432-5

Dhami, M. K., Hertwig, R., & Hoffrage, U. (2004). The role of representative design in an ecological approach to cognition. *Psychological Bulletin, 130,* 959–988. http://dx.doi.org/10.1037/0033-2909.130.6.959

Dresher, M. (1970). Probability of a pure equilibrium point in n-person games. *Journal of Combinatorial Theory, 8,* 134–145. http://dx.doi.org/10.1016/S0021-9800(70)80015-1

Duersch, P., Oechssler, J., & Schipper, B. C. (2012). Unbeatable imitation. *Games and Economic Behavior, 76,* 88–96. http://dx.doi.org/10.1016/j.geb.2012.05.002

Eichberger, J., Kelsey, D., & Schipper, B. C. (2008). Granny versus game theorist: Ambiguity in experimental games. *Theory and Decision, 64,* 333–362. http://dx.doi.org/10.1007/s11238-007-9053-3

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *The Quarterly Journal of Economics, 75,* 643–669. http://dx.doi.org/10.2307/1884324

Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review, 88,* 848–881.

Fisher, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika, 10,* 507–521. http://dx.doi.org/10.2307/2331838

Fudenberg, D., & Levine, D. K. (1998). *The theory of learning in games.* Cambridge, MA: MIT Press.

Fudenberg, D., & Liang, A. (2017). *Predicting and understanding initial play.* PIER Working Paper No. 18-009. http://dx.doi.org/10.2139/ssrn.3076682

Gigerenzer, G., & Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science, 1,* 107–143.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103,* 650–669.

Gigerenzer, G., Hertwig, R., & Pachur, T. (Eds.). (2011). *Heuristics: The foundations of adaptive behavior*. New York, NY: Oxford University Press.

Gigerenzer, G., & Selten, R. (Eds.). (2001). *Bounded rationality: The adaptive toolbox*. Cambridge, MA: The MIT Press.

Gigerenzer, G., Todd, P. M., & the ABC Research Group. (1999). *Simple heuristics that make us smart*. New York, NY: Oxford University Press.

Gilboa, I., & Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics, 18,* 141–153. http://dx.doi.org/10.1016/0304-4068(89)90018-9

Gode, D. K., & Sunder, S. (1993). Allocative efficiency of markets with zero-intelligence traders—Market as a partial substitute for individual rationality. *The Journal of Political Economy, 101,* 119–137. http://dx.doi.org/10.1086/261868

Goldberg, K., Goldman, A. J., & Newman, M. (1968). The probability of an equilibrium point. *Journal of Research of the National Bureau of Standards-B. Mathematical Sciences, 728,* 93–101.

Harrison, G. W. (1989). Theory and misbehavior of first-price auctions. *American Economic Review, 79,* 749–762.

Harsanyi, J. C. (1967). Games with incomplete information played by "Bayesian" players, I–III Part I. The basic model. *Management Science, 14,* 159–182. http://dx.doi.org/10.1287/mnsc.14.3.159

Harsanyi, J. C. (1973). Games with randomly disturbed payoffs: A new rationale for mixed-strategy equilibrium points. *International Journal of Game Theory, 2,* 1–23.

Harsanyi, J. C., & Selten, R. (1988). *A general theory of equilibrium selection in games*. Cambridge, MA: MIT Press Books.

Hart, S. (2005). Adaptive heuristics. *Econometrica, 73,* 1401–1430. http://dx.doi.org/10.1111/j.1468-0262.2005.00625.x

Hertwig, R., & Erev, I. (2009). The description-experience gap in risky choice. *Trends in Cognitive Sciences, 13,* 517–523. http://dx.doi.org/10.1016/j.tics.2009.09.004

Hertwig, R., & Herzog, S. M. (2009). Fast and frugal heuristics: Tools of social rationality. *Social Cognition, 27,* 661–698.

Hertwig, R., Hoffrage, U., & the ABC Research Group. (2013). *Simple heuristics in a social world*. New York, NY: Oxford University Press.

Hertwig, R., Pleskac, T. J., Pachur, T., & the Center for Adaptive Rationality. (2019). *Taming uncertainty*. Cambridge, MA: MIT Press.

Hogarth, R. M., & Karelaia, N. (2006). Regions of rationality: Maps for bounded agents. *Decision Analysis, 3,* 124–144. http://dx.doi.org/10.1287/deca.1060.0063

Humphrey, N. K. (1976). The social function of intellect. In P. P. G. Bateson & R. A. Hinde (Eds.), *Growing points in ethology* (pp. 303–317). Cambridge: Cambridge University Press.

Humphrey, N. K. (1988). The social function of intellect. In R. W. Byrne (Ed.), *Machiavellian intelligence: Social expertise and the evolution of intellect in monkeys, apes, and humans*. Oxford: Clarendon Press.

Imhof, L. A., Fudenberg, D., & Nowak, M. A. (2007). Tit-for-tat or win-stay, lose-shift? *Journal of Theoretical Biology, 247,* 574–580. http://dx.doi.org/10.1016/j.jtbi.2007.03.027

Johnson, E. J., Camerer, C. F., Sen, S., & Rymon, T. (2002). Detecting failures of backward induction: Monitoring information search in sequential bargaining. *Journal of Economic Theory, 104,* 16–47. http://dx.doi.org/10.1006/jeth.2001.2850

Kareev, Y. (1995). Through a narrow window: Working memory capacity and the detection of covariation. *Cognition, 56,* 263–269. http://dx.doi.org/10.1016/0010-0277(95)92814-G

Kareev, Y. (2000). Seven (indeed, plus or minus two) and the detection of correlations. *Psychological Review, 107,* 397–402. http://dx.doi.org/10.1037/0033-295X.107.2.397

Kareev, Y., Lieberman, I., & Lev, M. (1997). Through a narrow window: Sample size and the perception of correlation. *Journal of Experimental Psychology: General, 126,* 278–287. http://dx.doi.org/10.1037/0096-3445.126.3.278

Katsikopoulos, K. V., Schooler, L. J., & Hertwig, R. (2010). The robust beauty of ordinary information. *Psychological Review, 117,* 1259–1266. http://dx.doi.org/10.1037/a0020418

Knoepfle, D. T., Wang, J. T.-y., & Camerer, C. F. (2009). Studying learning in games using eye-tracking. *Journal of the European Economic Association, 7,* 388–398. http://dx.doi.org/10.1162/JEEA.2009.7.2-3.388

Luce, R. D., & Raiffa, H. (1957). *Games and decisions. Introduction and critical survey*. New York, NY: Wiley.

Mailath, G. J., & Samuelson, L. (2006). *Repeated games and reputations: Long-run relationships*. Oxford: Oxford University Press.

McLennan, A. (2005). The expected number of Nash equilibria of a normal form game. *Econometrica, 73,* 141–174. http://dx.doi.org/10.1111/j.1468-0262.2005.00567.x

McLennan, A., & Berg, J. (2005). Asymptotic expected number of Nash equilibria of two-player normal form games. *Games and Economic Behavior, 51,* 264–295. http://dx.doi.org/10.1016/j.geb.2004.10.008

Nagel, R., Bayona, A., Kheirandish, R., & Mousavi, S. (2016). Reinhard Selten, the dualist. In R. Frantz, S.-H. Chen, K. Dopfer, F. Heukelom, & S. Mousavi (Eds.), *Routledge handbook of behavioral economics* (pp. 67–88). Abingdon: Routledge.

Nowak, M., & Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature, 364,* 56–58. http://dx.doi.org/10.1038/364056a0

Ochs, J. (1995). Games with unique, mixed strategy equilibria: An experimental study. *Games and Economic Behavior, 10,* 202–217. http://dx.doi.org/10.1006/game.1995.1030

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge: Cambridge University Press.

Pearce, D. G. (1984). Rationalizable strategic behavior and the problem of perfection. *Econometrica, 52,* 1029–1050. http://dx.doi.org/10.2307/1911197

Peterson, C. R., & Beach, L. R. (1967). Man as an intuitive statistician. *Psychological Bulletin, 68,* 29–46.

Pleskac, T. J., & Hertwig, R. (2014). Ecologically rational choice and the structure of the environment. *Journal of Experimental Psychology: General, 143,* 2000–2019. http://dx.doi.org/10.1037/xge0000013

Polonio, L., & Coricelli, G. (2019). Testing the level of consistency between choices and beliefs in games using eye-tracking. *Games and Economic Behavior, 113,* 566–586. http://dx.doi.org/10.1016/j.geb.2018.11.003

Polonio, L., Di Guida, S., & Coricelli, G. (2015). Strategic sophistication and attention in games: An eye-tracking study. *Games and Economic Behavior, 94,* 80–96. http://dx.doi.org/10.1016/j.geb.2015.09.003

Powers, I. Y. (1990). Limiting distributions of the number of pure strategy Nash equilibria in n-person games. *International Journal of Game Theory, 19,* 277–286. http://dx.doi.org/10.1007/BF01755478

Pulford, B. D., & Colman, A. M. (2007). Ambiguous games: Evidence for strategic ambiguity aversion. *The Quarterly Journal of Experimental Psychology, 60,* 1083–1100. http://dx.doi.org/10.1080/17470210600866354

Rapoport, A. (1975). Research paradigms for studying dynamic decision behavior. In D. Wendt & C. Vlek (Eds.), *Utility, probability, and human decision making* (pp. 349–369). http://dx.doi.org/10.1007/978-94-010-1834-0_22

Rapoport, A., Guyer, M., & Gordon, D. G. (1976). *The 2 × 2 game*. Ann Arbor, MI: University of Michigan Press.

Rapoport, A., Seale, D. A., & Colman, A. M. (2015). Is tit-for-tat the answer? On the conclusions drawn from Axelrod's Tournaments. *PLoS ONE, 10.* http://dx.doi.org/10.1371/journal.pone.0134128

Rendell, L., Boyd, R., Cownden, D., Enquist, M., Eriksson, K., Feldman, M., . . . Laland, K. (2010). Why copy others? Insights from the social learning strategies tournament. *Science, 328,* 208–213. http://dx.doi.org/10.1126/science.1184719

Rey-Biel, P. (2009). Equilibrium play and best response to (stated) beliefs in normal form games. *Games and Economic Behavior, 65,* 572–585. http://dx.doi.org/10.1016/j.geb.2008.03.003

Roth, A. E., & Erev, I. (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior, 8,* 164–212. http://dx.doi.org/10.1016/S0899-8256(05)80020-X

Rubinstein, A. (1991). Comments on the interpretation of game theory. *Econometrica, 59,* 909–924. http://dx.doi.org/10.2307/2938166

Sargent, T. J. (1993). *Bounded rationality in macroeconomics*. New York, NY: Oxford University Press.

Savage, L. J. (1954). *The foundation of statistics*. New York, NY: Wiley and Sons.

Savani, R., & Von Stengel, B. (2006). Hard-to-solve bimatrix games. *Econometrica, 74,* 397–429. http://dx.doi.org/10.1111/j.1468-0262.2006.00667.x

Schelling, T. C. (1960). *The strategy of conflict*. Cambridge, MA: Harvard University Press.

Selten, R. (1975). Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory, 4,* 25–55. http://dx.doi.org/10.1007/BF01766400

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics, 69,* 99–118. http://dx.doi.org/10.2307/1884852

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review, 63,* 129–138. http://dx.doi.org/10.1037/h0042769

Simon, H. A. (1957). *Administrative behavior: A study of decision-making processes in administrative organization*. New York, NY: The Macmillan Company.

Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology, 41,* 1–19.

Smith, J. M. (1982). *Evolution and the theory of games*. Cambridge: Cambridge University Press.

Smith, V. L. (2003). Constructivist and ecological rationality in economics. *American Economic Review, 93,* 465–508. http://dx.doi.org/10.1257/000282803322156954

Spiliopoulos, L. (2011). Neural networks as a unifying learning model for random normal form games. *Adaptive Behavior, 19,* 383–408. http://dx.doi.org/10.1177/1059712311417636

Spiliopoulos, L. (2012a). Interactive learning in 2 × 2 normal form games by neural network agents. *Physica A: Statistical Mechanics and its Applications, 391,* 5557–5562. http://dx.doi.org/10.1016/j.physa.2012.06.017

Spiliopoulos, L. (2012b). Pattern recognition and subjective belief learning in a repeated constant-sum game. *Games and Economic Behavior, 75,* 921–935. http://dx.doi.org/10.1016/j.geb.2012.01.005

Spiliopoulos, L. (2013a). Beyond fictitious play beliefs: Incorporating pattern recognition and similarity matching. *Games and Economic Behavior, 81,* 69–85. http://dx.doi.org/10.1016/j.geb.2013.04.005

Spiliopoulos, L. (2013b). Strategic adaptation of humans playing computer algorithms in a repeated constant-sum game. *Autonomous Agents and Multi-Agent Systems, 27,* 131–160.

Spiliopoulos, L. (2015). Transfer of conflict and cooperation from experienced games to new games: A connectionist model of learning. *Frontiers in Neuroscience, 9,* 1–18. http://dx.doi.org/10.3389/fnins.2015.00102

Spiliopoulos, L., & Hertwig, R. (2019). Strategic uncertainty and incomplete information: The homo heuristicus does not fold. In R. Hertwig, T. J. Pleskac, T. Pachur, & The Center for Adaptive Rationality (Eds.), *Taming uncertainty* (pp. 89–110). Cambridge, MA: MIT Press.

Spiliopoulos, L., Ortmann, A., & Zhang, L. (2018). Complexity, attention, and choice in games under time constraints: A process analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44,* 1609–1640. http://dx.doi.org/10.1037/xlm0000535

Stahl, D. O. (1993). Evolution of smartn players. *Games and Economic Behavior, 5,* 604–617. http://dx.doi.org/10.1006/game.1993.1033

Stahl, D. O. (1999). Evidence based rules and learning in symmetric normal-form games. *International Journal of Game Theory, 28,* 111–130. http://dx.doi.org/10.1007/s001820050101

Stahl, D. O. (2000). Rule learning in symmetric normal-form games: Theory and evidence. *Games and Economic Behavior, 32,* 105–138. http://dx.doi.org/10.1006/game.1999.0754

Stahl, D. O., & Wilson, P. W. (1994). Experimental evidence on players' models of other players. *Journal of Economic Behavior & Organization, 25,* 309–327. http://dx.doi.org/10.1016/0167-2681(94)90103-1

Stahl, D. O., & Wilson, P. W. (1995). On players' models of other players: Theory and experimental evidence. *Games and Economic Behavior, 10,* 218–254. http://dx.doi.org/10.1006/game.1995.1031

Sterelny, K. (2003). *Thought in a hostile world*. The Evolution of Human Cognition. Oxford: Wiley.

Stigler, S. M. (1986). *The history of statistics*. The Measurement of Uncertainty Before 1900. Cambridge, MA: Harvard University Press.

Sutter, M., Czermak, S., & Feri, F. (2013). Strategic sophistication of individuals and teams. Experimental evidence. *European Economic Review, 64,* 395–410. http://dx.doi.org/10.1016/j.euroecorev.2013.06.003

Thaler, R. H. (1991). *Quasi rational economics*. New York, NY: Sage.

Todd, P. M., Gigerenzer, G., & ABC Research Group. (2012). *Ecological rationality: Intelligence in the world*. New York, NY: Oxford University Press.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185,* 1124–1131. http://dx.doi.org/10.1126/science.185.4157.1124

von Winterfeldt, D., & Edwards, W. (1973). *Flat maxima in linear optimization models* (Technical report No. AD-770 569). Engineenrg Psychology Laboratory Institute of Science & Technology University of Michigan. http://dx.doi.org/10.1037/e406952004-001

Wald, A. (1945). Statistical decision functions which minimize the maximum risk. *The Annals of Mathematics, 46,* 265–280. http://dx.doi.org/10.2307/1969022

Wang, J. T.-y., Spezio, M., & Camerer, C. F. (2010). Pinocchio's pupil: Using eyetracking and pupil dilation to understand truth telling and deception in sender-receiver games. *American Economic Review, 100,* 984–1007. http://dx.doi.org/10.1257/aer.100.3.984

Whiten, A., & Byrne, R. W. (1988). Tactical deception in primates. *Behavioral and Brain Sciences, 11,* 233–244. http://dx.doi.org/10.1017/S0140525X00049682

Whiten, A., & Byrne, R. W. (Eds.). (1997). *Machiavelliani II*. Extensions and evaluations. Cambridge: Cambridge University Press.

Zizzo, D. J., & Tan, J. H. W. (2007). Perceived harmony, similarity and cooperation in 2 × 2 games: An experimental study. *Journal of Economic Psychology, 28,* 365–386. http://dx.doi.org/10.1016/j.joep.2006.06.008

Zizzo, D. J., & Tan, J. H. W. (2011). Gameh: A behavioral approach to predicting cooperation in games. *American Behavioral Scientist, 55,* 987–1013. http://dx.doi.org/10.1177/0002764211407905

(*Appendices follow*)

Appendix A

Definitions

Strategic Dominance

If a (pure) strategy a_i is better than another strategy a_j regardless of the strategy chosen by an opponent, then a_i is said to dominate a_j and the latter is referred to as a dominated strategy. If a strategy a_i dominated all other strategies in a game, then it is referred to as a dominant strategy.

Incomplete Information

In a game of incomplete information, some information about the players' payoffs, preferences, or available actions is not commonly known to all players.

Nash Equilibrium

A set of strategies constitute a Nash equilibrium if no player can gain (increase their payoffs) by a unilateral change of strategy.

Appendix B

Policy Calculations

The following calculations are from the perspective of the row player in the game presented in Table 1.

MaxMin

Find the lowest payoff in each row—these are 32, 23, and 23 in the first, second, and third rows, respectively. Find the maximum of these (32). Therefore, choose action Middle, thereby guaranteeing a payoff of at least 32.

MaxMax

Find the highest own payoff in the whole game matrix (94). Therefore, choose action Up.

SocMax

For each cell, find the collective sum of payoffs to both players (e.g., 58 + 57 = 115 for the upper–left outcome). Find the maximum of these total payoffs (lower–left outcome, 144) and choose action Down.

L1

Sum the own payoffs in each row—these are 184, 94, and 134 for the first, second, and third rows, respectively. Therefore, choose the row with the largest sum, action Up.

D1

Examine whether the opponent has any dominated strategies. Player 2's Left action dominates the Right action, since regardless of Player 1's actions, the former always leads to a higher payoff than the latter. Eliminate the dominated strategy Right from consideration and perform an L1 calculation over the remaining strategies, Left and Center. The payoffs from playing Up are 58 + 32 = 90, for Middle are 34 + 23 = 57, and for Down are 70 + 41 = 111. Consequently, the row player chooses Down, which yields the highest mean payoff over the opponent's nondominated strategies.

L2

Sum the opponent's payoffs for each (column) action—these are 177, 132, and 110 for the first, second, and third columns, respectively. Assume that Player 2 uses L1 and will choose the action with the highest sum, Left. Find the highest own payoff in the column corresponding to Left (70). Play the corresponding action Down.

L3

Assume the opponent is an L2 player. By definition, an L2 player assumes that the opponent is L1. Find the solution to L1 for your own payoffs—from above this is Up. As an L2 player, the opponent (Player 2) will best-respond to Up by searching for their highest own payoff in the row corresponding to Up, and will play Center. Finally, choose the best response to this action by searching for the maximum own payoff within the column associated with Center. The corresponding action is Down with a payoff of 41.

(*Appendices continue*)

Nash Equilibrium

Find the maximum payoffs in each row (i.e., a best response to the assumption that the opponent has played each column action). These are 70, 41, and 94 for actions Left, Center, and Right, respectively. The corresponding best-response actions are Down, Down, and Up, respectively. Perform the same operations for the opponent (Player 2) payoffs for each row action of Player 1. The corresponding best responses are Center, Left, and Left for Player 2's actions Up, Middle, and Down, respectively. Determine for which actions these two best responses coincide. In this case, this occurs for the combination Down, Left; therefore, Player 1 would choose Down and Player 2 Left.

Appendix C

Definitions of the Performance Criteria

Let the action space $A(A')$ of a player (opponent) denote the set of actions $a_n(a_n')$ available. The number of actions for each player (or size of the action space) is denoted by N and N'. We assume that this is the same for both players and, therefore, refer to the common value N as the size of a game. A normal-form game g is defined by a mapping from the action spaces of both players to payoff functions $\pi_g(a_n, a_n')$ and $\pi_g'(a_n, a_n')$. That is, the combination of actions (a_n, a_n') determines the payoffs of both players. Let the decision policy space of a player (opponent) D (D') denote the set of policies d (d') available.

The first payoff in each cell of a normal form game, as presented in Table 1, corresponds to the Row player and the second to the Column player. For example, if the Row player chooses a_1 and the Column player chooses a_2' then the former receives a payoff of 32 and the latter 89.

The Indifference Criterion $\overline{\pi}$

According to the Indifference criterion, the performance of a decision rule d in an environment e consisting of a set of games G_e is given by:

$$\overline{\pi}(d \mid e) = \frac{1}{|D'|} \sum_{d \in D'} E_{G_e}[\pi_g(a(d), a(d'))].$$

As we argued earlier, the L1 and D1 policies converge in their recommendations as the size of the games increases. Because L2 is by definition a best response to these policies, including them separately with equal weighting in the population would bias the criterion in favor of L2. We decided to include L1 and D1 as a single "policy," by attributing half the weight to each and combining them, that is, taking the mean of the individual performance of the two and then weighting it as if it were a single policy (L1/D1). Consequently, $|D'| = 9$ since $D' = \{$Random, MaxMax, MaxMin, Eq, SocMax, L1/D1, L2, L3, NE$\}$. However, we still calculate the performance $\overline{\pi}$ ($d \mid e$) of L1 and D1 separately, that is, $|D| = 10$.

The Wald Criterion π_w

According to the Wald criterion, the performance of a decision rule d in an environment e consisting of a set of games G_e is given by:

$$\pi_w(d \mid e) = \min_{d' \in D'} E_{G_e}[\pi_g(a(d), a'(d'))].$$

The Composition-Robustness Criterion $\pi_\%$

Denote the proportion of a decision rule d' in the population by $p_{d'}$ and the set of proportions by δ. Assume that the set of possible proportions of decision rules is distributed uniformly over the unit simplex $\Delta^{10} = \{\sum_{d' \in D'} p_{d'} = 1$ and $p_{d'} \geq 0$ for all $p_{d'}\}$. Let π ($d \mid \delta, e$) be the expected payoffs for a decision rule d playing the set of games G_e against a population of opponent decision rules comprised of δ proportions. This is essentially a weighted average of the performance of d against every other decision rule d'. Let $I(d \mid \delta, e)$ be the percentile ranking (normalized between 0 and 1) for a decision rule d's expected payoffs $\pi(d \mid \delta, e)$ compared with the expected payoffs of all other decision rules. According to the Composition-Robustness criterion, the performance of a decision rule d in an environment e consisting of a set of games G_e is given by the integration of the percentile ranking over all possible values of δ, that is, proportions of the decision rules in the population:

$$\pi_\%(d \mid e) = \int_{\delta \in \Delta^{10}} I(d \mid \delta, e) d\delta.$$

(Appendices continue)

Appendix D

Results: Composition-Robustness Criterion, $\pi_\%$

In this section, we investigate the decision policies' performance with respect to the Composition-Robustness criterion, $\pi_\%$, which is the percentile ranking of a decision policy's performance averaged across all possible population mixtures. The performance of the decision policies are presented in Figures D1, D2, and D3 below.

Neutral Environment ($\rho = 0$)

The most interesting cases are L1, D1, and L2. D1 significantly outperforms L1 in small games, because the likelihood of a dominated action existing in these games is higher. L1 and D1 are quite robust across the whole map. The Nash equilibrium, while performing better than random across the map, is only competitive with L1 and D1 for a game size of 2. Very simple heuristics such as MaxMin, MaxMax, and Eq generally perform poorly. Finally, SocMax is a very poor choice for small games but performs adequately in larger games, especially for high payoff uncertainty.

Harmonious Environment ($\rho = 0.5$)

The Nash equilibrium is robust across the whole map, but is surpassed by SocMax for high payoff uncertainty. However, Soc-

Max performs poorly in very small games with little payoff uncertainty. L1 and D1 are moderately robust, with D1 performing better than L1 in small games where satisfying opponent dominance is advantageous (more so because of the positive degree of harmony).

Discordant Environment ($\rho = -0.5$)

Environments with conflict significantly alter the performance landscape in terms of the Composition-Robustness criterion. SocMax, which derives its prowess from cooperative behavior, fails to be among the top-performing policies in such an environment. The Nash equilibrium also performs poorly and is not robust to game size or payoff uncertainty beyond games of size 2. Predictably, MaxMin is a better choice in discordant environments, but is dominated by L1 and D1. D1 is the best performer on average throughout the whole map, outperforming L1 primarily in small games where the likelihood of an opponent having a dominated strategy is still relatively high. L2 and L3 perform moderately well across the whole map.
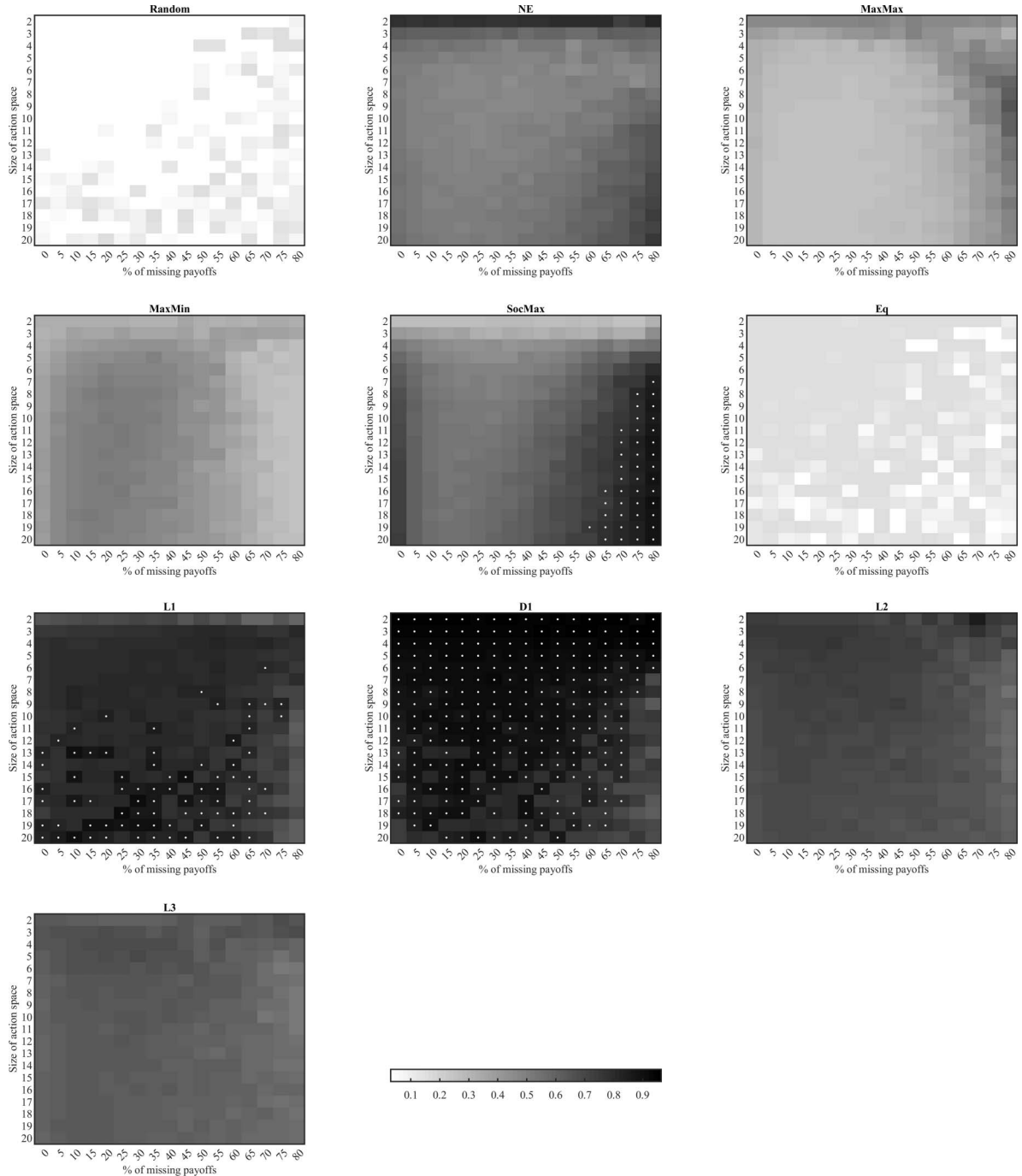
(*Appendices continue*)

Figure D1. Decision policy performance according to the $\pi_\%$ (Composition-Robustness) criterion for neutral environments ($\rho = 0$). The set of top-performing policies for each environment is marked by an overlaid white dot. NE = Nash equilibrium; L1 = Level-1; L2 = Level-2; L3 = Level-3; SocMax = social maximum; D1 = dominance-1; Eq = equality.
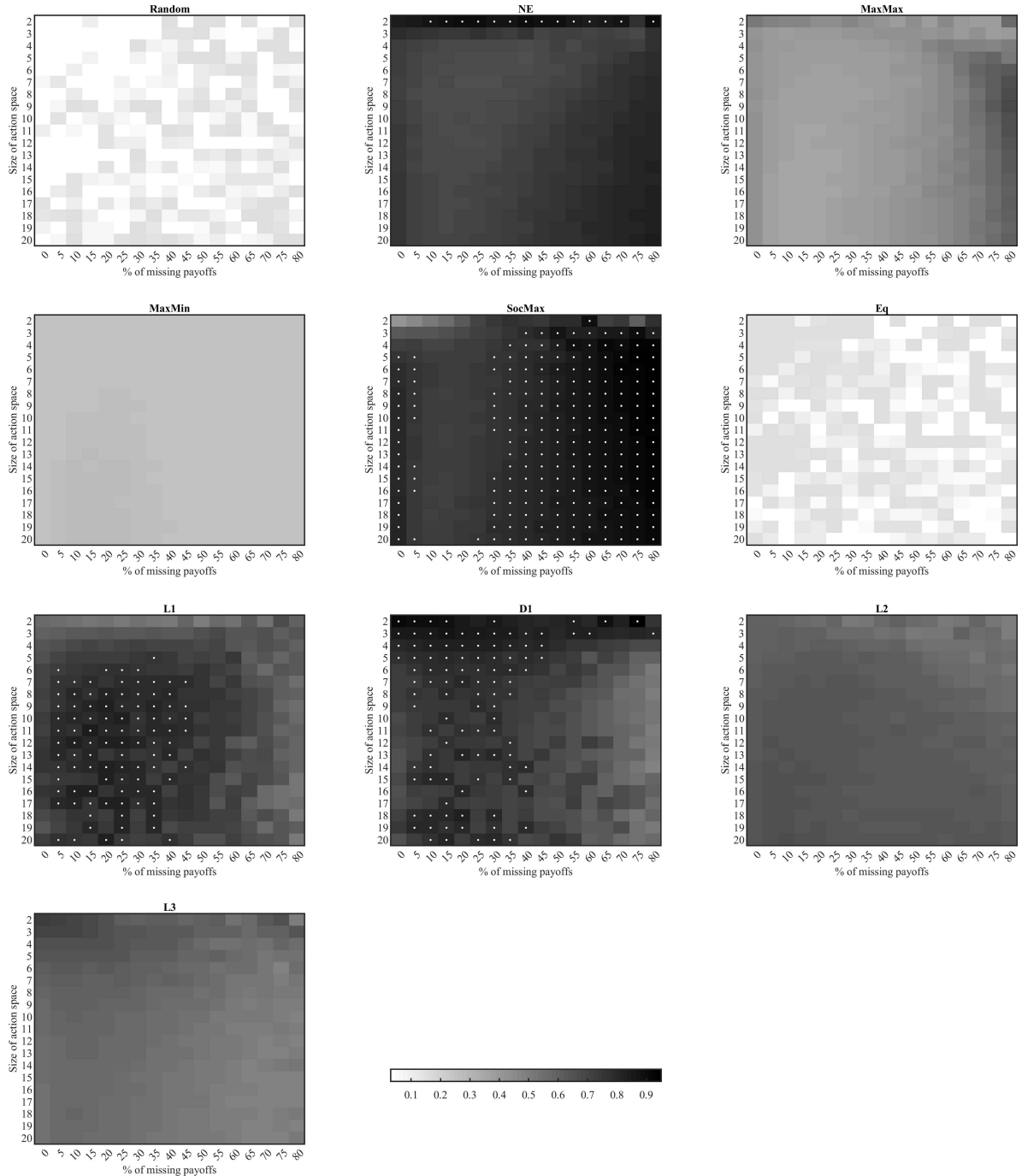
Figure D2. Decision policy performance according to the $\pi_\%$ (Composition-Robustness) criterion for harmonious environments ($\rho = 0.5$). The set of top-performing policies for each environment is marked by an overlaid white dot. NE = Nash equilibrium; L1 = Level-1; L2 = Level-2; L3 = Level-3; SocMax = social maximum; D1 = dominance-1; Eq = equality.

(*Appendices continue*)

Figure D3. Decision policy performance according to the $\pi_\%$ (Composition-Robustness) criterion for discordant environments ($\rho = -0.5$). The set of top-performing policies for each environment is marked by an overlaid white dot. NE = Nash equilibrium; L1 = Level-1; L2 = Level-2; L3 = Level-3; SocMax = social maximum; D1 = dominance-1; Eq = equality.
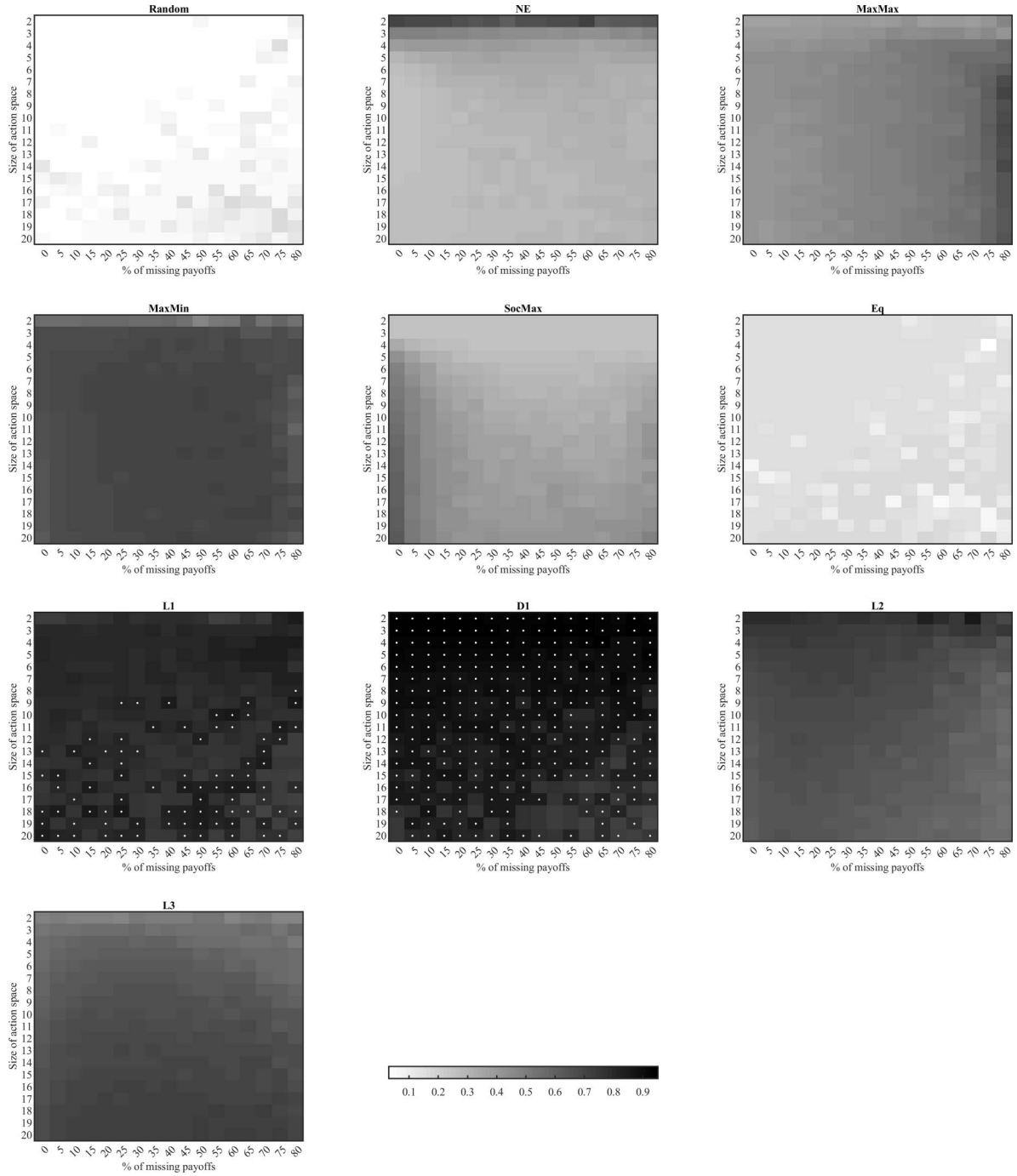
(*Appendices continue*)

Appendix E

Pairwise Comparisons of Decision Policies

In this section, we present pairwise comparisons of how decision policies perform when playing with each other. Because an exhaustive comparison of all possible pairs of decision policies for each environment is prohibitive, we present the performance averaged over all m and n for environments where games are generated with $\rho = 0$, 0.5, and -0.5. Tables E1–E3 present the average payoffs to a player using a decision policy (by rows) against an opponent using the policy designated in the columns. If there were no strategic uncertainty about the opponent, then a player could determine the expected best response to the opponent's decision policy by looking up this table. Payoffs in bold highlight these best-response policies for each of the opponent policies, including policies whose payoffs are within 5% of the best-response; we refer to these policies as the set of best-responses. For example, for $\rho = 0$ and an opponent using MaxMax, the best response would be to use the D1 policy with average payoffs of 25.3 (L1 is also in the set of best-response policies, with payoffs of 24.5).

The Nash equilibrium policy is only ever a best-responding policy to itself for $\rho = 0$ and 0.5, while it is not a best-response to itself for $\rho = -0.5$. However, the latter finding is because of the high incidence of games with no pure-strategy Nash equilibrium where the implementation of the Nash equilibrium policy chooses randomly—if one were to look only at games with pure-strategy Nash equilibrium, then the Nash equilibrium policy is a best-response to itself even in discordant environments. Note, that while L1 may receive a boost from Nash equilibrium opponents for games without pure-strategy Nash equilibrium—because L1 is a best response to L0 or random choice–it is clear that L1 still exhibits high performance even against the remaining policies, completely excluding Nash equilibrium opponents.

The documented strength of the L1 policy (and D1) according to the Indifference criterion can be broken down into its source components, that is, which opponent policies it performs best against. L1 belongs to the set of best-response policies against four opponent policies in $\rho = 0$, two policies for $\rho = 0.5$ and five policies for $\rho = -0.5$. Furthermore, when L1 does not belong to the set of best-response policies, it is very often the next best policy. This is the case, for example, for $\rho = 0.5$, against MaxMin, L1/D1, L2, and L3.

SocMax is always a best-response to itself and against MaxMax for $\rho = 0.5$. This highlights the fragility of the SocMax policy to strategic uncertainty, particularly in terms of the Wald criterion, which we discussed in the manuscript. Equality, MaxMax and MaxMin are never best-responding policies for any degree of ρ. As expected, sophisticated Level-k policies, L2 and L3, are typically best-responses to other Level-k policies, where k is one level lower.

Table E1

Pairwise Comparisons of Average Decision Policy Performance for Neutral Environments ($\rho = 0$), Averaged Across n and m

Player	Opponent policies									
	Random	MaxMax	MaxMin	SocMax	Eq	L1	D1	L2	L3	NE
Random	.0	.0	.0	7.3	.0	.0	.0	−.1	.3	1.1
NE	6.2	14.1	8.1	37.1	6.3	12.7	12.9	14.8	18.4	**50.7**
MaxMax	10.4	11.6	9.9	38.3	9.5	10.6	10.6	10.5	11.4	16.0
MaxMin	15.7	15.9	15.5	22.7	14.7	15.7	15.7	15.5	15.7	16.8
SocMax	7.3	16.1	8.9	**73.9**	7.3	13.4	13.5	11.8	13.7	30.8
Eq	.0	.3	.9	7.3	.0	1.2	1.2	.5	.8	1.8
L1	**24.1**	**24.5**	23.8	37.6	**22.9**	24.0	24.0	23.4	**24.1**	26.7
D1	23.5	**25.3**	24.5	37.8	22.4	24.7	24.7	24.2	**24.9**	27.2
L2	8.2	21.7	**27.0**	24.7	7.7	**74.5**	**72.6**	15.5	15.9	15.8
L3	8.5	15.0	11.0	23.0	8.2	16.0	16.2	**74.3**	20.4	17.8

Note. NE = Nash equilibrium; L1 = Level-1; L2 = Level-2; L3 = Level-3; SocMax = social maximum; D1 = dominance-1; Eq = equality. Each cell presents the expected payoffs to the policy (in a specific row) against an opponent playing the policy in a specific column. Numbers in bold indicate the best-response policies for each possible opponent policy.

(*Appendices continue*)

Table E2
Pairwise Comparisons of Average Decision Policy Performance for Harmonious Environments (ρ = .5), Averaged Across n and m

Player	Opponent policies									
	Random	MaxMax	MaxMin	SocMax	Eq	L1	D1	L2	L3	NE
Random	.0	7.5	9.9	11.0	.0	15.2	14.9	5.2	8.8	7.9
MaxMax	11.6	32.2	22.4	66.7	10.9	33.7	33.9	29.4	30.5	52.0
MaxMin	16.7	23.8	21.5	28.0	16.3	27.4	27.3	31.3	23.6	24.4
SocMax	11.0	**55.7**	22.1	**114.3**	10.9	38.9	39.0	30.2	35.2	84.9
Eq	.0	7.3	10.2	10.9	−.1	15.7	15.4	5.3	9.1	8.2
L1	**25.7**	37.8	32.8	49.1	**25.1**	41.0	40.9	74.1	37.9	41.9
D1	**25.1**	38.5	33.3	49.4	**24.5**	41.5	41.4	73.6	38.5	42.7
L2	8.7	32.2	**37.4**	35.3	8.4	**94.8**	**93.0**	24.5	**75.4**	30.6
L3	14.3	33.6	24.4	41.1	14.0	36.6	36.7	**93.8**	38.5	39.0

Note. NE = Nash equilibrium; L1 = Level-1; L2 = Level-2; L3 = Level-3; SocMax = social maximum; D1 = dominance-1; Eq = equality. Each cell presents the expected payoffs to the policy (in a specific row) against an opponent playing the policy in a specific column. Numbers in bold indicate the best-response policies for each possible opponent policy.

Table E3
Pairwise Comparisons of Average Decision Policy Performance for Discordant Environments (ρ = −.5), Averaged Across n and m

Player	Opponent policies									
	Random	MaxMax	MaxMin	SocMax	Eq	L1	D1	L2	L3	NE
Random	.0	−7.5	−9.9	4.0	−.1	−15.3	−15.0	−5.3	−3.1	−2.0
NE	3.5	−1.4	−4.8	11.3	3.4	−8.7	−8.5	1.0	4.2	8.9
MaxMax	11.6	3.8	−4.5	25.7	10.2	−8.4	−8.5	−4.4	6.2	8.2
MaxMin	16.7	14.4	8.0	19.9	14.9	6.6	6.7	3.8	13.2	15.0
SocMax	4.1	−6.6	−4.2	39.0	4.0	−9.7	−9.5	−2.9	1.3	5.1
Eq	.0	−6.5	−8.3	4.1	.0	−13.1	−12.8	−4.3	−2.7	−1.4
L1	**25.7**	**21.5**	12.7	32.4	**23.6**	10.4	10.6	−23.5	**18.2**	**23.3**
D1	**25.2**	**22.1**	13.1	32.3	**23.1**	10.9	11.0	−21.6	**19.0**	**23.5**
L2	8.9	17.1	**17.7**	18.8	7.9	**66.1**	**63.9**	7.2	−39.1	8.7
L3	5.4	1.8	−1.1	12.4	5.2	−1.7	−1.9	**73.8**	7.7	5.4

Note. NE = Nash equilibrium; L1 = Level-1; L2 = Level-2; L3 = Level-3; SocMax = social maximum; D1 = dominance-1; Eq = equality. Each cell presents the expected payoffs to the policy (in a specific row) against an opponent playing the policy in a specific column. Numbers in bold indicate the best-response policies for each possible opponent policy.

(Appendices continue)

Appendix F

Other Results

Table F1

The Probability of a Randomly Drawn Game With No Pure-Strategy Nash Equilibria

	ρ		
n	0	.5	−.5
2	.13	.05	.21
3	.22	.07	.41
4	.26	.07	.55
5	.28	.06	.64
6	.30	.05	.70
7	.31	.05	.75
8	.32	.04	.78
9	.32	.03	.81
10	.33	.03	.84
11	.33	.03	.86
12	.34	.02	.87
13	.33	.02	.88
14	.34	.02	.89
15	.34	.02	.90
16	.34	.01	.91
17	.35	.01	.92
18	.35	.01	.93
19	.35	.01	.93
20	.34	.01	.93

Table F2

Policies' Consistency Averaged Across (n, m) for Environments With ρ = {0, 0.5, −0.5} and m > 0 (%)

	ρ		
Policies	0	.5	−.5
Random	14	14	14
NE	28	22	17
MaxMax	31	30	31
MaxMin	33	32	33
SocMax	31	27	26
Eq	39	39	39
L1	37	35	37
D1	36	34	36
L2	27	26	27
L3	27	24	23

Note. NE = Nash equilibrium; L1 = Level-1; L2 = Level-2; L3 = Level-3; SocMax = social maximum; D1 = dominance-1; Eq = equality.