

A Model of Noisy Introspection

*Jacob K. Goeree and Charles A. Holt**

Department of Economics, Rouss Hall, University of Virginia, Charlottesville, VA 22901

February 2000

Abstract. This paper presents a theoretical model of noisy introspection designed to explain behavior in games played only once. The equilibrium 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. The paper contains a convergence proof and reports estimated introspection and error parameters for data from 37 one-shot matrix games. The accuracy of the model is compared with that of two alternative approaches: the Nash equilibrium and the logit quantal response equilibrium.

Keywords: game theory, introspection, Nash equilibrium, experiments.

JEL Codes: C72, C92

I. 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. An 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 was formalized by John Nash, in a 1950 *Proceedings of the National Academy of Sciences* paper that became the basis of his Nobel Prize in Economics almost fifty years later. In graduate school, Nash and his fellow students learned about John von Neumann and Oscar Morgenstern's work on zero-sum games. While applications of zero-sum games were limited to simple parlor games in which one player's gain is another one's loss, the U.S. Air Force was already trying to apply game theory to strategic bombing plans. Ironically, it was Nash, the graduate student, whose work helped game theory achieve the central role that von

* This research was funded in part by the National Science Foundation (SBR-9818683).

Neumann and Morgenstern had envisioned. In economics, the Nash equilibrium is now as frequently used to model small-group interactions as the classic notions of supply and demand are used in the analysis of "thick" markets with large numbers of traders.

Biological and economic market interactions often occur repeatedly, and the notion of equilibrium is appropriate when decision makers have learned what others can be expected to do. Many recent applications of game theory outside of economics, however, are more accurately modeled as games played only once, e.g. legal disputes 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. In fact, it is well known that precise iterated thinking of this kind will lead 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.

II. 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).

		Column Player's Decision	
		S	R
Row Player's Decision	S	6, 12	25, -4
	R	-4, 18	9 / 18 / 24, 25

Table 1. A One-Sided Prisoner's Dilemma Game (Row Payoff, Column Payoff):
Three Variations Affecting Row's "Incentive to Defect"

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

		Column Player's Decision	
		<i>S</i>	<i>R</i>
Row Player's Decision	<i>S</i>	10, 19	4, -7
	<i>R</i>	-14 / -7 / -1, 4	19, 10

Table 2. A Coordination Game with Multiple Equilibria:
Three Variations that Reduce the Risk for Row to "Force" Column

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 .39. The crossings of the best response 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 "x" 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 John Harsanyi and Reinhard Selten (1988), who were the others sharing the game theory Nobel prize with Nash. 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

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

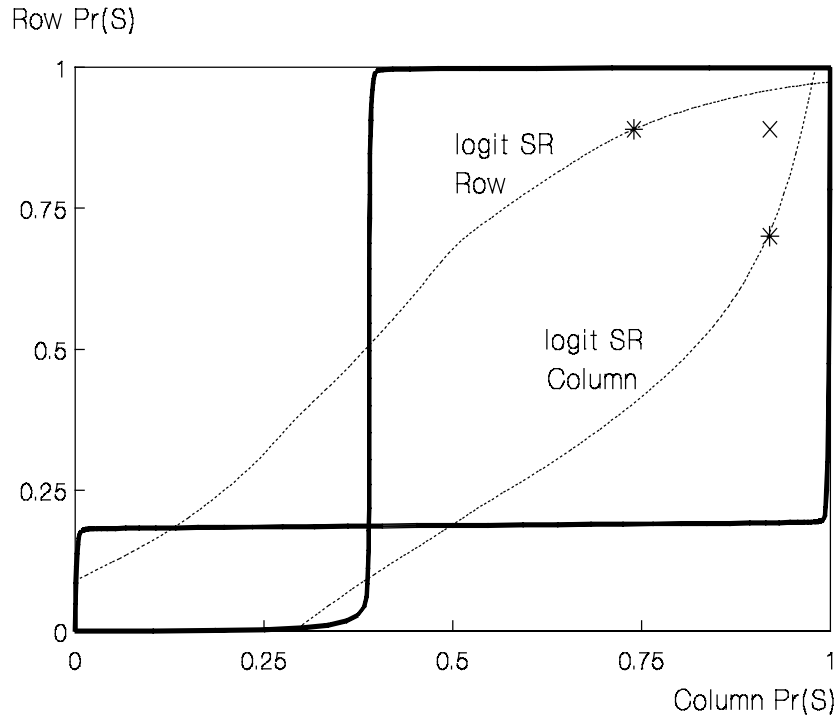


Figure 1. A Coordination Game:
Best Responses (Dark Lines) and Logit Stochastic Responses (Dotted Lines)

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 When there are only two possible options with expected payoffs $\pi^e(S)$ and $\pi^e(R)$, the logit probability of choosing strategy S , say, is given by:

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 (1967-68) in an equilibrium analysis, which is closely related to the quantal response equilibrium discussed below.

$$p = \frac{\exp(\pi^e(S)/\mu)}{\exp(\pi^e(S)/\mu) + \exp(\pi^e(R)/\mu)}, \quad (1)$$

and the probability of choosing R is simply $1 - p$. 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. Indeed, the slight rounding off of the corners of the response functions in Figure 1 is due to the fact that these were drawn for a low μ value of 0.1 (instead of 0). An 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 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 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).³ 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 Figures 3, that map belief probabilities into actions that occur with the same probability. In all but the simplest games, equilibrium concepts will have

³ 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.

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 Figure 1. The logit quantal response equilibrium 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 .70, 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 .7 chance of S , whereas Column plays S with probability 0.92. Notice that the asterisk points pull decisions away from the logit 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 Figure 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.⁴

The next section presents a model of noisy introspection that formalizes the intuition from these examples. We used data from one-shot games to estimate the model parameters, and section IV contains an assessment of how the model compares with the Nash and logit quantal response predictions. The final section concludes.

⁴ These slope effects may be either negated or reinforced if the relevant value of μ is different for logit and introspection equilibrium, since lower error rates will push the logit intersection closer to a Nash equilibrium. Estimates for μ are reported in section IV below.

III. 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. At the same time, we relax the assumption of perfect maximization by injecting some noise into the system via the logit choice function in (1), which predicts that decision probabilities are positively (but not perfectly) related to expected payoffs.

The expected payoffs in (1) are determined by a player's beliefs about the rival. For the games considered in this paper there are only two possible choices (R and S), and a player's beliefs can therefore be represented by a single probability, q , which is the probability with which the opponent chooses S . For instance, when a player thinks both options are equally likely to be chosen by the opponent, q is simply $1/2$. This belief probability, q , determines the expected payoffs for decisions R and S , and the logit rule (1) maps these expected payoffs into choice probabilities. Also the choice probabilities can be represented by a single number, p (the probability of choosing S). A convenient way to represent the transformation from belief to choice probabilities is: $p = \phi_\mu(q)$, where ϕ_μ represents the map on the right side of (1) and the noise parameter, μ , determines how sensitive choice probabilities are to payoff differences.

Of course, the process does not necessarily stop after one iteration. The logit best responses to uniform probabilities can again be used to determine expected payoffs, which lead to the next level of conjectured probabilities via (1). Iterating in this manner, the thought process will produce more refined conjectures.⁵ However, since there is likely to be more error associated with beliefs about others' beliefs about ..., etc., every iteration will introduce more noise. In other words, the iterative procedure will become increasingly complex and we shall assume that the error rate grows (geometrically) with each further iteration. For example, with

⁵ For an alternative approach, see Capra (1998). In her model, beliefs are represented by degenerate distributions that put all probability mass at a single point. The location of the belief points is, *ex ante*, stochastic. A deterministic model of introspection in 2×2 games is presented in Olcina and Urbano (1994). This model uses an axiomatic approach to select a prior distribution, which is revised by a simulated learning process. This latter process is essentially a partial adjustment from current beliefs to best responses to current beliefs. The model has the attractive property that it selects the risk-dominant Nash equilibrium. 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 of the model in light of the one-shot data reported below.

a vector of initial belief probabilities, p_0 , the twice iterated choice probability is given by a noisy (μ) response to an even noisier ($t\mu$) response to p_0 , i.e. $p = \phi_\mu(\phi_{t\mu}(p_0))$. This process can be iterated backwards, with the "telescoping" parameter $t > 1$ determining how fast the noise parameter blows up with further iterations; the error rate for the n th iteration is given by $t^{n-1}\mu$. We are interested in the choice probabilities in the limit as the number of iterations goes to infinity:

$$p = \lim_{n \rightarrow \infty} \phi_\mu(\phi_{t\mu}(\cdots \phi_{t^{n-1}\mu}(p_0))). \quad (2)$$

In the Appendix we use continuity arguments to show that this limit is well defined for $t > 1$. Since ϕ_∞ maps the whole probability simplex to a single point, the process is independent of the initial belief vector p_0 . The geometrically increasing error rate in (2) captures the idea that it becomes more and more complex to think further back. 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.⁶ Finally, the limit case $t = 1$ is of special interest. 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: $\phi_\mu(p^*) = p^*$. 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 case of the noisy introspective process in (2). When $t > 1$, the choice probabilities on the left side of (2) generally do not match the belief probabilities at any stage of the iterative process on the right. In other

⁶ The convergence proof in the Appendix can easily be extended to allow the telescope parameter to be person specific and to differ for different levels of introspection. The only restriction is that the telescope parameters be strictly positive and that there be more noise at higher levels of iteration. Instead of increases in error parameters from μ to $t\mu$ to $t^2\mu$, for example, the increase can be from μ to $t_1\mu$ to $t_2\mu$, where $1 < t_1 < t_2$. This formulation is flexible and allows many special cases. For instance, if $\mu = \infty$, (2) produces uniform choice probabilities as is the case for Stahl and Wilson's (1995) "level-0 rationality." If t goes to infinity, we have a logit best response to a uniform distribution, which corresponds to Stahl and Wilson's "level 1 rationality." Higher "levels of rationality" can be generated similarly. Indeed, in its most general form (2) allows for individual differences and for telescope parameters that are zero up to some point and then go to infinity. Rather than assuming a fixed number of iterations, (2) allows parsimonious representation of a wide range of rationality levels. 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.

words, the introspective process allows for surprises, which are likely to occur in oneshot games.

Given the payoff parameters, the introspective process in (2) 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 the next section, we use experimental data to obtain maximum-likelihood estimates these parameters. These "pooled estimates" allow us to compare the introspective model with two equilibrium theories, i.e. logit and Nash equilibria.

IV. 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 Figures 1 and 2.⁷ 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 Figure 2 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 this game. 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 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, so there are black dots at the top, middle, and bottom parts of the graph for this series of games. The remaining games, "chicken" (CK) and "reverse chicken" (RCK), only have a single symmetric Nash equilibrium, which is in mixed strategies

⁷ 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 permutation, 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.

(indicated by the dots). Finally, the asymmetric matching pennies (AMP) games have a unique Nash equilibrium in mixed strategies. The structure of these games will be discussed below.

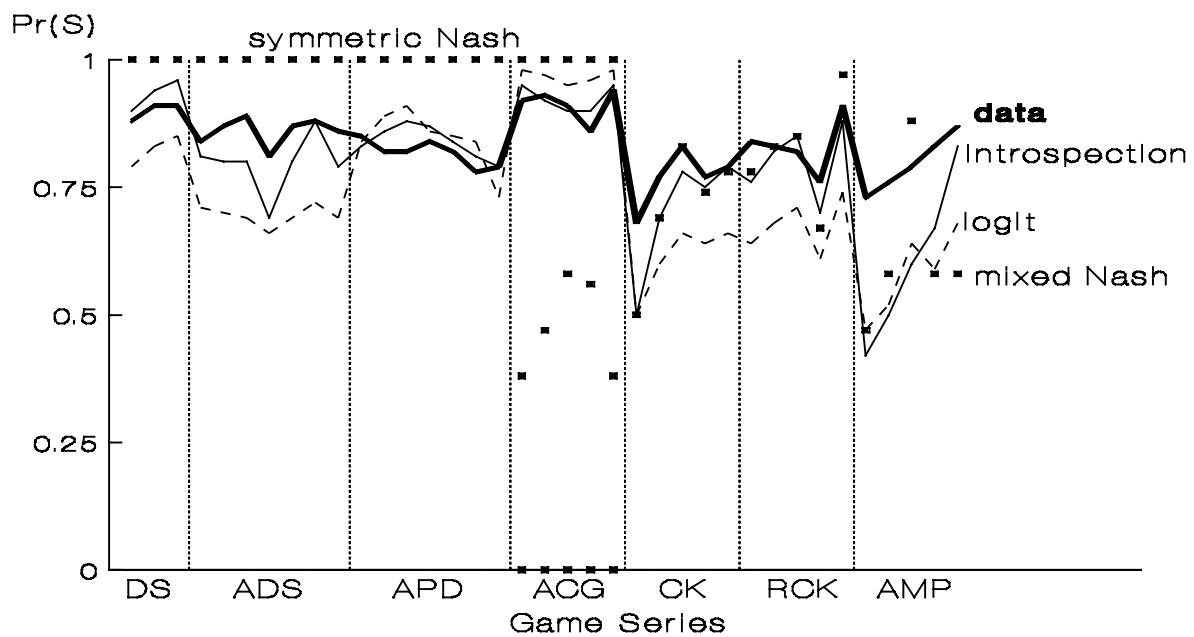
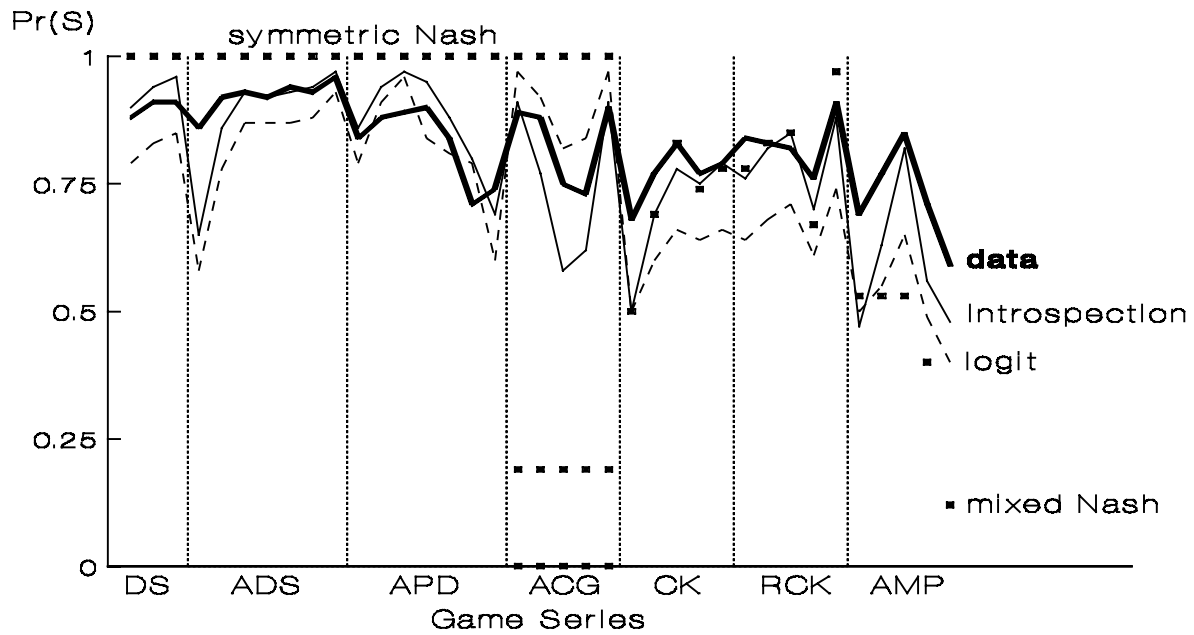


Figure 2. Choice Probabilities for Row (Top) and Column (Bottom): Guyer and Rapoport Data (dark line), Introspection (thin line), Logit (dashed line), Nash (dots)

The general picture that emerges from Figure 2 is that choice proportions fall short of Nash predictions in the first four series of games, and choice proportions generally exceed Nash predictions in the final matching pennies games.⁸ In the other 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. The large drop on 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. 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). Each specific value of the error parameter, μ , produces a probability for each decision in all games, and the product of the probabilities of the observed decisions is a likelihood function that is maximized by the μ of 6.6 (.1) that was used to construct the logit response lines in Figure 1. The standard error of the estimate is .1, which allows rejection of null hypothesis of $\mu = 0$ (Nash). The logit predictions for the 37 games are plotted in Figure as dashed lines.⁹ 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). Using percentages rather than probabilities, this mean-squared distance (MSD) is 379 for logit as compared to 490 for Nash (see also Table 5). 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. Maximum

⁸ Since each of the averages is determined on the basis of more than a thousand decisions, these differences between Nash predictions and actual choice frequencies are highly significant.

⁹ 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.

likelihood techniques were also used to obtain parameter estimates for introspection model: $\mu = 4.4$ (.1) and $t = 4.1$ (.1). The standard errors in parentheses are small enough to allow rejection of the special cases of Nash ($\mu = 0$) and logit ($t = 1$). 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 5). This improvement in fit is apparent in Figure 2; 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 Figure 1: introspection predictions are generally lower than logit when logit response functions are positively sloped and are higher when they are negatively sloped. 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.

		Column Player's Decision	
		<i>S</i>	<i>R</i>
Row Player's Decision	<i>S</i>	12, 12	15, 32
	<i>R</i>	32, 15	-5, -5

Table 3. A Game of Chicken (Row Payoff, Column Payoff)

The games where the introspection model predicts poorly are the first chicken game, shown just to the right of the dotted divider line in Figure 2, and the matching pennies games. In almost all of these cases, the predicted probability of the *S* outcome is too low. Consider the chicken game with payoffs shown in Table 3.¹⁰ For both players the sum of payoffs for either

¹⁰ 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 Figure 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.

decision is 27, so 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 Figure 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.

		Column Player's Decision	
		<i>S</i>	<i>R</i>
Row Player's Decision	<i>S</i>	9 / 15 / 24, 5	5, -10
	<i>R</i>	26, 9	-10, 26

Table 4. Matching Pennies (Row Payoff, Column Payoff)

Table 4 shows another game in which the each player's payoffs from making an *S* choice are considerably less variable than the payoffs for *R*. Each of the three variations in the table are "matching pennies games" in the sense that, starting from a match (*S*, *S*) or (*R*, *R*), Row would want to deviate to create a mismatch, but then Column would want to deviate to create a match. In the first variation, the payoff sums for each player's decisions are approximately equal, (9 + 5) and (26 - 10). Therefore, the mixed strategy Nash probabilities, which make each player indifferent and willing to randomize, are close to 0.5, i.e. 0.53 for Row and 0.47 for Column. These probabilities are determined by the intersection of the two dark best response lines in Figure 3. Adding some noise to the model ($\mu = 6.6$) rounds off the corners of the best response lines, making two thin lines that produce a logit equilibrium intersection at about the same location in the graph (please ignore the dotted lines for now). The introspection predictions are nearby. To summarize, all three models predict similar outcomes in the center of the figure, but the actual proportions of *S* choices were .69 for Row and .73 for Column, located at the

lower left asterisk in the figure.

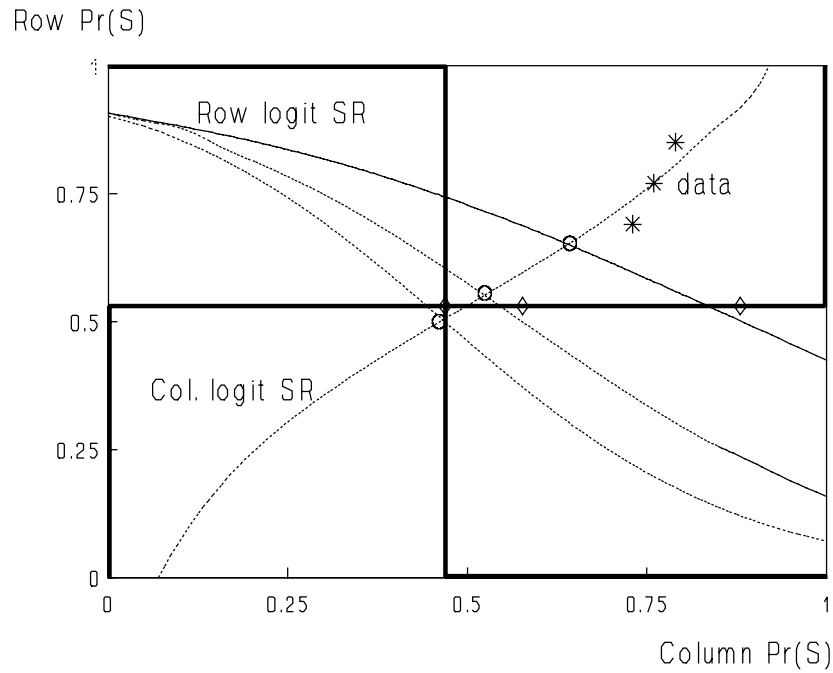


Figure 3. Best Responses for a Matching Pennies Game (Dark Lines), Logit Responses (Thin Lines), Nash Equilibria (Diamonds), Logit Equilibria (Circles), and Data (Asterisks)

The three variations of the matching pennies game in Table 4 increase the attractiveness of the (S, S) outcome for Row, from 9 to 15 to 24. Not surprisingly, this shift raises the proportion of S choices, as indicated by the shift of the asterisk data points upward in the figure. This "own-payoff effect" for Row is not predicted by a Nash analysis. Without noise, an increase in the attractiveness of the S decision for Row just shifts the (vertical) Row best response line to the right, producing new intersections at the diamond markers. The dark best response line for Column is horizontal, so the shifts in Row's best response line do not increase the probability that Row plays S in a mixed-strategy Nash equilibrium. This is analogous to an economic market in which an increase in demand does not raise price when supply is perfectly elastic (horizontal). The situation is qualitatively different with noise, however. In this case, Column's logit response line is positively sloped, and the right-ward shifts in Row's logit response (dotted) lines produce higher logit predictions for proportions of S decisions. These predictions (circles) show that Row's logit choice probabilities are sensitive to Row's own payoffs when noise is introduced.

But the three intersection points do not track the high propensity for players to use their safe, S , choices, as indicated by the three asterisk markers for the observed choice proportions in the upper-right part of the figure. Again, one possible explanation is an aversion to high-risk decisions.

Risk aversion is a concept that dates back at least to Daniel Bernoulli (1738), who observed that people were unwilling to pay large amounts for very high-risk gambles with infinite expected values. He proposed that the utility of money be a concave function of wealth, exhibiting a diminishing marginal value of additional money. This feature causes the money value of a gamble to be less than its mathematical expected value. The notion of nonlinear utility was formalized by von Neumann and Morgenstern (1944). Risk aversion is determined by the concavity of the function, and the most commonly used empirical model of risk aversion is probably the constant relative risk aversion form that specifies the utility of an amount x to be x^{1-r} , where the risk aversion parameter is in the unit interval: $0 < r < 1$. Estimates of this risk aversion parameter generally range from 0.3 to something close to 1. For example, Goeree, Holt, and Palfrey (1999) estimate a risk aversion parameter of about 0.5 for data from auction experiments. Risk aversion can be incorporated into the introspection model by replacing expected payoffs with expected utility expressions.

In order to obtain results that could be compared with our previous estimates, we used the constant relative risk aversion function shown above, with a constant (17) added to all payoffs in order to ensure that the lowest payoff for any of the 37 games would be at least 1. The results of this estimation are: $r = .46$ (.02), $\mu = .62$ (.05), and $t = 4.6$ (.1). The risk-aversion parameter obtained in this manner is approximately the same as that obtained for the auction data, a very different game with a different subject pool conducted over 25 years later.¹¹ 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.¹² The improved fit is largely in the first chicken

¹¹ 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.

¹² 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.

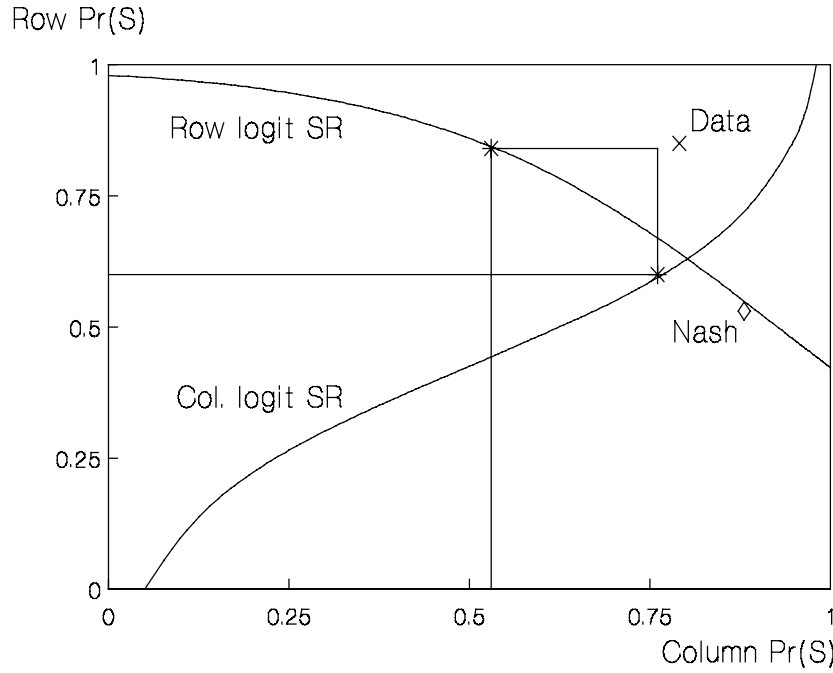


Figure 4. Logit Stochastic Best Responses for $\mu = 4.4$ (Curved Lines), Introspection Beliefs (Asterisks), and Guyer and Rapoport Data (\times)

game and the asymmetric matching pennies (AMP) games.¹³

Figure 4 illustrates the way in which the introspection model with risk aversion tracks the data for the most asymmetric of the three matching pennies games in Table 4, i.e. the one with payoffs of (24, 5) for the (S, S) outcome. The effect of risk aversion ($r = .46$) is to increase the probability of choosing the safe decision, which shifts the logit best response lines up for Row and to the right for Column. With the lower μ value estimated for the introspection model, the new intersection is not far from the Nash mixed equilibrium (the diamond marker) for this game. Unlike a logit equilibrium where the belief probabilities are determined by the intersection, belief probabilities in the introspection model are determined by iterated stochastic responses. It is possible to calculate beliefs about the other player's decision, stopping just before the final iteration in equation (2). In this case, Row's beliefs are that Column will play S with probability 0.53, as shown by the vertical dashed line in Figure 4, which intersect Row's stochastic best

¹³ Of course, adding an extra parameter increases the danger of "data-mining," and the reader will have to decide whether the improved fit and consistency with some previous estimates is worth the cost.

response at a point marked with an asterisk. Similarly, column's beliefs, calculated iteratively, are that Row will choose S with probability 0.6, as indicated by the horizontal dashed line in the figure. The two asterisk marks on the stochastic response functions determine the prediction of the introspection model, where the dashed lines meet just to the left of the "x" that marks the actual averages. Here again we see the slope effects, which tend to cause the introspective prediction to be lower than the logit equilibrium when the logit response line is positively sloped (Column) and higher than the logit equilibrium when the logit response line is negatively sloped (Row). The introspection model with risk aversion also improves the fit for all of the other matching pennies games.

Table 5. Mean-Squared Distances and Loglikelihoods For Alternative Models.

	MSD ^a	Loglikelihood ^b
Nash	490	NA
Logit QRE	379	-25,165
Logit QRE + Risk Aversion	343	-25,093
Introspection	168	-23,603
Introspection + Risk Aversion	78	-23,138

^a Mean of squared distances between predicted and actual percentages.

^b The loglikelihood is given by $\sum_g N_g (p \ln(p_m) + (1-p) \ln(1-p_m))$, with N_g the number of decisions made in game g ($\sum_g N_g = 52,216$), p is the actual frequency of choice S , and p_m is the frequency of S predicted by the model. The maximum possible loglikelihood occurs for $p = p_m$ and is given by -22,460. The random model, in which each decision is chosen with probability .5, results in a loglikelihood of -33,193.

One issue is whether the introspection model, with or without risk aversion, can explain behavior in more complex games. The results thusfar are encouraging. In Goeree and Holt (2000) we use the introspection parameters estimated from these 37 matrix games to explain the qualitative features of deviations from Nash equilibria in a wide variety of one-shot game experiments, both static and dynamic, with both complete and incomplete information about others' payoffs (large matrix games, multi-stage bargaining and threat games, and signaling).

V. 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. Parsimonious versions of this model were estimated using data from 37 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, beliefs do not seem to move far enough toward equilibrium levels, and 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. The most notable failures of all of these theories are in games where one decision is safe and the other can result in very high or very low payoffs, and the inclusion of an estimated risk aversion parameter reduces the mean squared error by more than half.

Appendix: Proof of Convergence

Here we prove that the noisy introspection process converges as the number of iterations tends to infinity. The proof applies to symmetric two-player games in which both players have K strategies (the generalization to other games is straightforward). Let S_K denote the K -dimensional simplex and let p_0 be its centroid, i.e. all elements of the vector p_0 are $1/K$. We wish to study $\Phi = \lim_{n \rightarrow \infty} \Phi_n$, where Φ_n is a composition of $(n+1)$ stochastic best response functions:

$$\Phi_n \equiv \phi_\mu \circ \phi_{t\mu} \circ \phi_{t^2\mu} \cdots \circ \phi_{t^n\mu},$$

and $\mu > 0$, $t > 1$. We consider the following generalization of the logit best response functions:

$$\phi_\mu^i(p_0) = \frac{g(\pi^e(i)/\mu)}{\sum_{k=1}^K g(\pi^e(k)/\mu)}, \quad i = 1, \dots, K, \quad (\text{A1})$$

where ϕ_μ^i is the probability that option i is selected and g is some strictly positive, strictly increasing, differentiable function on \mathbb{R} . (Note that (A1) reduces to the logit rule when $g(x) = \exp(x)$.) Since $\phi_\infty(p) = p_0$ for all $p \in S_K$, the limit function Φ (if it exists) maps the whole simplex into a single point. Convergence is proved by showing that the sequence $\{\Phi_n(p_0)\}$ for $n = 1, 2, \dots$, is a Cauchy sequence. Let the distance between two points p and q on the simplex be defined as: $d(p, q) = \max_{i=1, \dots, K} |p_i - q_i|$. For all $m > n$ we have:

$$d(\Phi_n(p_0), \Phi_m(p_0)) = d(\Phi_n(p_0), \Phi_n(p')),$$

where p' is defined as:

$$p' = \phi_{t^{n+1}\mu}(\phi_{t^{n+2}\mu}(\cdots \phi_{t^m\mu}(p_0))).$$

Convergence is established in two steps. First, we show that the distance between p_0 and p' tends to zero as n tends to infinity, and then we use the Mean-Value Theorem to prove that also $d(\Phi_n(p_0), \Phi_n(p')) \rightarrow 0$ as $n \rightarrow \infty$.

Let $p'_{\max} \geq 1/K$ and $p'_{\min} \leq 1/K$ be the largest and smallest elements of p' respectively. The distance between p_0 and p' is the greater of $p'_{\max} - 1/K$ and $1/K - p'_{\min}$. Hence, the distance

is less than or equal to the sum, which equals $p'_{\max} - p'_{\min}$. Let π_{\max} and π_{\min} denote the highest and lowest possible payoffs for the game, then we can bound $p'_{\max} - p'_{\min}$ by:

$$d(p_0, p') \leq \frac{g(\pi_{\max}/t^{n+1}\mu) - g(\pi_{\min}/t^{n+1}\mu)}{K g(\pi_{\min}(k)/t^{n+1}\mu)}.$$

Hence the distance between p_0 and p' tends to 0 as $n \rightarrow \infty$, because the numerator converges to 0 and the denominator converges to $K g(0) > 0$.

The next step is to bound the distance $d(\Phi_n(p_0), \Phi_n(p'))$. Using the Mean-Value Theorem we have:

$$\Phi_n(p') = \Phi_n(p_0) + \nabla \Phi_n(\lambda) \cdot (p' - p_0),$$

where ∇ denotes the gradient and λ is a point on the line connecting p_0 and p' . Hence, $d(\Phi_n(p_0), \Phi_n(p'))$ is at most $K d(p_0, p')$ times the largest element of the gradient matrix $\nabla \Phi_n(\lambda)$. Since

$$\nabla \phi_{t^n \mu} = \frac{1}{t^n \mu} \nabla \phi_1,$$

the chain rule implies the following upper bound for the elements of $\nabla \Phi_n(\lambda)$:

$$|\nabla \Phi_n(\lambda)|_{ij} \leq \frac{1}{\mu^{n+1} t^{\frac{1}{2}n(n+1)}} (\max |\nabla \phi_1|_{ij})^{n+1}, \quad (\text{A2})$$

where the $t^{n(n+1)/2}$ in the denominator results from the product $t^0 * t^1 * \dots * t^n$. Since g is differentiable, elements of $\nabla \phi_1$ will be bounded when payoffs are, so the right side of (A2) tends to zero as n goes to infinity.

To summarize: $d(\Phi_n(p_0), \Phi_m(p_0)) = d(\Phi_n(p_0), \Phi_n(p')) \leq K d(p_0, p') \max |\nabla \Phi_n(\lambda)|$, and since both $d(p_0, p')$ and $\nabla \Phi_n(\lambda)$ tend to zero, the distance between $\Phi_n(p_0)$ and $\Phi_m(p_0)$ becomes arbitrarily small in the limit as n tends to infinity. Q.E.D.

References

- Bernoulli, Daniel (1738) "Specimen Theoriae Novae de Mensura Sortis," *Comentarii Academiae Scientiarum Imperialis Petropolitanae*, 5, 175-192, translated by L. Sommer in *Econometrica*, 1954, 22, 23-36.
- Capra, C. Monica (1998) "Noisy Expectation Formation in One-Shot Games: An Application to the Entry Game," working paper, University of Virginia.
- Goeree, Jacob K. and Charles A. Holt (1999) "Stochastic Game Theory: For Playing Games, Not Just For Doing Theory," *Proceedings of the National Academy of Sciences*, 96, September, 10564-10567.
- Goeree, Jacob K. and Charles A. Holt (2000) "Ten Little Treasures of Game Theory and Ten Intuitive Contradictions," Discussion Paper, University of Virginia.
- Goeree, Jacob K., Charles A. Holt, and Thomas R. Palfrey (1999) "Quantal Response Equilibrium and Overbidding in Private-Value Auctions," Discussion Paper, University of Virginia.
- Guyer, Melvin J. and Anatol Rapoport (1972) " 2×2 Games Played Once," *Journal of Conflict Resolution*, 16(3), 409-431.
- Harsanyi, John (1967-1968) "Games with Incomplete Information Played by Bayesian Players," *Management Science*, 14, 159-182, 320-334, 486-502.
- Harsanyi, John C. and Reinhard Selten (1988) *A General Theory of Equilibrium Selection in Games*, Cambridge, Mass.: MIT Press.
- Luce, D. (1959). *Individual Choice Behavior*, New York: Wiley.
- McKelvey, Richard D. and Thomas R. Palfrey (1995) "Quantal Response Equilibria for Normal Form Games," *Games and Economic Behavior*, 10, 6-38.
- Nasar, Sylvia (1998) *A Beautiful Mind*, New York: Simon and Schuster.
- Nash, John (1950) "Equilibrium Points in N -Person Games," *Proceedings of the National Academy of Sciences*, 36, 48-49.
- Olcina, Gonzalo, and Amparo Urbano (1994) "Introspection and Equilibrium Selection in 2×2 Matrix Games," *International Journal of Game Theory*, 23, 183-206.
- Stahl, Dale and P. Wilson (1995) "On Players' Models of Other Players: Theory and Experimental Evidence," *Games and Economic Behavior*, 10, 218-254.
- Von Neumann, John, and Oscar von Morgenstern (1944) *Theory of Games and Economic Behavior*, Princeton: Princeton University Press.