Rock-Scissors-Paper and Evolutionarily Stable Strategies

Simon Loertscher*

November 12, 2012

Abstract

This paper argues that Rock-Scissors-Paper is a stochastic game with discounting. Provided the discount factor is less than 1, it has an evolutionarily stable strategy (ESS). This result contrasts with the one-shot normal form game, which is the customary representation of Rock-Scissors-Paper. It reconciles the finding that mutant players who tie against each other forever are never observed in real-world play of Rock-Scissors-Paper with a basic prediction of evolutionary game theory.

Keywords: Rock-Scissors-Paper, ESS.

JEL-Classification: C72, D72.

*Department of Economics, Arts West, University of Melbourne, Victoria 3010, Australia. Email: simonl@unimelb.edu.au. I am grateful to an anonymous referee for comments and suggestions that helped to clarify the exposition and the argument.

1 Introduction

Rock-Scissors-Paper (RSP) is a two-player game played by, amongst others, children and agents David and DiNozzo on the TV show NCIS. The purpose of playing RSP is often, though not exclusively, to break some tie. The actions available to every player are *Rock* (R), *Scissors* (S) and *Paper* (P). In every period, the two players choose an action simultaneously. If both choose the same action, the game moves to the next period. The game ends when the actions differ, in which case the winner is determined by the rule that rock beats scissors, scissors beats paper and paper beats rock. As defined by Shapely (1953) and the subsequent literature, this is a *stochastic game* with discounting in case both players prefer the game to end earlier than later, keeping fixed its outcome. Discounting is a sensible assumption when the game is played to break a tie.

A typical short-cut is to represent RSP as a normal form game like the one displayed in Table 1. Here the payoff of losing is set to 0, the payoff of winning is normalized to 2 and the payoff of a tie is equal to the average of the payoff of winning and the payoff of losing. As is well known, the game displayed in Table 1 has a unique Nash equilibrium. In this equilibrium, every player plays the strategy μ^* according to which each action is played with uniform probability. It is also well known that the game displayed in Table 1 fails to exhibit an *evolutionarily stable strategy (ESS)* (see Weibull 1995). That is, a population consisting of players who all play the Nash equilibrium strategy can be invaded by mutants who, say, always play R. Since these mutants fare as well against themselves as μ^* does against itself, there is no (or insufficient) evolutionary pressure against this kind of mutants. That RSP should have no ESS is disquieting as players of little sophistication like children (or, some might argue, agent DiNozzo) seem to be fairly good at playing in a way that, at least to the casual observer, is consistent with the mixed strategy equilibrium. In particular, pairs of children playing RSP each of them sticking stubbornly to the same fixed pure strategy are never observed.

This paper models RSP as a stochastic game with discounting, abbreviated RSP-D, and shows that this game has an ESS for any discount factor less than 1. Section 2 contains the setup while the analysis and the main result are in Section 3. Section 4 provides intuition.

2 Setup

Let $\delta \in (0,1)$ be the discount factor that, for simplicity, is assumed to be common across players, and denote by $x_t^i \in \{R,S,P\}$ the action player i chooses in period t, which includes the possibility that x_t^i is the realization of a random variable such as a mixed strategy. Denote by $\mathcal{H}^\tau$ the relevant history of the stochastic game up to period τ, relevant in the sense that the game is still going on. For RSP-D, $\mathcal{H}^\tau$ is a collection of action-pairs $(\{(x_t^1, x_t^2)\}_{t=0}^{\tau}$ with the property that $x_t^1 = x_t^2$ for all $t \leq \tau$. A strategy $\boldsymbol{\mu} = \{\mu(t)\}_{t=0}^{\infty}$ consists of a $\mu(0) \in \Delta$ and $\mu(t) : \mathcal{H}^{t-1} \to \Delta$ for $t > 0$, where Δ is the set of all probability distributions over the set of actions (or pure strategies in the stage game), which is $\{R,S,P\}$ for RSP. Let $\mathcal{M}$ denote the set of all strategies. The Nash equilibrium strategy of the stage game is denoted $\mu^* = (1/3, 1/3, 1/3)$. The stationary strategy according to which the Nash equilibrium strategy of the stage game is played in every period after every history is denoted $\boldsymbol{\mu}^* = \{\mu(t)^*\}_{t=0}^{\infty}$ with $\mu^*(t) = \mu^*$ for all periods t. The stochastic game RSP-D has a subgame perfect Nash equilibrium in which both players play $\boldsymbol{\mu}^*$.[1] The stage game of RSP-D for period t is displayed

[1] Whether this is the unique equilibrium is an interesting question. The results reported by Walker, Wooders, and Amir (2011) suggest that it may be unique. However, their analysis does not allow for discounting.

Player 1 \ Player 2	R	S	P
R	$1,1$	$2,0$	$0,2$
S	$0,2$	$1,1$	$2,0$
P	$2,0$	$0,2$	$1,1$

Table 1: RSP as a Normal Form Game.

Player 1 \ Player 2	R	S	P
R	$\delta V^{t+1}, \delta V^{t+1}$	$2,0$	$0,2$
S	$0,2$	$\delta V^{t+1}, \delta V^{t+1}$	$2,0$
P	$2,0$	$0,2$	$\delta V^{t+1}, \delta V^{t+1}$

Table 2: The Stage Game of RSP-D in Period t.

in Table 2 under the assumption that both players' expected continuation values V^{t+1} are the same.

Denote by $U(\boldsymbol{\mu}, \hat{\boldsymbol{\mu}})$ the expected discounted payoff of playing $\boldsymbol{\mu}$ against $\hat{\boldsymbol{\mu}}$. Adopting the notion of evolutionary stability from one-shot games (see, for example, Weibull 1995) to stochastic games with discounting, a strategy $\boldsymbol{\mu}$ is naturally said to be an *evolutionarily stable strategy (ESS)* in a stochastic game with discounting if and only if it satisfies the following first-order and second-order best-reply conditions:

$$U(\boldsymbol{\mu}, \boldsymbol{\mu}) \geq U(\hat{\boldsymbol{\mu}}, \boldsymbol{\mu}) \quad \text{for all} \quad \hat{\boldsymbol{\mu}} \in \mathcal{M} \tag{1}$$

$$U(\boldsymbol{\mu}, \boldsymbol{\mu}) = U(\hat{\boldsymbol{\mu}}, \boldsymbol{\mu}) \Rightarrow U(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\mu}}) < U(\boldsymbol{\mu}, \hat{\boldsymbol{\mu}}) \quad \text{for all} \quad \hat{\boldsymbol{\mu}} \neq \boldsymbol{\mu}. \tag{2}$$

3 Analysis and Results

We can now state the main result.

Proposition 1 *For any $\delta < 1$, the stationary strategy $\boldsymbol{\mu}^*$ is an ESS of the Rock-Scissors-Paper game with discounting.*

Proof: As $\boldsymbol{\mu}^*$ constitutes a subgame perfect Nash equilibrium, condition (1) is satisfied. However, because $\boldsymbol{\mu}^*$ induces the play of μ^* in every period, we have $U(\boldsymbol{\mu}^*, \boldsymbol{\mu}^*) = U(\hat{\boldsymbol{\mu}}, \boldsymbol{\mu}^*)$ for every $\hat{\boldsymbol{\mu}}$. Therefore, we have to show that $U(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\mu}}) < U(\boldsymbol{\mu}^*, \hat{\boldsymbol{\mu}})$ for all $\hat{\boldsymbol{\mu}} \neq \boldsymbol{\mu}^*$ holds.

Let

$$A^t = \begin{pmatrix} \delta V^{t+1} & 2 & 0 \\ 0 & \delta V^{t+1} & 2 \\ 2 & 0 & \delta V^{t+1} \end{pmatrix}$$

be the payoff matrix for the stage game in period t for a player who expects a continuation payoff of V^{t+1} if the game proceeds to the next period. We first show that μ^* is the unique maximizer of $\mu A^t \mu^T$, that is $\mu^* = \arg\max_{\mu \in \Delta} \mu A^t \mu^T$ for arbitrary V^{t+1} satisfying $\delta V^{t+1} < 1$. Second, we show that for any symmetric strategy profile $(\boldsymbol{\mu}, \boldsymbol{\mu})$ and any $\delta < 1$ it is indeed the case that $\delta V^{t+1} < 1$ holds for all t. Third, we use these two observations together with the stationarity of the problem to conclude that $V^t = V^{t+1} = V^* = 2/(3-\delta)$ under $\boldsymbol{\mu}^*$ and that, therefore, $U(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\mu}}) < U(\boldsymbol{\mu}^*, \hat{\boldsymbol{\mu}}) = U(\boldsymbol{\mu}^*, \boldsymbol{\mu}^*)$ for all $\hat{\boldsymbol{\mu}} \neq \boldsymbol{\mu}^*$ holds.

1. The first-order conditions for a maximum of $\mu A^t \mu^T$ are $0 = 2\mu_i + \mu_j - 1$ for $i \neq j$ and $i, j = 1, 2$, yielding μ^* as unique solution. The Hessian is negative semidefinite if $\delta V^{t+1} < 1$, 0 if $\delta V^{t+1} = 1$ and positive semidefinite if $\delta V^{t+1} > 1$. Thus, provided $\delta V^{t+1} < 1$, μ^* is indeed the unique maximizer of the stage game payoffs.

2. Assume to the contrary that $\delta V^{t+1} \geq 1$. Then the stage game payoffs are maximized with the strategy $(1, 0, 0)$ and $V^t = \delta V^{t+1}$ follows. But for $\delta < 1$, $V^t = \delta V^{t+1}$ can only hold for $V^t = V^{t+1} = 0$, contradicting the assumption $\delta V^{t+1} \geq 1$ we started with. Hence, we conclude that $\delta V^{t+1} < 1$.

3. The continuation value of the game when $\boldsymbol{\mu}^*$ plays against itself is $V^* = \mu^* A^t \mu^{T*}$ with $V^* = V^t = V^{t+1}$ because of the stationarity of the problem. Solving for V^* yields $V^* = 2/(3 - \delta)$. Because μ^* is the unique maximizer of the stage game payoff $\mu A^t \mu^T$, it follows that V^* is the value of the game, that is, V^* is larger than the continuation value of the game induced by any other strategy $\hat{\boldsymbol{\mu}} \neq \boldsymbol{\mu}^*$ (both the stage game payoffs and the continuation value being strictly larger under $\boldsymbol{\mu}^*$), whence we conclude that $U(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\mu}}) < U(\boldsymbol{\mu}^*, \boldsymbol{\mu}^*)$ for all $\hat{\boldsymbol{\mu}} \neq \boldsymbol{\mu}^*$. The last thing to show is that $U(\boldsymbol{\mu}^*, \hat{\boldsymbol{\mu}}) = U(\boldsymbol{\mu}^*, \boldsymbol{\mu}^*)$ for all $\hat{\boldsymbol{\mu}}$. But this is true because $\mu^* A^t \mu^T = \mu A^t \mu^{*T} = (2 + \delta V^{t+1})/3$ for all $\mu \in \Delta$. Therefore, in every period t the stage game payoffs from $\hat{\boldsymbol{\mu}}$ against $\boldsymbol{\mu}^*$ will be the same as the stage game payoffs of $\boldsymbol{\mu}^*$ against itself, given the same continuation value, and therefore the continuation values will be the same. Thus, $U(\boldsymbol{\mu}^*, \hat{\boldsymbol{\mu}}) = U(\boldsymbol{\mu}^*, \boldsymbol{\mu}^*)$ holds for all $\hat{\boldsymbol{\mu}}$. ■

Observe also that for any $\delta < 1$, μ^* is an ESS of the stage game of RSP-D that is induced by $\boldsymbol{\mu}^*$.

4 Discussion

The intuition for why playing the stage game equilibrium strategy no matter what is an evolutionarily stable strategy in the rock-scissors-paper game with discounting, provided the discount factor is less than one, is that ties bring the play back to square one. With discounting this is bad because it delays the play of the game by one period. The probability that ties occur is uniquely minimized with the uniform probability distribution induced by the Nash equilibrium strategy of the stage game. Therefore, any other strategy does strictly worse against itself than against the stationary strategy of uniform randomization in every stage.

References

Shapley, L. (1953): Stochastic Games, Proceedings of the National Academy of Science, 38, 1095 - 1100.

Walker, M., J. Wooders, and R. Amir (2011): Equilibrium Play in Matches: Binary Markov Games, Games and Economic Behavior, 71, 487 - 502.

Weibull, J. (1995): Evolutionary Game Theory. MIT Press.