Available online at www.sciencedirect.com
SCIENCE DIRECT®
Games and Economic Behavior 46 (2004) 365–382
GAMES and Economic Behavior
www.elsevier.com/locate/geb

A model of noisy introspection

Jacob K. Goeree [a,*] and Charles A. Holt [b]

[a] *University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands*
[b] *Department of Economics, University of Virginia, 114 Rouss Hall, Charlottesville, VA 22901, USA*

Received 10 October 2000

Abstract

We present a theoretical model of noisy introspection designed to explain behavior in games played only once. The model determines layers of beliefs about others' beliefs about ..., etc., but allows for surprises by relaxing the equilibrium requirement that belief distributions coincide with decision distributions. Noise is injected into iterated conjectures about others' decisions and beliefs, which causes the predictions to differ from those of deterministic models of iterated thinking, e.g., rationalizability. The paper contains a convergence proof that implies existence and uniqueness of the outcome of the iterated thought process. In addition, estimated introspection and noise parameters for data from 37 one-shot matrix games are reported. The accuracy of the model is compared with that of several alternatives.
© 2003 Elsevier Inc. All rights reserved.

JEL classification: C72; C92

Keywords: Game theory; Introspection; Rationalizability; Nash equilibrium; Experiments; Noisy behavior; Bounded rationality

1. Introduction

Game theory is a collection of mathematical models that are used to predict and explain behavior in strategic situations where players' optimal decisions depend on what other players are expected to do. A Nash equilibrium in a game is a state of rest in the sense that no player would want to change their own strategy unilaterally, knowing what strategies others are using. This notion of equilibrium is appropriate when decision makers have learned from repeated interactions what others can be expected to do. Many recent

* Corresponding author.
E-mail address: jkg@fee.uva.nl (J.K. Goeree).

0899-8256/$ – see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/S0899-8256(03)00145-3

applications of game theory outside of economics, however, are more accurately modeled as games played only once, e.g., legal disputes, election campaigns, and international conflicts. If game theory is to become a unifying theory of social science, we must develop models that predict behavior in one-shot interactions where it is not possible for players to learn what to expect. In such cases, it may or may not be appropriate to assume that some process of introspection leads players to a Nash equilibrium. Humans are capable of many layers of speculation about possible actions and reactions, like the inspector in Edgar Allen Poe's *The Purloined Letter* who tries to think about where the thief thinks the inspector will look, and where the thief thinks the inspector thinks the thief will hide the letter, etc. Iterated reasoning of this type corresponds to considering a sequence of best responses to best responses. It is well known that this sequence converges to a Nash equilibrium, if it converges at all. More analysis is needed, however, if there are multiple Nash equilibria, or if people are more and more uncertain about what others think, about what others think they think, etc. This paper presents a new model of noisy introspection designed to explain behavior in single-shot games.

2. Some simple games

Since game theory makes precise predictions on the basis of the payoffs and game structure, it is natural to test this theory in controlled experiments where the players' money payoffs depend on their own and others' decisions. Many of the initial tests of the Nash's theory involved repeated plays of a game for which the outcome with the highest total payoffs for both players is not an equilibrium. In Table 1, for example, each player chooses between decisions labeled R or S, with Row's payoffs listed on the left. The lower-right box shows three variations of Row's payoffs that do not alter the fact that this outcome maximizes the sum of the players' payoffs. This (R, R) outcome is risky in the sense that a unilateral deviation by Row would increase Row's payoffs to 25. Anticipating that, Column may also consider switching to S, and (S, S) is the only equilibrium with no unilateral incentive to deviate. Notice that iterated thinking in this manner focuses attention on the Nash equilibrium. In experiments with repeated plays of games like this, it is common to observe a significant amount of cooperative (R, R) outcomes. Nash criticized initial tests of his theory using repeated plays with the same individuals, arguing that the theory should be applied to the whole multiple-round interaction (Nasar, 1998).

In this paper, we will calibrate our analysis using data from 37 one-shot games reported by the psychologists Guyer and Rapoport (1972), including seven asymmetric prisoner's

Table 1
A one-sided prisoner's dilemma game (Row payoff, Column payoff): three variations affecting Row's "incentive to defect"

Row player's decision	Column player's decision: S	Column player's decision: R
S	6, 12	25, −4
R	−4, 18	9/18/24, 25

Table 2
A coordination game with multiple equilibria: three variations that reduce the risk for Row to "force" Column

Row player's decision	Column player's decision	
	S	*R*
S	10, 19	4, −7
R	−14/−7/−1, 4	19, 10

dilemma games, three of which are shown in Table 1. The changes in Row's payoffs from 9 to 18 to 24 successively increase the attractiveness of the joint-maximizing outcome for Row, without changing the location of the unique Nash equilibrium at (S, S). These changes caused a sharp reduction in the incidence of S choices by Row players, from 90% to 84% to 71%. While the best responses that determine a Nash equilibrium only depend on the signs, not the magnitudes, of payoff differences, the data seem to be affected by magnitudes in an intuitive manner.

Much recent work in game theory pertains to games with multiple equilibria, as in each of the three variations of the "coordination game" in Table 2, also taken from Guyer and Rapoport (1972). The S decision is the "maximin" strategy; it maximizes the minimum payoff, so we will sometimes refer to S as "safe" and R as "risky." The risky decision R is best for each if the other will also choose R, but S is best if the probability of the other playing S is sufficiently high. These best response functions for the first variation of this game are shown in Fig. 1 as dark lines, with Row's probability of S on the vertical axis and Column's probability of S on the horizontal axis (please ignore the dotted lines for now). Row's best response stays on the bottom of the figure on the left side, but jumps to 1 as soon as Column's probability of S exceeds 0.39. The crossings of the best response

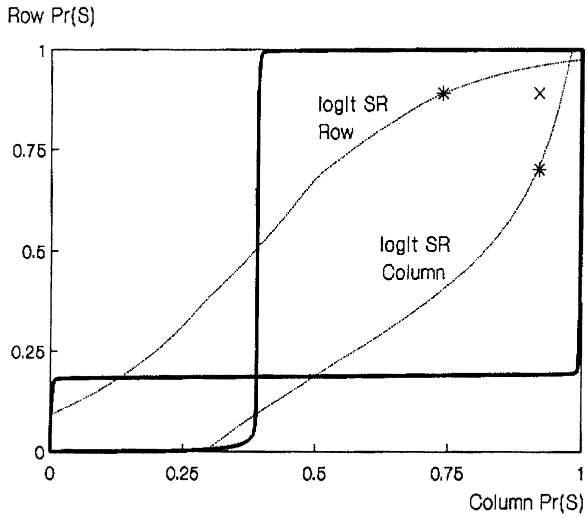

Fig. 1. A coordination game: best responses (dark lines) and logit stochastic responses (dotted lines).

functions at $(0,0)$ and $(1,1)$ represent Nash equilibria in pure strategies, and the crossing at $(0.39, 0.19)$ is an equilibrium in randomized strategies.

In the first variation of this game, the percentages of S choices were 89% and 92% for Row and Column players respectively in the one-shot games reported by Guyer and Rapoport (1972). These proportions are graphed as an "$\times$" in the upper-right part of the figure. This outcome is near the (S, S) Nash outcome that is "risk-dominant" in the sense of Harsanyi and Selten (1988). This risk dominance criterion is based on the intuitive notion that it is more risky to deviate from the (S, S) outcome: in the first variation a unilateral deviation costs 24 for Row and 26 for Column, whereas unilateral deviations from the (R, R) outcome only cost 15 for Row and 6 for Column. The three variations of this game, going from top to bottom, reduce Row's "deviation loss" at the (S, S) equilibrium, but it is still the risk dominant outcome in all cases.[1] These payoff changes do reduce the riskiness of Row's R decision, and not surprisingly, the incidence of S play falls from 89% to 88% to 75%. Again we see that magnitudes of payoff differences seem to matter, even when the "preferred" Nash equilibrium is unaffected.

Theoretical justifications for payoff-magnitude effects can be devised by introducing some random noise into the decision-making process. Without such noise, the probability of choosing a decision jumps "sharply" from 0 to 1 as soon as its expected payoff is the highest available. Following Luce (1959), suppose instead that the probability of choosing each decision is a smoothly increasing function of the expected payoff for that decision. Luce provided an axiomatic derivation of the popular "logit" rule, which is based on exponential functions.[2] Suppose there are only two possible options, S and R, which yield expected payoffs $\pi^{\mathrm{e}}_{i,\mathrm{S}}$ and $\pi^{\mathrm{e}}_{i,\mathrm{R}}$ to player i. Then the logit probability of choosing strategy S is given by

$$p_i(S) = \frac{\exp(\pi^{\mathrm{e}}_{i,\mathrm{S}}/\mu)}{\exp(\pi^{\mathrm{e}}_{i,\mathrm{S}}/\mu) + \exp(\pi^{\mathrm{e}}_{i,\mathrm{R}}/\mu)}, \tag{1}$$

and the probability that player i chooses R is simply 1 minus the probability in (1). The denominator in (1) ensures that probabilities lie between 0 and 1, and μ is a "noise" or "error" parameter. As μ goes to zero, the decision with the highest expected payoff is selected with probability one. The slight rounding off of the corners of the response functions in Fig. 1 is due to the fact that these were drawn for a low μ value of 0.1 (instead of 0). A further increase in μ softens the corners even more, and the dotted lines are for logit stochastic responses with $\mu = 6.6$, which we estimated from the Guyer and Rapoport data as explained below. With this higher amount of noise, the lines only intersect once, close to the risk-dominant Nash equilibrium. This intersection at $(0.97, 0.98)$ is a *quantal*

[1] Risk dominance selects the equilibrium for which the product of Row's and Column's deviation losses is greatest.

[2] The necessary axioms are that choice probabilities be unaffected by adding a constant to all payoffs, and that ratios of probabilities for two decisions be independent of the payoffs associated with any other decision. An alternative derivation of the logit rule is based on an assumption that the payoffs for each decision are augmented by an unobserved preference shock, with a double exponential distribution. This random-preference model was used by Harsanyi (1967–1968) in an equilibrium analysis, which is closely related to the quantal response equilibrium discussed below.

response equilibrium proposed by McKelvey and Palfrey (1995); it is an equilibrium in the sense that each of these probabilities is a logit stochastic response to the other one. This equilibrium is a generalization of the Nash equilibrium, and it converges to a Nash outcome as μ goes to zero (perfect rationality). In the other extreme, as the noise parameter μ goes to infinity (perfect randomness), the intersection of the logit stochastic response lines converges to the centroid, with probabilities of 0.5 for each decision.[3]

In some cases, laboratory experiments produce outcomes that are reasonably close to Nash predictions, both in one-shot play as for the coordination game discussed above, and in settings where players are randomly rematched with each other for a series of (approximately) one-shot interactions. There are many situations, however, in which observed play shows systematic deviations from Nash, and these are often tracked nicely by the quantal response equilibrium (see McKelvey and Palfrey, 1995; and for continuous games, Goeree and Holt, 1999).[4] For the middle variation of the asymmetric prisoner's dilemma in Table 1, the unique Nash outcome is (S, S), but the percentage of S choices (84% for Row and 82% for Column) are quite close to the logit quantal response predictions (81% for Row and 85% for Column).

Both the Nash equilibrium and its quantal response generalization are *equilibrium* concepts, e.g., fixed point intersections in Fig. 3, that map belief probabilities into actions that occur with the same probability. In all but the simplest games, equilibrium concepts will have the most explanatory power when people have the opportunity to learn about others' decision probabilities through experience. Such "rational expectations" assumptions may not be appropriate in one-shot interactions with no chance for learning and adaptation.

To see why surprises may occur in disequilibrium situations, consider again the dotted stochastic response lines for the coordination game in Fig. 1. The logit quantal response equilibrium (for the pooled estimate, $\mu = 6.6$) is almost at an extreme corner where the probabilities of S are essentially 1. Prior to the first and only play of a game like this, it may be the case that players are not so sure about others' decisions. If the Row player thinks that Column will only play S with a probability of about 0.7, for example, then the logit response is represented by the asterisk on the dotted line for Row. A similar asterisk is shown on Column's stochastic response line, and together these beliefs produce choice probabilities that are somewhat smaller than the logit and Nash predictions. In fact, beliefs of about 0.7 produce stochastic responses that are close to the actual choice percentages marked with the "$\times$." Since these two asterisk points do not coincide, the expectations are not in equilibrium, e.g., Row expects a 0.7 chance of S, whereas Column plays S with probability 0.92. Notice that the asterisk points pull decisions away from the logit

[3] In fact, there is a locus of quantal response equilibria, connecting the center of Fig. 1 with the upper-right corner, where each point on the locus corresponds to a quantal response equilibrium for a particular value of μ. This locus will pass very close to the "$\times$" that represents the data average, and in this sense, a quantal response model can explain the data in this game. The approach taken in Section 5, however, is to estimate a single value of μ using data from 37 different games.

[4] These papers report situations in which both quantal response equilibria and the observed choice data may be located far from Nash outcomes, in some cases on the opposite side of the set of feasible decisions, see also Goeree and Holt (2001).

intersection toward lower probabilities of S. This "pull to the center" is caused by (1) the tendency for uncertainty about other's actions to push beliefs towards 0.5, and (2) the fact that the dotted line logit response functions in Fig. 1 have positive slopes. In games where the logit response functions have negative slopes, however, the effect is reversed: greater uncertainty about other's decisions will pull decisions toward higher probabilities of S than are implied by the logit equilibrium. This "push to the edge" effect is revisited below in the context of "chicken" games where it is best to play S (safe) when the other player is playing R (risky) and vice versa.[5]

The next section presents a model of noisy introspection that formalizes the intuition from these examples. This model is essentially a noisy version of the Bernheim (1984) and Pearce (1984) notion of rationalizability, as explained in Section 4. We used data from one-shot games to estimate the model parameters, and Section 5 contains an assessment of how the model compares with the Nash and logit quantal response predictions. The final section concludes.

3. Iterated noisy introspection

Play in many types of one-shot games is likely to contain surprises, no matter how carefully players think about the payoffs before deciding. We therefore relax the equilibrium condition of consistency of actions and beliefs by introducing a process of iterated conjectures. Consider the one-sided prisoner's dilemma game in Table 1, where Row's R decision is never a best response for any beliefs about Column's decision. Assuming that Row is rational, Column anticipates Row choosing S and hence Column also chooses S. Stated differently, the unique "rationalizable" outcome of this game is the Nash outcome (S, S). In more complicated games, the notion of rationalizability corresponds to an iterated process of eliminating strategies that are never best responses for any beliefs (Bernheim, 1984; Pearce, 1984).[6] We will also consider iterated reasoning of this type but the noise observed in laboratory data motivates us to incorporate stochastic

[5] These slope effects may be either negated or reinforced if the relevant value of μ is different for the logit and introspection model, since lower error rates will push the logit intersection closer to a Nash equilibrium. Estimates for μ are reported in Section 5 below.

[6] Another well known model of introspection is Harsanyi and Selten's (1988) "tracing procedure." This procedure involves an axiomatic determination of players' common priors (the "preliminary theory") and the construction of a modified game with payoffs for each decision that are weighted averages of those in the original game and of the expected payoffs determined by the prior distribution. By varying the weight on the original game, a sequence of best responses for the modified game are generated. This process is used to select one of the Nash equilibria of the original game. Olcina and Urbano (1994) also use an axiomatic approach to select a prior distribution, which is then revised by a simulated learning process that is essentially a partial adjustment from current beliefs to best responses to current beliefs. The model has the attractive theoretical property that it selects the risk-dominant Nash equilibrium in 2×2 games. Since the simulated learning process has no noise, it will converge to the Nash equilibrium in games with a unique equilibrium, which is an undesirable feature in light of the one-shot data reported below. For an alternative approach, see Capra (1998) who introduces stochastic elements. In her model, beliefs are represented by degenerate distributions that put all probability mass at a single point, and the introspective process stops when a point is mapped into itself by a linked pair of stochastic best response functions. Our model, first described in Goeree and Holt (1999), is closer in spirit to the one considered

elements into the best responses.[7] This is done by injecting some noise into the system via the logit choice function in (1).

To illustrate our "noisy introspection" model we start by considering symmetric 2×2 games. This has the advantage that a player's choice can be represented by a single probability, e.g., the probability of choosing S. Moreover the symmetry assumption allows us to drop the player-specific subscripts in (1). The expected payoffs in (1) depend on a player's belief about the other's play, i.e., the probability, q, with which the other chooses S. We write $\pi_{\mathrm{S}}^{\mathrm{e}}(q)$ and $\pi_{\mathrm{R}}^{\mathrm{e}}(q)$ to make this dependence more clear. Let $\phi^{\mu_0}(q)$ denote a player's logit best response given the player's belief, q:

$$\phi^{\mu_0}(q) = \frac{\exp(\pi_{\mathrm{S}}^{\mathrm{e}}(q)/\mu_0)}{\exp(\pi_{\mathrm{S}}^{\mathrm{e}}(q)/\mu_0) + \exp(\pi_{\mathrm{R}}^{\mathrm{e}}(q)/\mu_0)}, \tag{2}$$

where μ_0 is the error rate associated with a player's decision. Equation (2) determines a player's choice probability for decision S as a function of the player's "first-order" belief about the other's decision. The other's decision, in turn, depends on the other's belief about the player's own decision, etc. This naturally leads us to consider higher-order beliefs, denoted by $B^0, B^1, B^2, \ldots,$ where:

- B^0 represents a player's choice probability,
- B^1 represents a player's (first-order) belief about the other's choice,
- B^2 represents a player's (second-order) belief about the other's belief about the player's own choice, etc.

(Note that B^0 and B^1 correspond to $\phi^{\mu_0}(q)$ and q in (2).) The logit best response function can be used iteratively to construct a player's higher-order beliefs. In particular, we model a player's first-order belief as the other's logit best response, $\phi^{\mu_1}(B^2)$, given the player's second-order belief, B^2. In other words, the thought process that produces a player's first-order belief about what the other will do is modeled as the other's noisy best response given what the player thinks the other thinks the player will do. This second level of introspection about what the other person is thinking is likely to be somewhat imprecise, so we assume that the error rate μ_1 associated with the transformation $\phi^{\mu_1}(B^2)$ is larger than the error parameter μ_0 for the transformation $\phi^{\mu_0}(B^1)$ that determines a player own choice probabilities.

To define a player's higher-order beliefs it will prove useful to introduce the following composition of logit best responses for $n > k$:

$$\Phi_n^k = \phi^{\mu_k} \circ \phi^{\mu_{k+1}} \circ \cdots \circ \phi^{\mu_n}, \tag{3}$$

by Kübler and Weizsäcker (2000), which has two error parameters, one pertaining to decisions and one pertaining to beliefs.

[7] One important use of introspective theories is to model beliefs in the first period of an experiment. In some papers, we have initialized computer simulations and learning models by assuming that players make stochastic best responses to uniform distributions of others decisions. Alternatively, one could assume that others are making stochastic responses to uniform distributions.

which maps a player's nth order belief into her kth order belief: $B^k = \Phi_n^k(B^n)$. Presumably, noise in successive iterations increases since there is likely to be more error associated with beliefs about others' beliefs about ..., etc. We therefore assume that the error rates associated with higher levels of iterated reasoning form an increasing sequence ($\mu_0 \leqslant \mu_1 \leqslant \mu_2 \leqslant \cdots$), which diverges as the number of iterations increases to infinity: $\mu_\infty = \infty$. A player's kth order belief, B^k, can now be defined as the limit:

$$B^k = \lim_{n\to\infty} \Phi_n^k(p_0) = \lim_{n\to\infty} \phi^{\mu_k} \circ \phi^{\mu_{k+1}} \circ \cdots \circ \phi^{\mu_n}(p_0). \tag{4}$$

The probability, p_0, that appears on the right side of (4) is the starting point of the iterative thought process, which can be chosen arbitrarily. The reason is that logit best response ϕ^μ for $\mu = \infty$ maps *any* initial belief probability to a uniform probability of one-half (perfectly noisy behavior). In other words, the assumption that the sequence of error rates diverges to infinity implies that higher-order beliefs become more and more diffuse, i.e., players effectively "start out" reasoning from a uniform prior.[8,9]

The noisy introspection model defined by (4) is flexible and allows many special cases. For instance, if $\mu_0 = \infty$ (and, hence, $\mu_k = \infty$ for all k), the model produces uniform choice probabilities as is the case for Stahl and Wilson's (1995) "level-0 rationality." This case corresponds to the left-most vertical line in Fig. 2. The second vertical line in the figure represents the case $\mu_0 = 0$ and $\mu_k = \infty$ for $k \geqslant 1$, which is "level-1 rationality,"

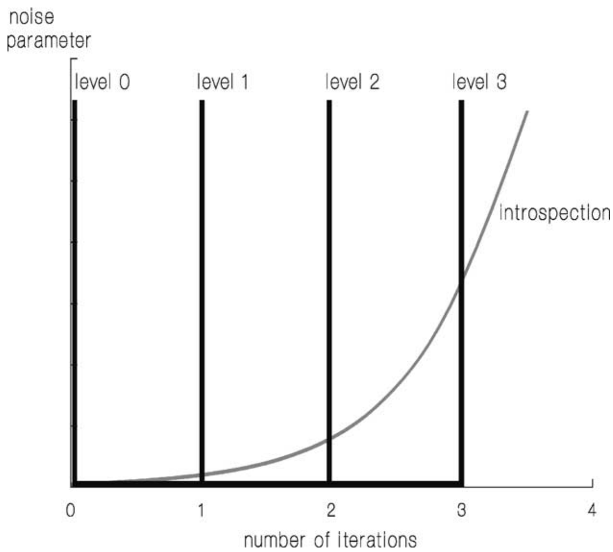

Fig. 2. A representation of models of noisy iteration.

8 The assumption that the error rate diverges to infinity may not be reasonable for all games, e.g., the prisoner's dilemma where arbitrarily high-order beliefs may still put more weight on the "defect" strategy.

9 An alternative approach would be to let individuals have probability distributions over the distributions that characterize others' beliefs, etc. This "distributions over distributions" approach is theoretically appealing but intractable for all but the simplest games.

corresponding to a rational best response to uniform beliefs. Higher levels of rationality can be generated similarly.[10]

Rather than assuming a fixed number of iterations with extreme values of the error parameters, the error rates could be increasing smoothly as shown by the curved line in Fig. 2. In Section 5 below, we report estimates of $\mu_1 = 4.4$, $\mu_2 = 17.6$, and $\mu_3 = 53.8$, which produce a pattern of increases that roughly corresponds to that shown in Fig. 2. To obtain a parsimonious specification, we assume that the error rates grow geometrically with each iteration: $\mu_k = t^k \mu_0$, where the "telescoping" parameter, $t > 1$, determines how fast the noise parameter blows up with further iterations. This geometric series allows for a wide range of rationality levels, as shown by the smooth increasing line in Fig. 2.

The limit case $t = 1$ is of special interest. This would correspond to a flat line in Fig. 2 at height μ_0. For some games (e.g., matching pennies) the process will not converge when $t = 1$, but when it does, the limit probabilities, p^*, must be invariant under the logit map: $\phi^{\mu_0}(p^*) = p^*$. A fixed point of this type constitutes a "logit equilibrium," which is a special case of the quantal response equilibrium defined in McKelvey and Palfrey (1995). It is in this sense that the logit equilibrium arises as a limit of the noisy introspective process in (4) as $t \to 1$. When $t > 1$, the choice probabilities on the left side of (4) generally do not match the belief probabilities at any stage of the iterative process on the right. In other words, the introspective process allows for surprises, which are likely to occur in one-shot games.

For a t value between 2 and 4, say, the process converges quickly and the iterated probabilities remain more or less the same after several steps. Given the payoff parameters of the game, the introspective process in (4) predicts the probability with which a player chooses strategy S, and this prediction will vary systematically with the values of the error and telescope parameters. In Section 5 we use experimental data from 37 different matrix games to obtain maximum-likelihood estimates of the error and telescope parameters. These "pooled estimates" allow us to compare the introspective model with two equilibrium theories, i.e., logit and Nash equilibria. First, we extend the model to allow for general N-person games and discuss its relation with the notion of rationalizability.

4. Noisy rationalizability

We start by reviewing the concept of rationalizability, which requires some notation. Consider the normal-form game, G, which can be represented by the triplet (N, S, π), where N is a finite set of players indexed by i, each of whom chooses from a pure-strategy set S_i, with $|S_i|$ elements denoted by $s_{i,k}$, where $k = 1, \ldots, |S_i|$. Player i's payoff is given by the ith component of the payoff mapping, $\pi(.)$, which maps players' strategies into von Neumann–Morgenstern utilities, $\pi : S \to R^N$, where $S = \prod_k S_k$ is the Cartesian product of the strategy spaces. Player i's opponents are denoted $-i$ and P_{-i} denotes the projection from S to $S_{-i} = \prod_{k \neq i} S_k$.

[10] Selten's (1991) model of "anticipatory learning" applies the notion of different levels of rationality to a dynamic learning process in the context of a repeated, large-population game.

Rationalizability is based on the idea of iteratively eliminating those strategies that are never best responses for any (consistent) set of beliefs (Bernheim, 1984; Pearce, 1984). The set of rationalizable strategies can be constructed by defining $\lambda : S \to S$ as $\lambda(S) = \prod_i f_i(P_{-i}(S))$, where $f_i(s_{-i})$ is player i's rational best response to her opponents' strategy s_{-i}. Bernheim shows that the set of (point-) rationalizable strategies, *RS*, can be obtained by recursively applying the λ mapping. In other words, the (point-) rationalizable strategies are given by the limit set

$$RS = \lim_{n\to\infty} \lambda^n(S), \tag{5}$$

where λ^n is defined recursively as $\lambda^n(.) = \lambda(\lambda^{n-1}(.))$.[11]

The outcomes of the noisy introspection model can be defined analogously by replacing players' rational best responses, f_i, by logit best responses, ϕ_i^{μ}. When $\mu > 0$, the logit best responses assign non-zero probabilities to all pure strategies so we are naturally led to consider the extension of S to mixed strategies. Let M_i denote player i's mixed-strategies, i.e., the set of all probability distributions over S_i, and let $M = \prod_k M_k$. Suppose player i's first-order belief about rivals' play is given by $B_i^1 \in M_{-i}$. The expected payoff of choosing pure-strategy $s_{i,k}$ is given by $\pi^{\mathrm{e}}_{i,k}(B_i^1)$ and the probability of selecting $s_{i,k}$ follows from a generalization of (2):

$$\phi^{\mu}_{i,k}\big(B_i^1\big) = \frac{\exp(\pi^{\mathrm{e}}_{i,k}(B_i^1)/\mu)}{\sum_{l=1}^{|S_i|}\exp(\pi^{\mathrm{e}}_{i,l}(B_i^1)/\mu)}, \quad k = 1, \ldots, |S_i|. \tag{6}$$

The $|S_i|$ dimensional vector ϕ_i^{μ}, with elements $\phi^{\mu}_{i,k}$ for $k = 1, \ldots, |S_i|$, maps elements from M_{-i} to M_i. Define $\phi^{\mu}(M)$ to be the Cartesian product $\prod_i \phi_i^{\mu}(P_{-i}(M))$; it is this ϕ^{μ}-mapping that replaces the λ-mapping used by Bernheim (1984) in the construction of rationalizable strategies.

Recall that the notion of rationalizability assumes perfectly rational decision-making at any level of introspection, i.e., irrespective of the number of iterations. For this reason, the rationalizable strategies follow from applying the same λ-mapping recursively in (5). In contrast, we assume that higher levels of introspection become increasingly more noisy. The set of noisy rationalizable strategies, *NRS*, is obtained by recursively applying the ϕ^{μ}-mapping to M, using a higher error rate at every step:

$$NRS = \lim_{n\to\infty} \phi^{\mu_0} \circ \phi^{\mu_1} \circ \cdots \circ \phi^{\mu_n}(M), \tag{7}$$

where the $\{\mu_n\}_{n=0}^{\infty}$ form in increasing sequence with $\mu_\infty = \infty$. The latter assumption implies that the set of noisy rationalizable outcomes consists of a single element of M, since ϕ^{μ} for $\mu = \infty$ maps M into a single point, corresponding to uniform belief probabilities for all players. The interpretation is that players' higher-order beliefs about what others think about what others think about ... etc, become more and more diffuse and their (infinite) thought processes start out with uniform beliefs. The *outcome* of

[11] Bernheim (1984) and Pearce (1984) also extend the notion of rationalizability to include mixed strategies. We focus on the pure-strategy or point-rationalizable strategy set, to clarify the relation with our noisy introspection model.

their thought processes, however, generally corresponds to a different element of M, i.e., players' choice probabilities are not uniform. To summarize, while the set of rationalizable strategies generally consists of more than one point, the noisy rationalizable (mixed-) strategy is always unique, even in games with multiple Nash equilibria.

So far, we have implicitly assumed that the iterative process in (7) converges. The next theorem, which summarizes the main result of this section, establishes that this is indeed the case (see Appendix A for a proof).[12]

Theorem. *Let $\mu_0, \mu_1, \mu_2, \ldots$ denote a sequence of increasing and strictly positive error rates that diverges to infinity, and let ϕ^μ be the vector of logit best response mapping whose components are defined in* (6). *The sequence $\phi^{\mu_0}(p_0)$, $\phi^{\mu_0}(\phi^{\mu_1}(p_0))$, $\phi^{\mu_0}(\phi^{\mu_1}(\phi^{\mu_2}(p_0)))$, ... converges to a unique point* (*the noisy rationalizable strategy*) *independent of the initial starting point p_0.*

5. Experimental evidence

Guyer and Rapoport (1972) report an experiment in which 214 subjects played a large number of 2×2 matrix games, without feedback, in order to preserve the "one-shot" nature of the interaction. There were 37 basic games, six of which are shown in Tables 1 and 2.[13] In each game, strategy S is the maximin strategy, and the proportions of S choices for the games are shown by the dark lines in Fig. 3 for Row (top panel) and Column (bottom panel). The first three games, shown on the left side of each panel, are labeled DS at the bottom, which refers to the fact that S is a dominant strategy for these games. The dots at the top indicate that S is a Nash equilibrium for these three games. The next group of games also have dominant strategies, but these are asymmetric games, and hence are labeled as ADS at the bottom. Notice that the proportion of S choices (dark line) is high but not equal to 1 when it is a dominant strategy. The third group of games, labeled APD, are asymmetric prisoner's dilemma games, three of which are shown in Table 1. The payoffs that result from playing the dominant strategies are Pareto dominated by those of the "cooperative" outcome, which is not the case for the ADS games. The three coordination game variations in Table 2 are among those in the next group of asymmetric coordination games, labeled ACG. Recall that the coordination games have symmetric Nash equilibria at (S, S), (R, R), and a mixed equilibrium at an intermediate probability,

[12] The convergence proof is quite different from the fully rational case (e.g., Bernheim, 1984) where the recursive application of the λ-mapping produces a sequence of nested sets, i.e., $\lambda^n(S) \subseteq \lambda^{n-1}(S) \subseteq \cdots \subseteq \lambda(S) \subseteq S$. This is not the case for the recursive application of ϕ^μ-mapping in (7) where the error rate, μ, changes with each iteration. Instead we show in Appendix A that the sequence $\phi^{\mu_0}(p_0)$, $\phi^{\mu_0}(\phi^{\mu_1}(p_0))$, $\phi^{\mu_0}(\phi^{\mu_1}(\phi^{\mu_2}(p_0)))$, ... is a Cauchy sequence.

[13] Each game was permuted in all possible ways, by changing the labeling of players and decisions, for a total of 244 permutations. These were presented to subjects in a random order, by shuffling a deck of game cards for each person. Subjects made a decision for each of the 244 permutations, yielding a total of $214 \times 244 = 52{,}216$ decisions. After all decisions were made, subjects were paired, and their "point" earnings were determined by matching up the decisions for each of the games. Final earnings were determined by a \$2.50 fixed payment and a conversion of points into cash, with the conversion factor unreported.

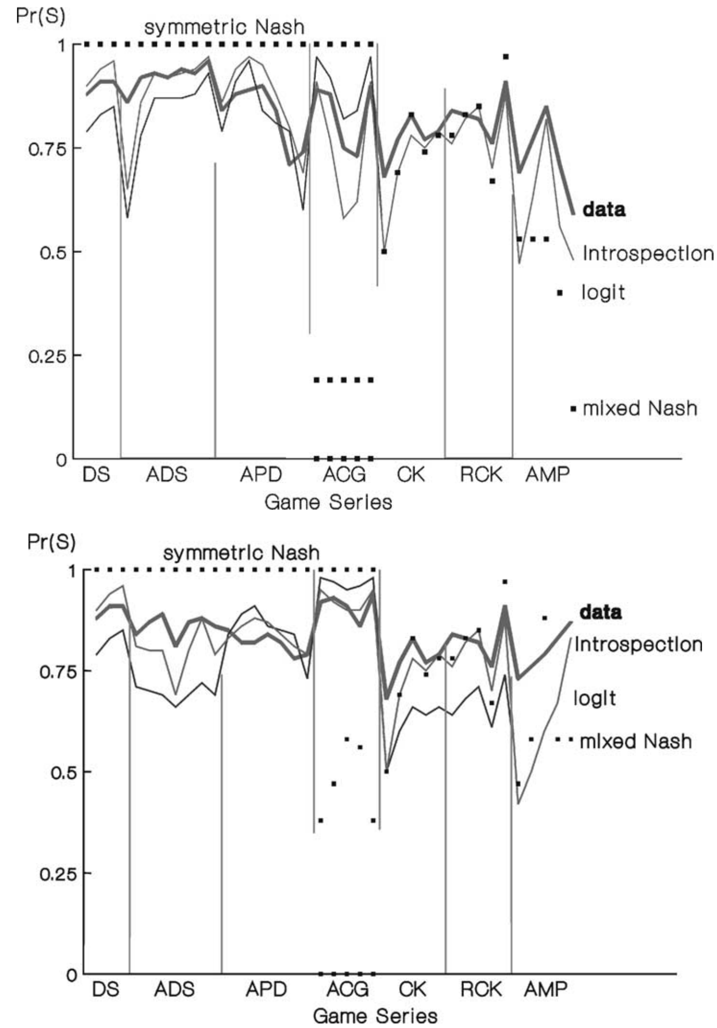

Fig. 3. Choice probabilities for Row (top) and Column (bottom): Guyer and Rapoport data (dark line), introspection (thin line), logit (dashed line), Nash (dots).

so there are black dots at the top, middle, and bottom parts of the graph for this series of games. The remaining games (discussed below), include games of "chicken" (CK) and "reverse chicken" (RCK). These games only have a single symmetric Nash equilibrium, which is in mixed strategies (indicated by the dots). Finally, the asymmetric matching pennies (AMP) games have a unique Nash equilibrium in mixed strategies.

The general picture that emerges from Fig. 3 is that choice proportions fall short of Nash predictions in the first three series of games, and choice proportions generally exceed Nash predictions in the final matching pennies games. In the two series of chicken games, the mixed-strategy Nash points are remarkably close to the data averages, a fact that seems to have gone unnoticed by Guyer and Rapoport, who focused instead on the proportions of maximin choices. In the other games behavior seems to be sensitive to changes in payoff asymmetries. For example, the large drop in the proportion of S choices by Row for the APD games occurs for the three variations in Table 1 that increase the attractiveness of the (R, R) outcome for Row. Similarly, the large drop in the incidence of safe choices in the first three ACG games is caused by the reduction in the riskiness of the R for Row, as shown in Table 2. Also notice that Column choices are relatively stable for these two series, which reflects the fact that "own payoff" effects seem to be more important.

The Nash equilibrium, strictly speaking, allows for no error, so any deviation is a rejection in an uninteresting technical sense. In order to evaluate the Nash concept statistically, it is necessary to append some randomness, and this was done with the logit formulation in (1). In the logit equilibrium, the probabilities that go into the expected payoffs on the right side of (1) should match the probabilities that come out on the left. Therefore, in asymmetric 2×2 games, there are two fixed-point equations (one for each player's probability of choosing S), which have to be solved, given μ. Therefore, each specific value of the error parameter, μ, produces equilibrium probabilities, $p_{\text{Row}}(\mu)$ and $p_{\text{Col}}(\mu)$, and the product of the equilibrium probabilities for each observed decision is used to form the likelihood function. Taking logs, the products become sums and the loglikelihood becomes:

$$\log L = \sum_{g=1}^{37} N_g \big(P^g_{\text{Row}} \log\big(p^g_{\text{Row}}(\mu)\big) + \big(1 - P^g_{\text{Row}}\big) \log\big(1 - p^g_{\text{Row}}(\mu)\big)\big) + \sum_{g=1}^{37} N_g \big(P^g_{\text{Col}} \log\big(p^g_{\text{Col}}(\mu)\big) + \big(1 - P^g_{\text{Col}}\big) \log\big(1 - p^g_{\text{Col}}(\mu)\big)\big), \tag{8}$$

where N_g is the number of decisions made by Row and Column in game $g = 1, \ldots, 37$, the capital P notation refers to the observed proportions of S choices for Row and Column players, and the lower case $p(\mu)$ notation refers to logit equilibrium probabilities.

The error parameter estimate is obtained by maximizing (8) with respect to μ, which yields the μ estimate of 6.6 that was used to construct the logit response lines in Fig. 1. The standard error of this estimate is 0.1, which allows rejection of the null hypothesis of $\mu = 0$ (Nash). The logit predictions for the 37 games are plotted in Fig. 3 as dashed lines. One way to measure how well the logit equilibrium tracks the observed data is to compute the mean of the squared distances between logit predictions and data averages (for both Row and Column, using all games).[14] Using percentages rather than probabilities, this mean-squared distance (MSD) is 379 for logit as compared to 490 for Nash (see also Table 3).[15] Even though the MSD for the logit predictions is lower than for Nash, the logit predictions are consistently too high or too low relative to the data in each of the game series, with the exception of the APD games.

[14] Some of the games with multiple Nash equilibria in the ACG, CK, and RCK series also have multiple logit equilibria for the μ value we estimated. We only plot the symmetric logit equilibria in these cases.

[15] This measure is calculated by adding the squared deviations (in percentages) for both Row and Column, and averaging across games.

Table 3
Mean-squared distances and loglikelihoods for alternative models

	Estimate	MSD[a]	Loglikelihood[b]
Nash ($\mu = 0, t = 1$)	NA	490	NA
Logit QRE ($t = 1$)	$\mu = 6.6$ (0.1)	379	$-25{,}165$
Introspection	$\mu = 4.4$ (0.1) $t = 4.1$ (0.1)	168	$-23{,}603$

[a] Mean of squared distances between predicted and actual percentages.
[b] The loglikelihood is given in (8) with N_g the number of decisions made in game g ($\sum_g N_g = 52{,}216$), P is the observed frequency of choice S, and p is the frequency of S predicted by the model. For comparison, the highest possible value of the loglikelihood is obtained when $P = p$ (i.e., the observed frequency equals the theoretical prediction), and equals $-22{,}460$. The random model, in which each decision is equally likely, results in a loglikelihood of $-33{,}193$.

Maximum likelihood techniques were also used to obtain parameter estimates for the introspection model. As before the loglikelihood is given by (8), where the logit equilibrium probabilities are replaced by introspection predictions, $p_{\mathrm{Row}}(\mu, t)$ and $p_{\mathrm{Col}}(\mu, t)$. These introspection probabilities are calculated by taking the limit of the composition of functions on the right side of (7) as the number of iterations goes to infinity. We approximate this by truncating the right side of (7) at ten iterations, which is justified by the fact that the introspection probabilities are virtually the same for every number of iterations greater than 5 for values of t that are greater than two. Using this procedure we obtain estimates of $\mu = 4.4$ (0.1) and $t = 4.1$ (0.1).[16] The standard errors in parentheses are small enough to allow rejection of the special cases of Nash ($\mu = 0$) and logit ($t = 1$).[17]

The introspection model further reduces the mean squared distance from 379 for the logit model to 168 for the introspection model. In addition, the introspection model has a much higher loglikelihood (see Table 3).[18] This improvement in fit is apparent in Fig. 3; whenever the logit predictions are too low, the introspection predictions tend to be higher (DS, ADS, CK, RCK, and AMP), and when the logit predictions are too high (ACG) the introspection predictions are lower. These qualitative comparisons are consistent with the intuition from Fig. 1: introspection predictions are generally lower than logit when logit response functions are positively sloped and are higher when they are negatively sloped.

[16] Notice that these estimates imply that $\mu_0 = 4.4$, $\mu_1 = 18.0$, and $\mu_2 = 73.8$. We also estimated a three-parameter model in which the levels of μ were not constrained to increase geometrically. The resulting estimates were somewhat similar $\mu_0 = 4.4$, $\mu_1 = 17.6$, and $\mu_2 = 53.8$, and these estimates yield essentially the same predicted introspection probabilities. Moreover, the loglikelihood of the three-parameter model, $-23{,}602$, is not significantly higher than the loglikelihood obtained from the two-parameter introspection model in the bottom row of Table 3.

[17] We also estimate the model separately for the three categories of games: games with a dominant strategy solution (games 1–17), games with multiple Nash equilibria (games 18–32), and games with a unique mixed-strategy equilibrium (games 33–37). The μ-estimates for these three categories of games are: 4.7 (0.1), 4.0 (0.1), and 3.4 (0.1), respectively. The t-estimates are 4.2 (0.3), 4.3 (0.1), and 3.9 (0.3), respectively.

[18] A standard loglikelihood ratio test involves computing twice the difference between the loglikelihoods of the nested models in Table 3. The test statistic associated with adding the introspection parameter, t, is greater than 3000, which exceeds the critical value for a chi-square test at any standard level of significance.

Table 4
A game of chicken (Row payoff, Column payoff)

	Column player's decision	
Row player's decision	S	R
S	12, 12	15, 32
R	32, 15	$-5, -5$

In addition, the introspection predictions match the accuracy of the Nash predictions in the chicken games, but like logit, generally do better than Nash in the other games.

One of the games where the introspection model predicts poorly is the first chicken game, shown just to the right of the dotted line that separates ACG and CK in Fig. 2. Consider the chicken game with payoffs shown in Table 4.[19] For both players the sum of payoffs for either decision is 27, so the mixed-strategy Nash equilibrium is to choose each decision with probability $1/2$. In this case, the best response functions intersect in the center of a graph like Fig. 1, at $(0.5, 0.5)$. The effect of adding noise is to round off the corners, leaving S-shaped logit response functions that still intersect in the center. This symmetry causes the symmetric logit and introspection equilibria to also be at 0.5. The Nash equilibrium produces an expected payoff of 14.5 for each decision, despite the fact that the payoff variance would be much higher for the risky decision. The data, in contrast to all three predictions, reveal that 67% of the choices were the safe decision. This suggests that the high rate of safe choices may be due to risk aversion.[20]

Finally, consider the data for the five asymmetric matching pennies games shown on the right side of Fig. 3. In the first three games, only Row's payoffs were changed. The Nash predictions are constant for Row, since Row's probability is determined by the requirement that Column be indifferent. Therefore, the dots that show Row's Nash predictions for the first three AMP games are on a horizontal line. The dark data line for Row is sharply increasing for these games, indicating that Row's choice probabilities are sensitive to "own-payoff" effects. Since Row's stochastic best-response function shifts in

[19] The chicken and reverse-chicken games are similar in that the best response to aggressive behavior (R) is passive (S) and vice versa, so there are asymmetric Nash equilibria (S, R) and (R, S), and there is a symmetric equilibrium in mixed strategies that is shown by the solid dots in Fig. 2. The only difference is that for each R/S outcome, the player choosing R earns more in the chicken game and the person choosing S earns more in the reverse chicken game.

[20] To test this conjecture we incorporated risk aversion into the noisy introspection model. We assumed constant relative risk aversion so that the utility of an amount x is x^{1-r}, where the risk aversion parameter, r, satisfies $0 < r < 1$. We added a constant (17) to all payoffs to ensure that the lowest payoff for any of the 37 games would be at least 1. The results of this estimation are: $r = 0.46$ (0.02), $\mu = 0.62$ (0.05), and $t = 4.6$ (0.1). (The estimated error parameter is lower than the risk-neutral estimate because the power function expected utility numbers are much lower than the expected payoffs.) The hybrid introspection/risk-aversion model has a much lower mean squared deviation of 78, as compared with 168 for the model without risk aversion. (We also estimated a risk aversion parameter for the logit model: $r = 0.45$, which reduces the mean squared deviation from 379 with risk neutrality to 343 with risk aversion.) The improved fit is largely in the first chicken game and the asymmetric matching pennies (AMP) games. Of course, adding an extra parameter increases the danger of "data-mining," and the reader will have to decide whether the improved fit is worth the cost.

the direction of the safe strategy when one of the payoffs for that strategy is increased, the logit predictions determined by the intersection of the smooth stochastic best responses also shifts towards higher probability of S. It is apparent from the top part of Fig. 3 that the introspection predictions also track these own-payoff effects, which is not surprising since these predictions are determined by the iteration of stochastic best responses. Similarly, the payoff changes in the second, fourth, and fifth AMP games affect only Column, and therefore the Nash predictions for Column are constant. Again the logit and introspection model correctly predict the own-payoff effects observed in the data for Column shown in the bottom part of Fig. 3.

6. Conclusion

Many strategic encounters are unique, non-repeated interactions. Equilibrium concepts that build in “rational expectations” about others’ decisions may not be appropriate in such cases. Without an opportunity to learn, players must think about others’ decisions, others’ theories of one’s own decisions, etc., but such speculation is likely to become increasingly noisy with successive iterations. In this paper we propose a general model of iterated noisy introspection and prove convergence (existence and uniqueness). Parsimonious versions of this model were estimated using data from thirty-seven 2×2 matrix games, and the model predictions are more accurate than those of equilibrium theories, both with noise (logit) and without noise (Nash).

The mixed-strategy Nash equilibrium is remarkably accurate in symmetric games (e.g., chicken), but it is quite inaccurate in some matching pennies games where the only Nash equilibrium is mixed. The reason for this difference is that human subjects do not seem to follow the mixed-strategy prediction that decision probabilities depend only on *other’s* payoffs. In the symmetric chicken games, this asymmetry bias does not occur because the parameter changes affect both players in the same manner. Moreover, Nash predictions do not pick up systematic “own-payoff” effects that alter quantitative but not qualitative payoff comparisons. In contrast, the logit (quantal response) equilibrium is sensitive to magnitudes of payoff differences. The logit equilibrium has provided remarkably accurate predictions of behavior in games with learning opportunities (McKelvey and Palfrey, 1995; Goeree and Holt, 1999). In one-shot games, however, the logit predictions tend to be systematically biased: above the data for games with negatively sloped stochastic response functions and below the data for games with positively sloped stochastic response functions. The model of noisy introspection follows the Nash predictions in games where they are on track, and it is generally much closer to the data in other games.

Acknowledgments

This research was funded in part by the National Science Foundation (SBR-9818683 and SBR-0094800). We are grateful to two anonymous referees and one associate editor for useful suggestions.

Appendix A. Existence and uniqueness

Theorem. *Let $\mu_0, \mu_1, \mu_2, \ldots$ denote a sequence of increasing and strictly positive error rates that diverges to infinity, and let ϕ^μ be the vector of logit best response mapping whose components are defined in* (6). *The sequence $\phi^{\mu_0}(p_0)$, $\phi^{\mu_0}(\phi^{\mu_1}(p_0))$, $\phi^{\mu_0}(\phi^{\mu_1}(\phi^{\mu_2}(p_0)))$, $\ldots$ converges to a unique point* (*the noisy rationalizable strategy*) *independent of the initial starting point p_0.*

Proof. Define

$$\Phi_n \equiv \phi^{\mu_0} \circ \phi^{\mu_1} \circ \cdots \circ \phi^{\mu_n},$$

where $\{\mu_n\}_{n=0}^\infty$ is an increasing sequence of error rates that diverges to infinity, and $\phi^\mu : M \to M$ is defined by $\phi^\mu(M) = \prod_i \phi_i^\mu(P_{-i}(M))$, where P_{-i} is the projection from S to $S_{-i} = \prod_{k \neq i} S_k$ and player i's noisy best response ϕ_i^μ has components:

$$\phi_{i,k}^\mu\big(B_i^1\big) = \frac{g(\pi_{i,k}^{\mathrm{e}}(B_i^1)/\mu)}{\sum_{l=1}^{|S_i|} g(\pi_{i,l}^{\mathrm{e}}(B_i^1)/\mu)}, \quad k = 1, \ldots, |S_i|, \tag{A.1}$$

where $\pi_{i,k}^{\mathrm{e}}(B_i^1)$ is player i's expected payoff from choosing the pure strategy $s_{i,k}$ when her beliefs about the others' play are given by (some arbitrary) $B_i^1 \in M_{-i}$. The $g(.)$ function on the right-side of (A.1) is some strictly positive, strictly increasing, differentiable function on $\mathbb{R}$. Note that (A.1) reduces to the logit rule discussed in the main text when $g(x) = \exp(x)$.

We have to show that the sequence $\{\Phi_n(p_0)\}_{n=0}^\infty$ converges and that the limit point is independent of p_0. The latter claim follows from the assumption that the error rates diverge to infinity and ϕ_i^μ for $\mu = \infty$ maps player i's entire probability simplex, M_i, into a single point, i.e., the centroid C_i corresponding to uniform probabilities: $C_{i,k} = 1/|S_i|$ for $k = 1, \ldots, |S_i|$. Let C denote the vector that results by concatenating the C_i for $i = 1, \ldots, N$. The limit point of $\{\Phi_n(C)\}_{n=0}^\infty$ (if it exists) will thus be the same as that of $\{\Phi_n(p_0)\}_{n=0}^\infty$ for all p_0.

We prove convergence of $\{\Phi_n(C)\}_{n=0}^\infty$ by showing that it is a Cauchy sequence, i.e., for all $\varepsilon > 0$ there exist n such that the distance $d(\Phi_n(C), \Phi_m(C)) < \varepsilon$ for all $m > n$. Here the distance between two points in M can be defined as: $d(p, q) = \sum_{i=1}^N \max_{i=1,\ldots,K} |p_i - q_i|$. Note that for $m > n$ we can write $\Phi_m(C) = \Phi_n(c(n, m))$, where the lower case $c(n, m)$ is defined as

$$c(n, m) = \phi^{\mu_{n+1}} \circ \cdots \circ \phi^{\mu_m}(C).$$

Since Φ_n is a composition of continuous functions it is itself continuous. Hence for all $\varepsilon > 0$ there exist δ such that $d(\Phi_n(C), \Phi_n(c(n, m))) < \varepsilon$ if $d(C, c(n, m)) < \delta$. So the proof follows if we can show that for all $\delta > 0$ there exist an n such that $d(C, c(n, m)) < \delta$ for all $m > n$.

Let C_i, $c_i(n, m)$ denote the projection of C, $c(n, m)$ onto M_i, and let $c_{i,\max}(n, m)$ and $c_{i,\min}(n, m)$ be the largest and smallest element of $c_i(n, m)$, respectively. The distance between C_i and $c_i(n, m)$ is the greater of $c_{i,\max}(n, m) - 1/|S_i|$ and $1/|S_i| - c_{i,\min}(n, m)$. Hence, the distance $d(C_i, c_i(n, m))$ is no greater than the sum of these two expressions, which equals: $c_{i,\max}(n,m) - c_{i,\min}(n,m)$. Denote the highest and lowest possible payoffs for player i by $\pi_{i,\max}$ and $\pi_{i,\min}$, respectively (which are assumed to be finite), then $d(C, c(n,m))$ can be bounded by

$$\begin{aligned} d\big(C, c(n,m)\big) &= \sum_{i=1}^{N} d\big(C_i, c_i(n,m)\big) \\ &\leqslant \sum_{i=1}^{N} \frac{g(\pi_{i,\max}/\mu_{n+1}) - g(\pi_{i,\min}/\mu_{n+1})}{|S_i| g(\pi_{i,\min}/\mu_{n+1})}, \end{aligned} \tag{A.2}$$

for all $m > n$. Since $g(.)$ is continuous and the error rate μ_{n+1} diverges as n grows large, the numerator on the far right side of (A.2) can be made arbitrarily small (while the denominator limits to $|S_i| g(0) > 0$) by choosing n large enough. □

References

Bernheim, D., 1984. Rationalizable strategic behavior. Econometrica 52, 1007–1028.

Capra, C.M., 1998. Noisy expectation formation in one-shot games: an application to the entry game. Discussion paper. Univ. of Virginia.

Goeree, J.K., Holt, C.A., 1999. Stochastic game theory: for playing games, not just for doing theory. Proc. Nat. Acad. Sci. 96, 10564–10567.

Goeree, J.K., Holt, C.A., 2001. Ten little treasures of game theory and ten intuitive contradictions. Amer. Econ. Rev. 91, 1402–1422.

Guyer, M.J., Rapoport, A., 1972. 2 × 2 games played once. J. Conflict Resolution 16, 409–431.

Harsanyi, J., 1967–1968. Games with incomplete information played by Bayesian players. Manage. Sci. 14, 159–182, 320–334, 486–502.

Harsanyi, J., Selten, R., 1988. A General Theory of Equilibrium Selection in Games. MIT Press, Cambridge.

Kübler, D., Weizsäcker, G., 2000. Limited depth of reasoning and failure of cascade formation in the laboratory. Discussion paper. Harvard University.

Luce, D., 1959. Individual Choice Behavior. Wiley, New York.

McKelvey, R., Palfrey, T., 1995. Quantal response equilibria for normal form games. Games Econ. Behav. 10, 6–38.

Nasar, Sylvia, 1998. A Beautiful Mind. Simon and Schuster, New York.

Olcina, G., Urbano, A., 1994. Introspection and equilibrium selection in 2 × 2 matrix games. Int. J. Game Theory 23, 183–206.

Pearce, D., 1984. Rationalizable strategic behavior and the problem of perfection. Econometrica 52, 1029–1050.

Selten, R., 1991. Anticipatory learning in two-person games. In: Selten, R. (Ed.), Game Equilibrium Models I. Springer-Verlag, Berlin, pp. 98–153.

Stahl, D., Wilson, P., 1995. On players' models of other players: theory and experimental evidence. Games Econ. Behav. 10, 218–254.