State-Dependent Risk Preferences in Evolutionary Games

Patrick Roos[1,2] and Dana Nau[1,3,2]

[1] Department of Computer Science
[2] Institute for Advanced Computer Studies
[3] Institute for Systems Research
University of Maryland, College Park MD 20742, USA
{roos,nau}@cs.umd.edu

Abstract. There is much empirical evidence that human decision-making under risk does not correspond the decision-theoretic notion of "rational" decision making, namely to make choices that maximize the expected value. An open question is how such behavior could have arisen evolutionarily. We believe that the answer to this question lies, at least in part, in the interplay between risk-taking and sequentiality of choice in evolutionary environments.

We provide analytical and simulation results for evolutionary game environments where sequential decisions are made between risky and safe choices. Our results show there are evolutionary games in which agents with *state-dependent* risk preferences (i.e., agents that are sometimes risk-prone and sometimes risk-averse depending on the outcomes of their previous decisions) can outperform agents that make decisions solely based on the local expected values of the outcomes.

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 [1]. Researchers have invested much effort into constructing utility functions that appropriately model human decision making under risk (e.g. [2–4]). 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 [1, 5], regret theory [6], and SP/A (Security-Potential/Aspiration) theory [7–9].

A question that has received much less attention 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 or been learned in societies. We believe that one part of the answer to this question is the interplay between risk-taking and *sequentiality of choices*; and in this paper we present analytical and simulation results to support this hypothesis.

Our 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 whose risk preference depends on the success or failure of its previous choices.

2 Evolutionary Lottery Games

We now describe a class of evolutionary games based on a finite, homogeneous population model in which agents 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 certain outcome of payoff 4 (with probability 1), we call this the *safe* lottery. The other lottery gives a payoff of 0 with probability 0.5 and 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 reproduction dynamics.

2.1 Number of Choices

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 (i.e., $n = 2$).

When $n = 1$ there are two possible pure strategies, as shown in Table 1 When $n = 2$, there are six possible pure strategies, as 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
R-WS	choose risky	choose safe if 1st lottery was won, risky otherwise
R-WR	choose risky	choose risky if 1st lottery was won, safe otherwise

2.2 Reproduction Dynamics

Our evolutionary model uses 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 also be a treated as a model for the process of learning [10] or the social spread and adoption of cultural memes or behavioral traits [11], e.g. [12]. We consider two different variants of our games, using two widely used reproduction mechanisms: the replicator dynamic and an imitation dynamic.

The *replicator dynamic*, originating from biology, is the most widely used reproduction mechanism in the literature on evolutionary game theory. The payoffs received by agents are considered to be a measure of the agent's fitness, and agent types reproduce proportional to these payoffs [13, 14]:

$$p^{new} = p^{curr} pay(agent_i)/\overline{pay} \tag{1}$$

where p^{curr} is the proportion of agents of type i in the current population, p^{new} is the corresponding proportion in the next generation, $pay(agent_i)$ is the average payoff an agent of type i received from all games played, and $\overline{pay}$ 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.

Imitation dynamics are probably the second most widely used kind of reproduction mechanism, and are arguably more appropriate in modeling reproduction of strategies in the context of games played in societies [13]. We use the imitation process commonly referred to as *tournament selection*[15–17]. 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 of the matched agents is equal, one of the two agents is chosen at random to reproduce.

3 Analytical Results

We now analyze how well the various strategies should perform under all four combinations of the following parameters: the number of sequential choices ($n = 1$ or $n = 2$), and the reproduction mechanism (imitation or replicator dynamic).

3.1 Case $n = 1$

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 the replicator dynamic, by equation (1) we expect neither type of agent to have an advantage. Under the imitation dynamic, 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.

Table 3. Payoff distributions for all agent types in the sequential lottery game.

agent	*R-WS*			*R-WR*			*SR*		*RS*		*SS*	*RR*		
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

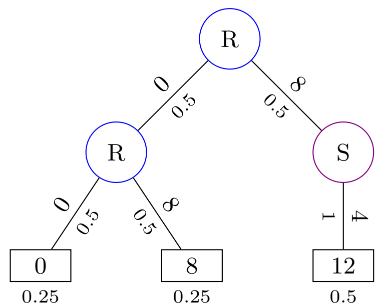


Fig. 1. Sequential lottery tree illustrating payoff distribution achieved by the *R-WS* strategy. Nodes represent lotteries (R for risky, S for safe). Edges are labeled with the payoff dispensed and the associated probability. Nodes are labeled with the final accumulated payoff and the probability for each.

3.2 Case $n = 2$

The situation is more complicated when $n = 2$. Recall from Table 2 that in this case there are six pure strategies. Table 3 gives, for each strategy, its possible numeric payoffs, and the probabilities of these payoffs. We can see by Figure 1 that the *R-WS* agent has 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 the imitation dynamic, *R-WS* has an advantage over the other strategies because it has 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. *R-WS* pays for this enlarged chance of being above the threshold through a small chance of doing much worse (payoff 0) than the summed expected values, which occurs when the first and the second risky choice is lost.

The replicator dynamic defines 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 *R-WS* of being significantly below the expected value balances against the agent's larger chance of being slightly above it. Thus, under the replicator dynamic, the *R-WS* agents have no advantage. All six strategies have an expected value of 4 at each lottery choice, thus a total expected value of 8 for the sequence of two choices. Consequently, we would expect all six strategies to do equally well when using the replicator dynamic.

Since the imitation dynamic 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.

If we compare the payoff distribution of *SR* and *RS* with that of *R-WS*, we see that if agents of these strategies are matched up with each other under the imitation dynamic, there is an equal chance that either of the agent reproduces. But an *R-WS* has a significantly higher chance of beating an agent from the rest of the population. Against *SS* for example, *R-WS* 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), *R-WS* is favored. *SR* and *RS* on the other hand only have a 50% chance of winning against *SS*. Similar relations hold for *RR* and *R-WR*.

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

– In an environment that contains *SR*, *RS*, and *R-WS* and no other strategies, all three should do equally well.
– In an environment that contains *SR*, *RS*, *SS* and *RR* and no other strategies, all four should do equally well.
– In an environment that contains *SR*, *RS*, *SS*, *RR*, and *R-WS*, *R-WS* will increase until *SS* and *RR* become extinct, at which point *SR* and *RS* and *R-WS* are at an equilibrium and remain at their current frequencies.

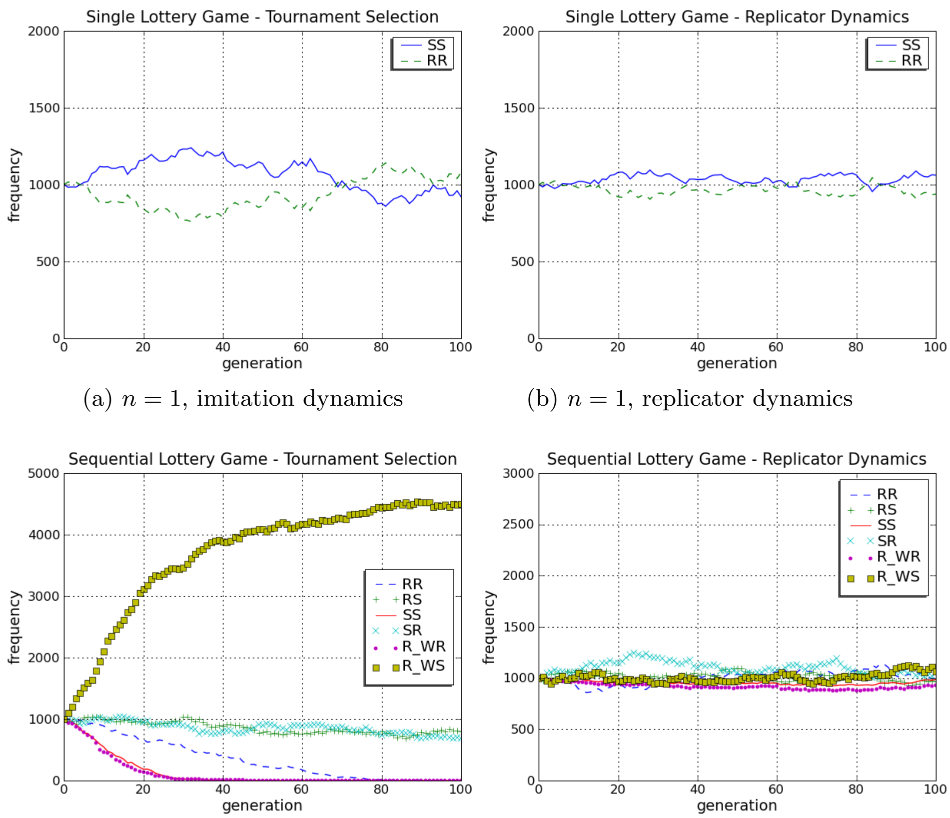
In the following section, we report on simulation results that confirm these predictions.

4 Simulation Results

To test the predictions at the end of the previous section, we have run simulations using all four combinations of the number of sequential choices ($n = 1$ or $n = 2$) and the reproduction mechanism (imitation or replicator dynamic). 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.

Figures 2(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.

For $n = 2$ (Figures 2(c,d)), the results are more interesting and differ depending on the reproduction mechanism used. Under the replicator dynamic, all of the strategies performed equally well and remained at their frequency in the original population. But under the imitation dynamic, the conditional strategy *R-WS* outperformed the other strategies. *R-WS* rose in frequency relatively quickly to comprise the majority ($> 2/3$) of the population and remained at this high frequency throughout subsequent generations. Furthermore, the two unconditional strategies *SR* and *RS* remained, comprising the proportion of the population not taken over by *R-WS*.



(a) $n = 1$, imitation dynamics (b) $n = 1$, replicator dynamics

(c) $n = 2$, imitation dynamics (d) $n = 2$, replicator dynamics

Fig. 2. Agent type frequencies for all four simulations over 100 generations.

5 Relations to Alternative Decision Making Models

The manner in which the *R-WS* 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 [1, 5] 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 [9], 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 [18]. 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, [19] 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 [20] 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 *R-WS* in our simulations.

The sequentiality of choices in our game simulations allow 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 3. 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 imitation dynamics model 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 the reproduction rate is not directly proportional to payoff (i.e., a reproduction mechanism other than the pure replicator dynamic),[4] 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 Conclusion

Our analytical and experimental results in several evolutionary lottery games demonstrate how sequentiality and reproduction can affect decision making under risk. Our results show that a strategy other than expected-value maximization can become prevalent in an evolutionary environment having the following characteristics:

[4] We say "pure" here because the replicator dynamic can be modified to make reproductive success not directly proportional to payoff. For example, if a death rate (e.g. [21]) 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.

- 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 an imitation dynamic known as tournament selection; but as pointed out in Section 5, we could have gotten similar results with many other reproduction mechanisms.

The most successful strategy in our analysis and experiments, namely the *R-WS* strategy, exhibits behavior that is sometimes risk-prone and sometimes risk-averse depending on its success or failure in the previous lottery. This kind of 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 Fig. 2 and the last paragraph of Section 3) 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 or societies.

In conclusion, 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, NAVAIR contract N6133906C0149, DARPA's Transfer Learning Program, DARPA IPTO grant FA8650-06-C-7606, and NSF grant IIS0412812. The opinions in this paper are those of the authors and do not necessarily reflect the opinions of the funders.

References

1. Kahneman, D., Tversky, A.: Prospect theory: An analysis of decision under risk. Econometrica **47** (1979) 263–291
2. Friedman, M., Savage, L.J.: The utility analysis of choices involving risk. The Journal of Political Economy **56**(4) (August 1948) 279–304
3. Arrow, K.J.: Essays in the theory of risk-bearing. Markham, Chicago (1971)
4. Rabin, M.: Risk aversion and Expected-Utility theory: A calibration theorem. Econometrica **68**(5) (September 2000) 1281–1292
5. Tversky, A., Kahneman, D.: Advances in prospect theory: Cumulative representation of uncertainty. Journal of Risk and Uncertainty **5** (1992) 297–323
6. Loomes, G., Sugden, R.: Regret theory: An alternative theory of rational choice under uncertainty. The Economic Journal **92** (December 1982) 805–824
7. Lopes, L.L.: Between hope and fear: The psychology of risk. Advances in Experimental Social Psychology **20** (1987) 255–295
8. Lopes, L.L.: Re-modeling risk aversion. In von Furstenberg, G.M., ed.: Acting under uncertainty: Multidisciplinary conceptions. Boston: Kluwer (1990) 267–299
9. Lopes, L.L., 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
10. Harley, C.B.: Learning the evolutionarily stable strategy. Journal of Theoretical Biology **89** (1981) 611–633
11. Dawkins, R.: The Selfish Gene. New York: Oxford University Press (1976)
12. Hales, D.: An open mind is not an empty mind: Experiments in the metanoosphere. Journal of Artificial Societies and Social Simulation **1**(4) (1998)
13. Hofbauer, J., Sigmund, K.: Evolutionary game dynamics. Bulletin of the American Mathematical Society **40**(4) (July 2003) 479–519
14. Gintis, H.: Game Theory Evolving: A Problem-centered Introduction to Modeling Strategic Behavior. Princeton University Press (2000)
15. Hales, D.: Evolving specialisation, altruism, and group-level optimisation using tags. In Sichman, J.S., Bousquet, F., Davidsson, P., eds.: MABS. Volume 2581 of Lecture Notes in Computer Science., Springer (2002) 26–35
16. Riolo, R.L., Cohen, M.D., Axelrod, R.: Evolution of cooperation without reciprocity. Nature **411** (2001) 441–443
17. Hales, D.: Searching for a soulmate - searching for tag-similar partners evolves and supports specialization in groups. In Lindemann, G., Moldt, D., Paolucci, M., eds.: RASTA. Volume 2934 of Lecture Notes in Computer Science., Springer (2002) 228–239
18. A L Houston, J M McNamara, M.D.S.: Do we expect natural selection to produce rational behaviour? Philosophical Transactions of the Royal Society B **362** (2007) 1531–1543
19. Diecidue, E., Ven, J.V.D.: Aspiration level, probability of success and failure, and expected utility. International Economic Review **49**(2) (May 2008) 683–700
20. 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
21. Nowak, M.A., 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