

Risk Preference and Sequential Choice in Evolutionary Games

Patrick Roos

Department of Computer Science
Institute for Advanced Computer Studies
University of Maryland, College Park MD 20740, USA
`roos@cs.umd.edu`

Abstract. There is much empirical evidence that human decision-making under risk does not coincide with expected value maximization, and much effort has been invested into the development of descriptive theories of human decision-making involving risk (e.g. Prospect Theory). An open question is how behavior corresponding to these descriptive models could have been learned or arisen evolutionarily, as the described behavior differs from expected value maximization. We believe that the answer to this question lies, at least in part, in the interplay between risk-taking, sequentiality of choice, and population dynamics in evolutionary environments. In this paper we provide the results of several evolutionary game simulations designed to study the risk behavior of agents in evolutionary environments. These include several evolutionary lottery games where sequential decisions are made between risky and safe choices, and an evolutionary version of the well-known stag hunt game. Our results show how agents that are sometimes risk-prone and sometimes risk-averse can outperform agents that make decisions solely based on the maximization of the local expected values of the outcomes, and how this can facilitate the evolution of cooperation in situations where cooperation entails risk.

1 Introduction

Empirical evidence of human decision making under risk shows that humans are sometimes risk averse, sometimes risk seeking, and even behave in ways that systematically violate the axioms of expected utility [17]. Such risk propensities can differ greatly from simple expected value considerations on prospective outcomes. Researchers have invested much effort into constructing utility functions that appropriately model human decision making under risk (e.g. [2, 8, 28]). Researchers have also constructed alternative descriptive theories of decision making that claim to correspond more closely to how humans make decisions involving risk, such as prospect theory [17, 34], regret theory [18], and SP/A (Security-Potential/Aspiration) theory [19–21]. One advantage of these models is that they more explicitly or perhaps more naturally model some of the mechanics involved in human decision making processes. For example, state-dependent

attitudes toward risk are modeled in prospect theory by using a reference point with respect to which prospective outcomes can be interpreted as potential gains or losses, and are modeled in SP/A theory by including an aspiration level as an additional decision criterion in decisions involving risk.

An important open question is how behaviors corresponding to the above decision-making models, or any other empirically documented risk-related behavior that differs from expected value maximization, could have arisen in human evolution or are learned in societies. We believe that the answer lies, at least in part, in the interplay between risk-taking, sequentiality of choice, and population dynamics in evolutionary environments. We present analysis and simulation results to support this hypothesis.

Several recent works speculate about the relation of risk-related behavior and biological evolutionary factors [1, 30, 32]. Our work differs from and expands such study by providing explicit analyses and simulations of risk behavior using evolutionary-game models intended to reflect both biological and cultural evolution. In the spirit of numerous previous studies that have used evolutionary game simulations to explore and derive explanations for how the phenomenon of cooperation can arise in populations of individuals (e.g. [3, 4, 11, 22–24]), we use an evolutionary games based approach to explore risk-related behavior. Our results include the following:

1. Simulations of versions of simple lottery games in which agents make sequential choices among lotteries that have equal expected value but different risks. The experimental results demonstrate that depending on the game’s reproduction mechanism, an agent that acts solely according to the local expected values of outcomes can be outperformed by an agent that varies its risk preference in ways suggested by descriptive models of human decision making.
2. Simulations of evolutionary stag hunt games that show how the principle results from our evolutionary lottery games can apply to and impact results in other, more complex games of social cooperation. The experimental results show how the advantage of conditionally risky behavior can promote the evolution of cooperation in a situation where the cooperation requires a risky decision (namely, choosing to cooperate).

Section 2 introduces our evolutionary lottery games. Analysis, simulation results, and discussion of results for these games follow in Sections 3-5. Section 6 describes our experiments with an evolutionary version of the well-known stag hunt game. Section 7 provides concluding remarks.

2 Evolutionary Lottery Games

Here we describe a class of evolutionary games that we use to investigate risk behavior. Agents in these evolutionary games acquire payoffs dispensed by lotteries. In each generation, each agent must make a sequence of n choices, where each choice is between two lotteries with equal expected value but different risks.

One lottery has a guaranteed payoff of 4, we call this the *safe* lottery. The other lottery gives a payoff of 0 with probability 0.5 or a payoff of 8 with probability 0.5, we call this the *risky* lottery. Both lotteries have an expected value of 4, the only difference is the payoff distribution.

Within this class, we can define different games by varying two important game features, both of which are discussed below: the number n of choices in the sequence, and the population dynamics. We will include all possible pure strategies in the environment in equal frequencies, as described in the following section. Through this setup, we don't model risk propensities of agents explicitly, but each of the agent types strategies can be interpreted as different attitudes towards risk: some risk-averse, some risk-seeking, and some varying depending on previous choice outcomes.

2.1 Number of Choices and Strategies

We consider two cases: $n = 1$, i.e., at each generation the agents make a single, one-shot choice among the two lotteries; and $n = 2$, i.e., at each generation the agents make two sequential choices. When $n = 1$ there are two possible pure strategies, shown in Table 1. When $n = 2$, there are six possible pure strategies, shown in Table 2.

Table 1. All of the possible pure strategies when $n = 1$.

Strategy	Choice
S	choose the safe lottery
R	choose the risky lottery

Table 2. All of the possible pure strategies when $n = 2$.

Strategy	1st lottery	2nd lottery
SS	choose safe	choose safe
RR	choose risky	choose risky
SR	choose safe	choose risky
RS	choose risky	choose safe
RwS	choose risky	choose safe if 1st lottery was won, risky otherwise
RwR	choose risky	choose risky if 1st lottery was won, safe otherwise

2.2 Population Dynamics

Our evolutionary model uses finite, non-overlapping populations of agents. Once all lottery choices have been made and payoffs have been dispensed, all agents reproduce into the next generation (a new population). Reproduction does not

necessarily mean biological reproduction, but can be treated as a model for the process of learning [13] or the social spread and adoption of cultural memes or behavioral traits [6] (e.g. [10]), possibly through imitation. These processes are integral parts of cultural evolution[5]. We consider two different variants of our games, using two widely used population dynamics according to which agents reproduce: the replicator dynamic and imitate-the-better.

The *replicator dynamic* is a commonly used population dynamic originating from biology. Under this dynamic the payoffs received by agents are considered to be a measure of the agent’s fitness, and agent types reproduce proportionately to these payoffs [9, 14]:

$$p_{new} = p_{curr} f_{(a)} / F \tag{1}$$

where p_{new} is the corresponding proportion in the next generation, p_{curr} is the proportion of agents of type a in the current population, $f_{(a)}$ is the average payoff an agent of type a received from all games played in the current generation, and F is the average payoff received by all agents in the population. An agent’s type is simply the strategy it employs to make choices among lotteries. In terms of cultural evolution through imitation or social learning, this model is equivalent to each agent randomly observing another agent in the population for potential imitation (adopting their strategy) and doing so with a probability proportional to the difference in the agent’s own payoff and the observed agent if the observed agent has a higher payoff [9].

Imitate-the-better ([14], used in, e.g.: [11, 12, 29]), also called “tournament selection”, is a reproduction mechanism most commonly used as an imitation dynamic for modeling reproduction of strategies in the context of games played in societies and social learning. Here, each agent in the population is matched up with a randomly drawn other agent in the population and the agent with the higher acquired payoff is reproduced into the next generation. If the payoffs are equal, one of the agents is chosen at random. In terms of imitation, this model is equivalent to each agent randomly observing another agent in the population for potential imitation and doing so if the observed agent has a higher payoff [9]. There is empirical evidence that this imitation strategy is a good model of human social learning [16, 15, 25, 26].

3 Analysis

We now consider four different versions of our evolutionary lottery game, using all four combinations of the following parameters: the number of sequential choices ($n = 1$ or $n = 2$), and the reproduction mechanism (imitate-the-better or replicator dynamics).

We are interested in the dynamics of agent type (i.e. strategy) frequencies as our population of agents evolves over time. To anticipate the performance of different agents in our simulations, it is important to take a closer look at the nature of the reproduction mechanisms. Under replicator dynamics, agents’ rate of change of population frequency is directly proportional to the payoff received relative to the average population payoff. Imitate-the-better on the other

hand acts like a probabilistic, population dependent, threshold step-function, where the particular threshold an agent needs to achieve in order to produce one offspring is the payoff of the randomly drawn opponent from the population.

Recall that for $n = 1$ (i.e., the single choice game) there are only two pure strategies, S and R . S will always receive a payoff of 4, while R will have a 50% chance to receive a payoff of 8 and a 50% chance to receive 0. Hence in each case, the expected value is 4. Thus under replicator dynamics, by equation (1) we expect neither type of agent to have an advantage. Under imitate-the-better, an R agent will have a 50% chance to beat an S agent and a 50% to lose, thus we expect neither agent to have an advantage here either.

For $n = 2$ (i.e., a sequence of two lottery choices), there are six pure strategies. As before, they all have an expected value of 4 at each lottery choice, thus a total expected value of 8 for the sequence of two choices. For each of these strategies except RwS and RwR , the probability of being above the expected value equals the probability of being below the expected value; but for RwS and RwR , the two probabilities differ. (Table 4 in Section 5.1 shows the payoff distributions for all strategies when $n = 2$.) RwS has a 50% chance of receiving a payoff of 12, since it always chooses the safe option after the first lottery was one. It has a 25% chance of receiving the exact expected value 8, which occurs when only the second risky choice is won. Hence it has only a 25% chance of acquiring a payoff below the expected value. (Similarly, RwR has only a 25% chance of acquiring a payoff above the expected value.) This lead us to hypothesize that there are circumstances in which RwS will do better than the other strategies. The purpose of our simulation, described in the next section, is to test this hypothesis.

4 Evolutionary Lottery Game Simulation Results

We have run simulations of our evolutionary lottery game using all four combinations of the following parameters: the number of sequential choices ($n = 1$ or $n = 2$), and the reproduction mechanism (imitate-the-better or replicator dynamics). The types of agents were the ones described in Section 2.1. All simulations started with an initial population of 1000 agents for each agent type and were run for 100 generations, which was sufficient for us to observe the essential population dynamics. The average over multiple simulation runs would not produce significantly different results, except for smoother lines due to noise being averaged out.

In these sets of simulation experiments, the agent types and numbers thereof included in the initial population are the control variables. The independent variables are the number of sequential choices ($n = 1$ or $n = 2$), and the population dynamics used (imitate-the-better and replicator dynamics). The dependent variables measured are the population frequencies of agent type throughout evolutions and the stable population states arrived at, if any.

Figures 1(a,b) show the frequency for each type of agent when $n = 1$. As we had expected, both S and R performed equally well (modulo some stochastic noise) regardless of which reproduction mechanism we used.

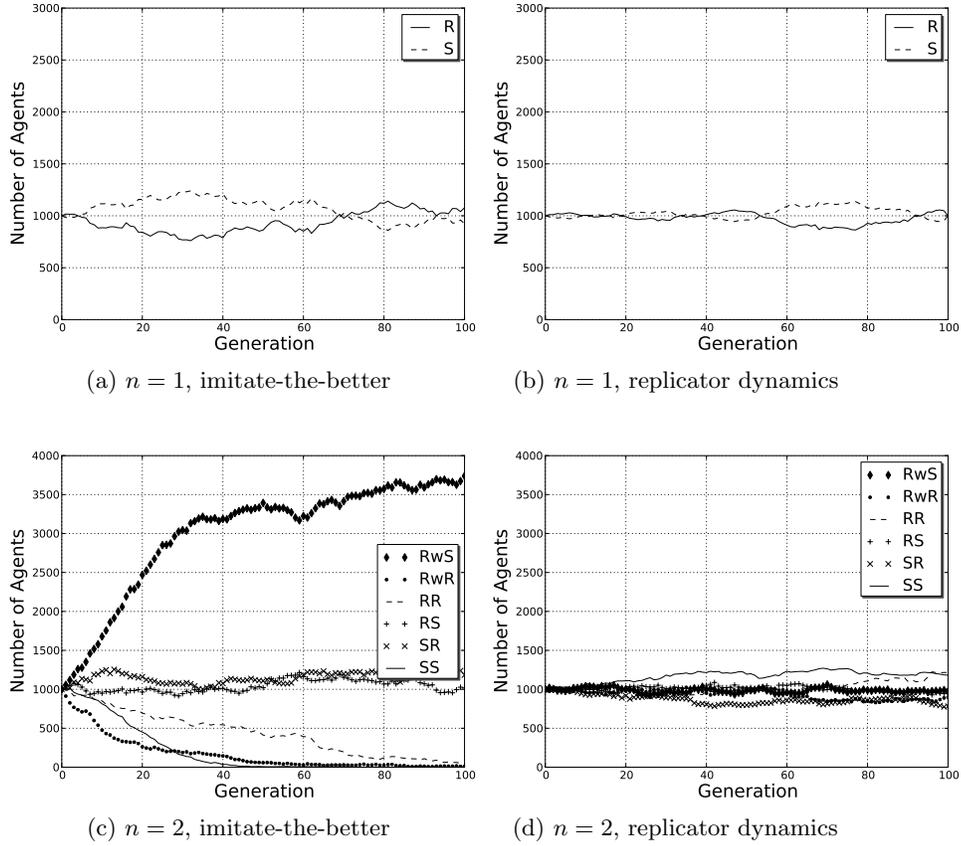


Fig. 1. Agent type frequencies for all four simulations over 100 generations.

For $n = 2$ (Figures 1(c,d)), the results are more interesting and differ depending on the reproduction mechanism used. Under replicator dynamics, all of the strategies performed equally well and remained at their frequency in the original population. But under imitate-the-better, the conditional strategy RwS outperformed the other strategies. RwS rose in frequency relatively quickly to comprise the majority ($> 2/3$) of the population and remained at this high frequency throughout subsequent generations. One surprise, which we discuss in the next section, was that the two unconditional strategies SR and RS fell slightly in population but then remained, comprising the proportion of the population not taken over by RwS .

5 Discussion

Table 3 summarizes the experimental results as they relate to the hypothesis we stated at the end of Section 3. The results confirm our hypothesis that there are circumstances under which the *RwS* agent performs better than other agent types. In particular, it performed much better than all other agents in the sequential lottery using imitate-the-better simulation.

Table 3. Performance of agent types for the single ($n = 1$) and sequential ($n = 2$) lottery game simulations under imitate-the-better and replicator dynamics.

	Imitate-the-better	Replicator
Single	<i>S, R same</i>	<i>S, R same</i>
Sequential	<i>RwS best</i>	all same

The following subsections discuss the impact of the reproduction mechanisms used, the population dynamics observed, and how the results relate to theoretical work concerning models of decision making under risk.

5.1 The Role of Reproduction Mechanisms

The reproduction mechanism played a central role in the results of our simulations. We now analyze the impact of the reproduction mechanisms by examining the expected payoff distributions of the agent types.

Table 4 lists these distributions of payoffs that agents are expected to receive from their choices in one generation of the sequential lottery game.

Table 4. Payoff distributions for all agent types in the sequential lottery game.

agent	<i>RwS</i>	<i>RwR</i>	<i>SR</i>	<i>RS</i>	<i>SS</i>	<i>RR</i>
payoff	12 8 0	16 8 4	12 4	12 4	8	16 8 0
probability	.5 .25 .25	.25 .25 .5	.5 .5	.5 .5	1	.25 .5 .25

We can see that the *RwS* agent had a 50% chance of acquiring a payoff of 12, a 25% chance of acquiring a payoff of 8, and a 25% chance of acquiring 0. Under imitate-the-better, *RwS* had an advantage over the other agents because it had an increased probability of achieving a payoff at or above a certain reproduction threshold. This threshold is the payoff of a randomly drawn opponent, which has an expected value of 8 equal to the expected value of the lotteries. *RwS* pays for this enlarged chance of being above the threshold through a small chance of

doing much worse (payoff 0) than the summed expected utilities, which occurs when the first and the second risky choice is lost.

Since imitate-the-better only considers whether or not the agent's payoff is better than another agent's in order to decide whether the agent reproduces, the *extent* to which the agent is better is not significant. In contrast, the replicator dynamics define reproduction to be directly proportional to the amount by which the agent's payoff deviates from the population average. In this case the small chance of *RwS* of being significantly below the expected value balances against the agent's larger chance of being slightly above it. Thus, under replicator dynamics, the *RwS* agents had no advantage.

Hofbauer and Sigmund [14] define an infinite class of reproduction dynamics called *imitation dynamics*, for which there is a single parameter $0 \leq \alpha \leq 1$ with the following meaning: $\alpha = 0$ is the imitate-the-better dynamic, $\alpha = 1$ is the replicator dynamic, and each value $0 < \alpha < 1$ is a reproduction dynamic whose effects are intermediate between the imitate-the-better and replicator dynamics. We have shown analytically [31] that for every $0 < \alpha < 1$, the *RwS* strategy has an evolutionary advantage over all five of the other pure strategies. In other words, in an evolutionary double lottery dynamic using any imitation dynamic for which $0 < \alpha < 1$, if the initial population includes the *RwS* strategy, then *RwS* will grow to 100% of the population and the other strategies will become extinct.

5.2 Further Population Dynamics

As noted in Section 4, the *SR* and *RS* agents did not go extinct in the sequential-lottery imitate-the-better simulation. The reason *SR* and *RS* remained can be explained by considering all the agents' payoff distributions expected in a generation (Table 4). If we compare the payoff distribution of *SR* and *RS* with that of *RwS*, we see that if these agents are matched up with each other in imitate-the-better there is an equal chance that either of the strategies reproduces, since each strategy has an equal chance of having a higher payoff than the other. Thus once all other strategies are extinct, the population frequencies remain approximately unchanged. The reason *RwS* rises in frequency so much faster early on is because *RwS* has a significantly higher chance of beating an agent from the rest of the population. Against *SS* for example, *RwS* has a 62.5% chance of winning: 50% of the time the payoff of 12 beats the sure payoff of 8 by *SS* and 1/2 of the time the two players are matched with equal payoff of 8 (25% chance), *RwS* is favored. *SR* and *RS* on the other hand only have a 50% chance of winning against *SS*. Similar relations hold for *RR* and *RwR*.

This shows an interesting dynamic of population-dependent success of agents:

- In an environment that contains *SR*, *RS*, and *RwS* and no other strategies, all three do equally well.
- In an environment that contains *SR*, *RS*, *SS* and *RR* and no other strategies, all four do equally well.

- In an environment that contains SR , RS , SS , RR , RwR , and RwS , RwS will increase until SS , RR and RwR become extinct, at which point SR and RS and RwS are at an equilibrium and remain at their current frequencies.

5.3 Relations to Alternative Decision Making Models

The manner in which the RwS strategy deviates from expected value maximization in our lottery game can be characterized as risk-averse (preferring the safe choice) when doing well in terms of payoff and risk-prone (preferring the risky choice) otherwise. Similar risk behavior is suggested by models such as prospect theory [17, 34] and SP/A theory. In prospect theory, people are risk-seeking in the domain of losses and risk-averse in the domain of gains relative to a reference point. In SP/A theory [21], a theory from mathematical-psychology, aspiration levels are included as an additional criterion in the decision process to explain empirically documented deviations in decision-making from expected value maximization.

One explanation for the existence of decision-making behavior as described by such models is that the described behavioral mechanisms are hardwired in decision makers due to past environments in which the behaviors provided an evolutionary advantage [1]. Another interpretation, not necessarily unrelated, is that the utility maximized by decision makers is not the payoffs at hand, but a different perhaps not obvious utility function. Along these lines, [7] proposes a model of decision making that includes probabilities of success and failure relative to an aspiration level into an expected utility representation with a discontinuous (at the aspiration level) utility function. Empirical evidence and analysis provided in [27] provide clear support for the use of probability of success in a model of human decision making. All these descriptive theories provide for agents to be sometimes risk-prone and sometimes risk-averse, depending on their current state or past outcomes, such as the RwS in our simulations.

The sequentiality of choices in our game simulations allows for such state-dependent risk behavior to be explicitly modeled. One could theoretically model the sequential lottery game in normal form, i.e. reduce the choices to a single choice between the payoff distributions listed in Table 4. Doing so would provide essentially equivalent results except that the asymmetry in the payoff distribution of lotteries would be the determining factor of agent successes. In such a representation however, the analysis of risky and safe choices, and agents' preferences among them becomes blurred. In fact, we believe that a tendency towards modeling games in normal form often leads people to overlook the impact of sequentiality on risk-related behavior.

We believe our results show that imitate-the-better models an important mechanism that can lead to the emergence of risk-taking behavior with similar characteristics to that captured in alternative, empirical evidence-based models of decision making like the ones discussed above. Whenever reproductive success is not directly proportional to payoff (i.e., a reproduction mechanism other

than the pure replicator dynamics),¹ risk propensities that differ from expected value maximization have the opportunity to be more successful than agents that solely consider expected value in their local choices. This suggests that there are many other reproduction mechanisms for which expected-value agents can be outperformed by agents that vary their propensities toward risk-taking and risk-averseness.

6 Conditionally Risky Behavior in an Evolutionary Stag Hunt Game

In this section we show how the principle observations from our lottery game experiments apply in a popular social dilemma game of safety and cooperation. We consider an evolutionary game in which agents play two sequential stag hunt games in a generation. Like the prisoner’s dilemma [3], the stag hunt is a game that models a dilemma between cooperation and noncooperation. We demonstrate how a strategy essentially equivalent to the *RwS* strategy from our lottery games can have an advantage in this evolutionary stag hunt environment, and how this advantage impacts the evolutionary results. (For an extensive discussion of the stag hunt game, see e.g. [33].)

6.1 Stag Hunt Environment

The stag hunt environment we consider is equivalent to our sequential lottery game environment, except now payoffs are acquired through two sequential two-player stag hunt games rather than through single-player lotteries. The payoff matrix we use for the stag hunt game is shown in Table 5.² Each generation, all agents are randomly paired to play a two-player stag hunt game. Agents receive payoff from the first game and then are randomly paired again for a second game, the payoffs of both games are accumulated. After these payoffs are accumulated, agents reproduce into the next generation according to population dynamics as before (which means an additional random pairing for imitation under the imitate-the-better dynamic). Since agents play two sequential stag hunt games, we will call this an evolutionary *double stag hunt* game.

6.2 Risk and Strategies

A significant difference between our stag hunt environment and our lottery game environment is that in the former payoffs are not stochastic due to probabilities

¹ We say “pure” here because replicator dynamics can be modified to make reproductive success not directly proportional to payoff. For example, if a death rate (e.g. [24]) is implemented as a payoff-dependent threshold function, we might expect risk propensities to differ depending on whether an agent is above or below that threshold, similar to an aspiration level in SP/A theory.

² Many different payoff matrixes may be used for the stag hunt game, as long as the payoffs satisfy certain constraints. We chose payoff values that coincide with our lottery games, but keep the relevant payoff relations of the stag hunt.

Table 5. Payoff matrix used in our stag hunt game. The payoff values are chosen as to coincide with our lottery games, but keep the relevant payoff relations of the stag hunt.

	Stag	Hare
Stag	8, 8	0, 4
Hare	4, 0	4, 4

on the payoffs themselves, but due to the probabilities of playing against a *stag* agent (i.e., cooperator, always hunts stag) or *hare* agent (i.e., defector, always hunts hare) in the social game. Assume the initial population consists of 50% *stag* agents and 50% *hare* agents. Hence, for a new agent entering the population, hunting stag is a risky choice that will pay 8 with probability of 0.5 and 0 with probability 0.5. Hunting hare on the other hand is a safe choice that will always pay 4. We can thus define the equivalent of an *RwS* strategy in this environment as follows: hunt stag (the risky choice) in the first stag hunt game; if the stag payoff was achieved in the first game, hunt hare (the safe choice) in the second game, otherwise hunt stag again in the second game.

Given what we have learned from our lottery game results in Section 4, we know that in a population approximately split equally between *stag* and *hare* players, the *RwS* strategy just described should have an evolutionary advantage under imitate-the-better (but not under replicator dynamics). This is because with 50% stag and 50% hare agents, the choices that an agent has to make in the two stag hunt games—as described in the previous paragraph—are equivalent in payoff distributions to those of our two-choice lottery game.

We describe simulation experiments that we have run to confirm this hypothesis and to investigate the impact it has on population evolution in Section 6.4. But first we provide some general analysis of the double stag hunt game environment required to explain our experiments and results.

6.3 Analysis

Consider a population consisting of *hare* and *stag* agents. Let s be the proportion of *stag* agents in the population. The payoff to a *hare* agent will be 4 in each stag hunt game, thus *hare* agents will accumulate a payoff of 8 in a generation of the double stag hunt environment. The payoff to a *stag* agent will depend on s , the probability of playing another *stag* player in each game. An accumulated payoff of 16 is only achieved if the agent plays another *stag* agent (getting a payoff of 8) in both games, which occurs with probability s^2 . If the agent plays a *hare* agent (getting a payoff of 0) in both games, it receives a total payoff of 0, which occurs with probability $(1 - s)^2$. Finally, if the agent plays a *hare* agent (payoff of 0) in one game and a *stag* (payoff of 8) agent in the other, it receives a total payoff of 8, which occurs with probability $2s(1 - s)$. Table 6 lists these

payoff distributions achieved by agents in a population consisting of *hare* and *stag* agents in the double stag hunt game environment.

Table 6. Payoff distributions for agents in a population of *stag* and *hare* agents in the double stag hunt environment. s denotes the proportion of *stag* agents in the population.

agent	<i>hare</i>	<i>stag</i>		
payoff	8	16	8	0
probability	1	s^2	$2s(1-s)$	$(1-s)^2$

It can be shown easily that in a population of (50%) *stag* and (50%) *hare* agents, neither strategy will have an advantage (on average) over the other under either replicator dynamics or imitate-the-better. Under replicator dynamics, the average payoff of both strategies is equal, and under imitate-the-better, the probabilities that either strategy will achieve a higher payoff than the other is equal. However, under both population dynamics, if one of the agent types increases in population proportion due to random variation, that agent type will bootstrap itself to take over the entire population.

Under replicator dynamics a random (arbitrarily small) increase in s will lead to a higher average payoff of *stag* agents, which in turn leads to more offspring, which again leads to a higher average payoff. More specifically, let s_i and s_{i+1} be the proportion of *stag* agents in generation i and $i+1$, respectively. Then the replicator equation (Eq. (1)) gives $s_{i+1} = s_i \cdot f_{i(s)}/F_i$, where $f_{i(s)}$ is the average payoff of *stag* agents and F_i is the average payoff of the population. Using the payoff distribution information from Table 6, we get:

$$\begin{aligned}
 f_{i(s)} &= 16s^2 + 16s(1-s) + 0(1-s)^2 = 16s, \\
 F_i &= sf_{i(s)} + 8(1-s) = 16s^2 - 8s + 8, \\
 s_{i+1} &= 16s_i^2 / (16s_i^2 - 8s_i + 8).
 \end{aligned}$$

Since we are dealing with a population of only *stag* and *hare* agents, the proportion of *hare* agents at any generation j is simply $h_j = (1 - s_j)$. Figure 2 plots s_{i+1} and h_{i+1} , the proportion of *stag* and *hare* agents in generation $i+1$ against s_i , the proportion of *stag* agents in the previous generation. We can see that if by random variation we arrive at a generation j in which $s_j \neq 0.5$, if $s_j < 0.5$, then s_j goes to 0, and if $s_j > 0.5$, then s_j goes to 1. Thus eventually one of the agents will bootstrap themselves to take over the entire population. If neither strategy has an advantage when $s = 0.5$, and we have a population split equally between *hare* and *stag* agents, the population converges to 100% *hare* or 100% *stag* agents with equal likelihood under replicator dynamics.

Similarly under imitate-the-better random variation in population proportion will lead to the population being taken over entirely by either *hare* or *stag*

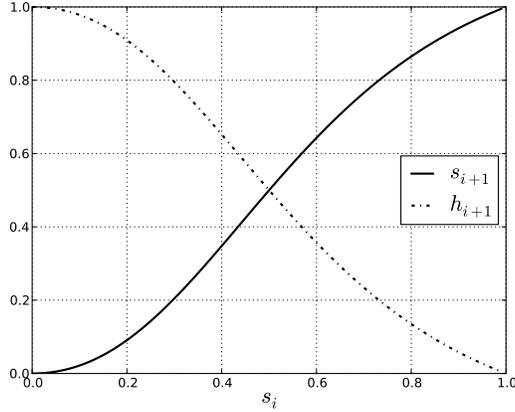


Fig. 2. Plot of s_{i+1} , the proportion of *stag* agents in generation $i + 1$, and h_{i+1} , the proportion of *hare* agents in generation $i + 1$ against s_i , the proportion of *stag* agents in generation i , under replicator dynamics.

agents. When pairing agents for imitate-the-better, we have the following possible pairing probabilities and resulting reproductions:

$$\begin{aligned}
 P(\textit{stag vs. stag}) &= s^2, \text{ whence } \textit{stag} \text{ reproduces.} \\
 P(\textit{hare vs. hare}) &= (1 - s)^2, \text{ whence } \textit{hare} \text{ reproduces.} \\
 P(\textit{stag vs. hare}) &= 2s(1 - s)^2, \text{ whence the agent with the higher payoff} \\
 &\text{reproduces, or a random agent if payoffs are equal.}
 \end{aligned}$$

We can calculate s_{i+1} under imitate-the-better by combining these pairing probabilities and the payoff distribution information from Table 6. Doing so gives:

$$\begin{aligned}
 s_{i+1} &= P(\textit{stag vs. stag}) \cdot 1 + P(\textit{hare vs. hare}) \cdot 0 + \\
 &\quad + P(\textit{stag vs. hare})[P(\textit{stag's payoff is 16}) + P(\textit{stag's payoff is 8})/2] \\
 &= s_i^2 + 2s_i(1 - s_i)[s_i^2 + s_i(1 - s_i)] \\
 &= 3s_i^2 - 2s_i^3.
 \end{aligned}$$

Figure 3 plots s_{i+1} and h_{i+1} against s_i under imitate-the-better. We can see that, as for replicator dynamics, an arbitrarily small increase in s will lead to a higher reproduction probability for *stag*, which will in turn increase s in the next generation. The opposite effect occurs for an arbitrary small decrease in s . As for the replicator dynamics, eventually one of the agents will bootstrap themselves to take over the entire population. If neither strategy has an advantage when $s = 0.5$, and we have a population split equally between *hare* and *stag* agents, the population converges to 100% *hare* or 100% *stag* agents with equal likelihood under imitate-the-better.

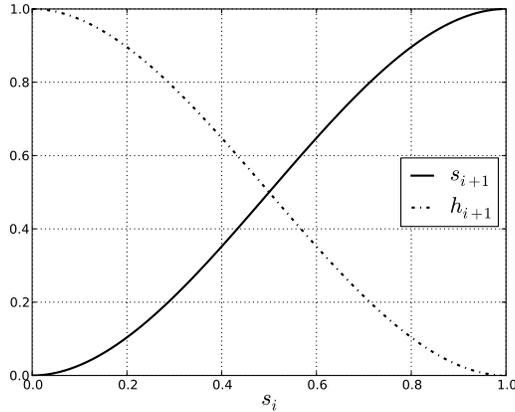


Fig. 3. Plot of s_{i+1} , the proportion of *stag* agents in generation $i + 1$, and h_{i+1} , the proportion of *hare* agents in generation $i + 1$ against s_i , the proportion of *stag* agents in generation i , under imitate-the-better.

Hence we have illustrated how under both replicator dynamics and imitate-the-better, in a population of *hare* and *stag* agents, if one of the agent types acquires a majority in the population (possibly due to random effects), that agent type will bootstrap itself into taking over 100% of the entire population.

6.4 Simulations and Results

Our first set of simulation experiments serves as a control and as to verify that in a population of 50% *stag* and 50% *hare* agents, neither agent type has an advantage on average. Since the above analyzed bootstrapping leads each simulation run to converge to 100% *stag* or 100% *hare* agents, we run a large number of simulation runs and count the amount of times the population is entirely taken over by either agent type. Figure 4 shows the counts of each for 200 simulation runs for an initial population of 3000 *stag* and 3000 *hare* agents under both replicator dynamics and imitate-the-better. We see that the counts are very close, confirming that neither agent type has an advantage under either population dynamic and the population is equally likely to evolve to full cooperation (100% *stag*) and full defection (100% *hare*).

We have hypothesized in Section 6.2 that in a population of 50% *stag* and 50% *hare* players, given the payoff matrix in Table 5, the *RwS* agent in our stag hunt environment should have an evolutionary advantage under imitate-the-better (but not replicator dynamics), as the two choices of hunting hare vs. stag are equivalent to the safe vs. risky lottery choices in our earlier lottery games. Our second set of experiments serves to verify this hypothesis and investigate the impact it has on population evolution. For this set of simulations, we used an initial population of 3000 *stag*, 3000 *hare*, and a small amount (30) of *RwS* agents.

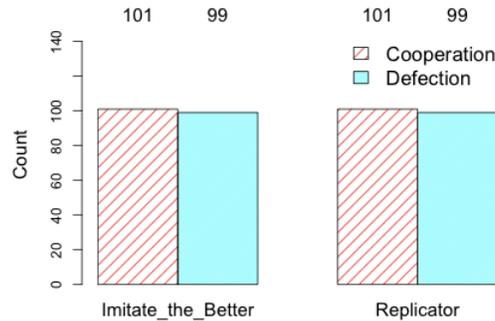


Fig. 4. Simulation results for an initial population of 3000 *hare* and 3000 *stag* agents. The plot shows the count of simulations in which the population resulted in all *stag* agents (cooperation) and all *hare* agents (defection) for 200 simulations under each imitate-the-better and replicator dynamics.

We again ran 200 simulations each under replicator dynamics and imitate-the-better (the independent variable is the population dynamics used) and compare results. The earlier described bootstrapping of *stag* or *hare* agents occurs just the same in a population with *RwS* agents as it does in a population without. Thus all of our simulations again lead to the population evolving to complete cooperation (100% *stag*) or complete defection (100% *hare*).

Figure 5 shows the number of times that the population evolved to complete cooperation and the number of times it evolved to complete defection under replicator dynamics and under imitate-the-better. Observe that under imitate-the-better the population evolves to all cooperators more often than under replicator dynamics. A Pearson’s Chi-squared test shows this difference in the number of cooperative outcomes between the two sets of simulations to be significant with a p-value of 0.005414 ($X^2 = 7.7356$).

6.5 Discussion

The reason a significantly higher amount of cooperation occurred under imitate-the-better is due to the fact that the *RwS* strategy (as expected from our lottery game results) had an advantageous risk behavior under the imitate-the-better dynamics. This led to growth in the number of *RwS* agents during the first few iterations (during which the *stag* and *hare* players occupied an approximately equal population proportion). The *RwS* agents in the population aid the cooperating *stag* players, since the *RwS* agents will play *stag* as long as they haven’t already received a *stag* payoff in an earlier game. Thus *RwS* agents serve as a catalyst to *stag* agents. Since the *RwS* agents initially increase under imitate-the-better, the chance they will boost the *stag* players and lead them to bootstrap themselves into taking over the population is higher under imitate-the-better than under replicator dynamics.

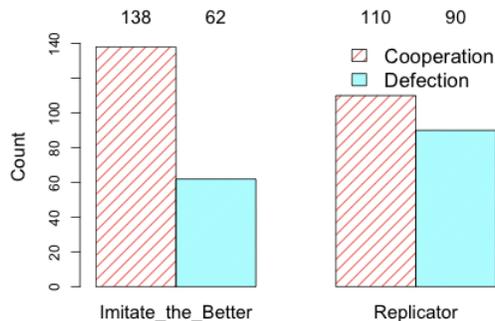


Fig. 5. Simulation results for an initial population of 3000 *hare* and 3000 *stag* agents and 30 *RwS* agents. The plot shows the count of simulations in which the population resulted in all *stag* agents (cooperation) and all *hare* agents (defection) for 200 simulations under each imitate-the-better and replicator dynamics.

Figure 6 shows a plot of the number of agents of each type from a typical simulation run under imitate-the-better in which this “boosting” occurs. We see that the *RwS* agents grew from the initial 30 to over 500 agents, which was enough aid to the cooperating *stag* players for them to take over the population. Once the *stag* agents grew to a significantly higher population proportion, hunting *stag* is no longer as much of a risky choice, and the *RwS* agents begin to decline in numbers. In summary, these experiments showed that the principle lessons learned from our lottery game simulations can apply and impact the results of other (social) evolutionary games, in this case promoting the emergence of cooperative behavior in an evolutionary double *stag* hunt environment.

7 Conclusion

We have explored risk behavior of agents through analysis and simulation of several evolutionary games. We provided results for two types of simulations: 1) simulations of simple evolutionary lottery games that we have proposed to study risk behavior of agents in evolutionary environments, and 2) simulations of evolutionary *stag* hunt games that show how our results from the lottery games can apply to a more complex social cooperation game.³

7.1 Evolutionary Lottery Game Simulations

Our results from several evolutionary lottery games demonstrate how the interplay between sequentiality of choice and population dynamics can affect decision

³ Simulation source code and result data used for this paper are made available for download on the author’s website at <http://www.cs.umd.edu/users/roos/materials/ACS2010.html>.

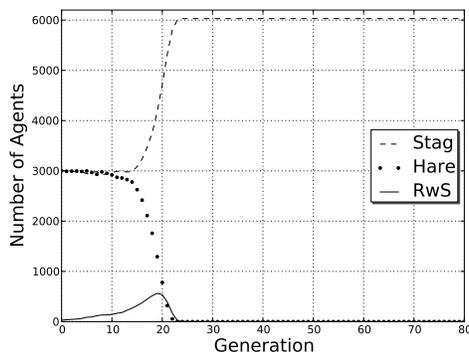


Fig. 6. Agent type frequencies for a typical stag hunt simulation run under imitate-the-better in which *RwS* agents grew and boost stag players, leading them to take over the entire population

making under risk. The simulations show that a strategy other than expected-value maximization can do well in an evolutionary environment having the following characteristics:

- At each generation, the agents must make a sequence of choices among alternatives that have differing amounts of risk.
- An agent’s reproductive success is not directly proportional to the payoffs produced by those choices. We specifically considered imitate-the-better; but as pointed out in Section 5.3, we could have gotten similar results with many other reproduction mechanisms.

The most successful strategy in our simulations, namely the *RwS* strategy, exhibited behavior that was sometimes risk-prone and sometimes risk-averse depending on its success or failure in the previous lottery. Such a behavioral characteristic is provided for in descriptive theories of human decision making based on empirical evidence. It is not far-fetched to suppose that when human subjects have exhibited non-expected-value preferences in empirical studies, they may have been acting as if their decisions were part of a greater game of sequential decisions in which the success of strategies is not directly proportional to the payoff earned. Apart from a purely biological interpretation, in which certain behavioral traits are hardwired in decision-makers due to past environments, perhaps such empirical studies capture the effects of the subjects’ learned habit of making decisions as part of a sequence of events in their daily lives.

Our results also demonstrate (see Section 5.2) that the population makeup can have unexpected effects on the spread and hindrance of certain risk propensities. This may be an important point to consider, for example, when examining decision-making across different cultures, societies, or institutions.

7.2 Evolutionary Stag Hunt Game Simulations

Our evolutionary stag hunt game simulations show how the results from our simple lottery games can apply in other, more complex and commonly studied games of social cooperation. The results show how the advantage of conditionally risky behavior under imitate-the-better can promote the evolution of cooperation in a situation where the cooperation requires a risky decision (namely, choosing to cooperate). We suspect that the interplay between risk taking, sequential choices, and population dynamics can impact a variety of other games (e.g. the Prisoner's Dilemma) similarly.

7.3 Future Work

General avenues for future work include investigating how a greater range of population dynamics and sequential choices can affect risk behavior as well as if and how such results apply to a variety of other games and situations. Our simple lottery game simulations are a first step in exploring evolutionary mechanisms which can induce behavioral traits resembling those described in popular descriptive models of decision making. A specific related topic to explore is how the prospect-theoretic notion of setting a reference point may relate to evolutionary simulations with sequential lottery decisions. In general, there is much more opportunity for future work to use simulation for the purpose of exploring or discovering the mechanisms which induce, possibly in a much more elaborate and precise manner, the risk-related behavior characteristics described by prospect theory or other popular descriptive decision making models based on aspiration levels.

Acknowledgements

This work was supported in part by AFOSR grant FA95500610405 and NAVAIR contract N6133906C0149. The opinions in this paper are those of the authors and do not necessarily reflect the opinions of the funders.

References

1. Alasdair I. Houston, J. M. M. and Steer, M. D., Do we expect natural selection to produce rational behaviour?, *Philosophical Transactions of the Royal Society B* **362** (2007) 1531–1543.
2. Arrow, K. J., *Essays in the theory of risk-bearing* (Markham, Chicago, 1971).
3. Axelrod, R. and Hamilton, W. D., The evolution of cooperation, *Science* **211** (1981) 1390–1396.
4. Bowles, S. and Gintis, H., The evolution of strong reciprocity: cooperation in heterogeneous populations, *Theoretical Population Biology* **65** (2004) 17–28.
5. Boyd, R. and Richerson, P., *Culture and the evolutionary process* (University of Chicago Press, 1988).
6. Dawkins, R., *The Selfish Gene* (New York: Oxford University Press, 1976).

7. Diecidue, E. and Ven, J. V. D., Aspiration level, probability of success and failure, and expected utility, *International Economic Review* **49** (2008) 683–700.
8. Friedman, M. and Savage, L. J., The utility analysis of choices involving risk, *The Journal of Political Economy* **56** (1948) 279–304.
9. Gintis, H., *Game Theory Evolving: A Problem-centered Introduction to Modeling Strategic Behavior* (Princeton University Press, 2000).
10. Hales, D., An open mind is not an empty mind: Experiments in the meta-noosphere, *Journal of Artificial Societies and Social Simulation* **1** (1998).
11. Hales, D., Evolving specialisation, altruism, and group-level optimisation using tags, in *MABS*, eds. Sichman, J. S., Bousquet, F., and Davidsson, P., *Lecture Notes in Computer Science*, Vol. 2581 (Springer, 2002), ISBN 3-540-00607-9, pp. 26–35.
12. Hales, D., Searching for a soulmate - searching for tag-similar partners evolves and supports specialization in groups, in *RASTA*, eds. Lindemann, G., Moldt, D., and Paolucci, M., *Lecture Notes in Computer Science*, Vol. 2934 (Springer, 2002), ISBN 3-540-20923-9, pp. 228–239.
13. Harley, C. B., Learning the evolutionarily stable strategy, *Journal of Theoretical Biology* **89** (1981) 611–633.
14. Hofbauer, J. and Sigmund, K., Evolutionary game dynamics, *Bulletin of the American Mathematical Society* **40** (2003) 479–519.
15. Huck, S., Normann, H., and Oechssler, J., Does information about competitors' actions increase or decrease competition in experimental oligopoly markets?, *International Journal of Industrial Organization* **18** (2000) 39–57.
16. Huck, S., Normann, H. T., and Oechssler, J., Learning in cournot oligopoly: An experiment, *SSRN eLibrary* (1997).
17. Kahneman, D. and Tversky, A., Prospect theory: An analysis of decision under risk, *Econometrica* **47** (1979) 263–291.
18. Loomes, G. and Sugden, R., Regret theory: An alternative theory of rational choice under uncertainty, *The Economic Journal* **92** (1982) 805–824.
19. Lopes, L. L., Between hope and fear: The psychology of risk., *Advances in Experimental Social Psychology* **20** (1987) 255–295.
20. Lopes, L. L., Re-modeling risk aversion., in *Acting under uncertainty: Multidisciplinary conceptions*, ed. von Furstenberg, G. M. (Boston: Kluwer, 1990), pp. 267–299.
21. Lopes, L. L. and Oden, G. C., The role of aspiration level in risky choice: A comparison of cumulative prospect theory and sp/a theory, *Journal of Mathematical Psychology* **43** (1999) 286–313.
22. Nowak, M., Five rules for the evolution of cooperation, *Science* **314** (2006) 1560–1563.
23. Nowak, M. A. and Sigmund, K., Tit for tat in heterogeneous populations, *Nature* **355** 250–253.
24. Nowak, M. A. and Sigmund, K., A strategy of win-stay, lose-shift that outperforms tit for tat in the prisoner's dilemma game, *Nature* **364** (1993) 56–58.
25. Offerman, T., Potters, J., and Sonnemans, J., Imitation and belief learning in an oligopoly experiment, *The Review of Economic Studies* **69** (2002) 973–997.
26. Offerman, T. and Schotter, A., Imitation and luck: An experimental study on social sampling, *Games and Economic Behavior* **65** (2009) 461–502.
27. Payne, J. W., It is whether you win or lose: The importance of the overall probabilities of winning or losing in risky choice, *Journal of Risk and Uncertainty* **30** (2005) 5–19.

28. Rabin, M., Risk aversion and Expected-Utility theory: A calibration theorem, *Econometrica* **68** (2000) 1281–1292.
29. Riolo, R. L., Cohen, M. D., and Axelrod, R., Evolution of cooperation without reciprocity., *Nature* **411** (2001) 441–443.
30. Rode, C. and Wang, X., Risk-Sensitive Decision making Examined within an Evolutionary Framework, *American Behavioral Scientist* **43** (2000) 926–939.
31. Roos, P., Carr, J. R., and Nau, D., Evolution of state-dependent risk preferences, *submitted for journal publication* (2010).
32. Stevens, J., *Rational Decision Making in Primates: The Bounded and the Ecological*, Vol. 2 (Oxford University Press, 2010), pp. 98–111.
33. Syrms, B., *The Stag Hunt and the Evolution of Social Structure* (Cambridge University Press, Cambridge, U.K., 2003).
34. Tversky, A. and Kahneman, D., Advances in prospect theory: Cumulative representation of uncertainty, *Journal of Risk and Uncertainty* **5** (1992) 297–323.