ABSTRACT

Title of dissertation: EVOLUTIONARY GAME THEORETIC
MODELING OF DECISION MAKING
AND CULTURE

Patrick Roos, Doctor of Philosophy, 2012

Dissertation directed by: Dana Nau
Department of Computer Science

Evolutionary Game Theory (EGT) has become an attractive framework for modeling human behavior because it provides tools to explicitly model the dynamics of behaviors in populations over time and does not require the strong rationality assumptions of classical game theory. Since the application of EGT to human behavior is still relatively new, many questions about human behavior and culture of interest to social scientists have yet to be examined through an EGT perspective to determine whether explanatory and predictive rather than merely descriptive insights can be gained. In this thesis, informed by social science data and under close collaboration with social scientists, I use EGT-based approaches to model and gain a qualitative understanding of various aspects of the evolution of human decision-making and culture. The specific phenomena I explore are i) risk preferences and their implications on the evolution of cooperation and ii) the relationship between societal threat and the propensity with which agents of societies punish norm-violating behavior.

First, inspired by much empirical research that shows human risk-preferences to be state-dependent rather than expected-value-maximizing, I propose a simple sequential lottery game framework to study the evolution of human risk preferences. Using this game model in conjunction with known population dynamics provides the novel insight that for a large range of population dynamics, the interplay between risk-taking and sequentiality of choices allows state-dependent risk behavior to have an evolutionary advantage over expected-value maximization. I then demonstrate how this principle can facilitate the evolution of cooperation in classic game-theoretic games where cooperation entails risk.

Next, inspired by striking differences across cultural groups in their willingness to punish norm violators, I develop evolutionary game models based on the Public Goods Game to study punishment behavior. Operationalizing various forms of societal threat and determining the relationship between these threats and evolved punishment propensities, these models show how cross-cultural differences in punishment behavior are at least partially determined by cultures' exposure to societal threats, providing support for social science theories hypothesizing that higher threat is a causal factor for higher punishment propensities.

This work advances the state of the art of EGT and its applications to the social sciences by i) creating novel EGT models to study different phenomena of interest in human decision-making and culture, and ii) using these models to provide insights about the relationships between variables in these models and their impact on evolutionary outcomes. By developing and analyzing these models under close consideration of relevant social science data, this work not only advances our understanding of how to use evolutionary game and multi-agent system models to study social phenomena, but also lays the foundation for more complex explanatory and predictive tools applicable to behaviors in human populations.

EVOLUTIONARY GAME THEORETIC MODELING OF DECISION MAKING AND CULTURE

by

Patrick Roos

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2012

Advisory Committee:
Dana Nau, Chair/Advisor
Michele Gelfand, Co-Advisor
Mohammad Hajiaghayi
Jim Reggia
V.S. Subrahmanian

Magna est veritas et prevalebit.

Acknowledgments

I am exceptionally thankful for my advisors Dana Nau and Michele Gelfand for their invaluable guidance that cannot be reflected in a few words and for being a great complement to each other as advisors. I'm grateful to Dana for giving me a chance as his graduate research assistant, and for Michele and Dana for making it possible for me to do the interdisciplinary research I had always hoped to be able to do. I am honored to work with you both. As heads of the Lab for Computational Cultural Dynamics, V.S. Subrahmanian and Dana have my gratitude for inspiring me and showing me that applying computer science to social science and culture is a realizable avenue of research. Bill Manaris has my endless gratitude as the person to recognize me and get me started on the path of research in the first place, for his skillful mentoring, and for encouraging me to continue research in graduate school.

I am grateful for my family and friends for their encouragement and support throughout the years of working on this research. Melissa Lee deserves special recognition for her invaluable role in helping me concretize and clarify my ideas and their presentation. I also want to thank my collaborators Ryan Carr and Paulo Shakarian, who enabled me to be more productive and made my graduate school research a pleasant collaborative experience, and my former post-doc Austin Parker who helped motivate me through conversations early on and was always available for encouragement and advice.

Funding Acknowledgments

The work in this thesis was supported in part by AFOSR Grant FA95501210021, ARO Grant W911NF1110344, and ONR Grant W911NF0810144. The views expressed in this thesis are those of the author and do not necessarily reflect the official policy or position of the funders.

Table of Contents

List of Figures

List of Abbreviations

General Terms:

EGT	Evolutionary Game Theory
PGG	Public Goods Game
ESS	Evolutionary Stable Strategy

2-Lottery Game Strategies:

RwS	Risky-Win-Safe Strategy
RwR	Risky-Win-Risky Strategy
RR	Risky-Risky Strategy
SS	Safe-Safe Strategy
SR	Risky-Safe Strategy
SR	Safe-Risky Strategy

Mathematical Symbols:

π_i or $\pi(i)$	Expected payoff acquired by agents using strategy i.
μ	mutation rate
α	imitation dynamic parameter
x	vector giving the population proportion of agent types
x_i	proportion of agents of type i in the population
$\nabla_\alpha(i,j)$	switching rate between agents of strategy i and j, see Section 2.3.2
p	probability of winning the risky lottery in a lottery game
$E(X,Y)$	expected payoff of strategy X when playing strategy Y in a game
s_i	proportion of *Stag* agents in the population at generation i
h_i	proportion of *Hare* agents in the population at generation i
N	PGG group size
r	PGG contribution multiplication factor
c	PGG contribution cost
l	punishment cost
ρ	punishment fine/cost of being punished
κ	action perception noise

Chapter 1

Introduction

In this thesis, I employ evolutionary game theoretic approaches that combine theoretical analysis and multi-agent system simulations to generate models of the evolution of various phenomena relating to human decision-making and culture. Following its original conception to study problems in evolutionary biology, Evolutionary Game Theory (EGT) has become an increasingly common tool used to model, explain, and predict phenomena of human behavior and societies. EGT is an attractive framework under which to model human behavior because it allows for the explicit modeling of changes in behaviors and norms over time based in a manner that is dependent on other behaviors in the population. Also, the social-science literature is filled with examples of empirical studies showing that humans violate the strong rationality assumptions of classical game theory. Under the EGT framework, such an assumption of *a priori* rational agents is not required.

This thesis advances the state of the art of evolutionary game theory and its applications to the social sciences by i) developing novel evolutionary game theoretic models to study different specific phenomena of interest in human behavior and culture that have not yet been explored in this manner, and ii) providing novel observations about the relationships between variables in these evolutionary game theoretic models and their impact on evolutionary outcomes, thereby providing new insights into the evolution of human decision-making behavior, social norms and culture.

Specifically, the models presented in this thesis show how a large range of population dynamics that model social imitation learning result in state-dependent risk preferences under sequential choice, how this principle facilitates the evolution of cooperation in classic game theoretic games where cooperation entails risk, and lastly how cross-cultural differences in punishment behavior to enforce cooperative norms are at least partially determined by cultures exposure to societal threats. These investigations are motivated and carried out under the close consideration of relevant social science data. The development and analysis of evolutionary game theoretic models of these aspects of human behavior and culture not only advances our understanding of how to use evolutionary game and multi-agent system models to study such problems, but also lays the foundation for more complex explanatory and predictive tools applicable to human populations and societies.

1.1 Motivation, Aim and Approach

Many questions about human behavior and culture of interest to social scientists have yet to be examined through an evolutionary game theoretic perspective in order to determine whether explanatory and predictive rather than merely descriptive insights can be gained. The primary goal of my evolutionary game theoretic models is to enhance our understanding of human behaviors and societies by shedding light on the relationships of various environmental factors and adaptation (social learning) processes to evolutionary behavioral outcomes and dynamics. This research is grounded in empirical social science data and conducted in close interdisciplinary collaboration with social scientists. Such an interdisciplinary approach is often crucial for the validity and usefulness of evolutionary game theoretic models.

Integrating evolutionary game theoretic models into social science can provide a complementary method to those that exist in the field. In order to understand our complex human social world, social scientists often create descriptive models of human behavior and decision making. Based on empirical data, these studies seek primarily to describe observed behavior rather than to understand their underlying dynamics and reasons for emergence. Similarly, empirical studies often find correlations among various historical, socio-cultural, and behavioral factors in human populations. However, no causal relationships in these correlations can be determined. By providing explanatory models of the emergence of different observed behaviors, evolutionary game theoretic models have the ability to describe the process and dynamics under which different behaviors emerge, and establish support for causal relationships among socio-cultural, historical, and behavioral factors that are often difficult or impossible to test or infer empirically. Furthermore, such models can be used to test hypothetical scenarios and to create predictive tools of human behavior and population norms.

Throughout the development of the evolutionary game theoretic models presented in this thesis, I used mathematical analysis and multi-agent system simulations in a symbiotic fashion. Simulations can be used as an exploratory tool to inspire theoretical investigations, to validate theoretical predictions, or to produce results when the model is too complex to meaningfully analyze mathematically. In return, when possible and useful, mathematical analyses serve to formalize the dynamics and relationships in the presented models and validate simulation results.

1.2 Evolutionary Game Theory

Generally, Lewontin [62] and Maynard-Smith [69] are credited with the pioneering of evolutionary game theory and its application in biological contexts. They recognized that the fitness of an organism (and hence its genes) is dependent on the abundance of other organisms in the environment (termed *frequency-dependent fitness*), and this fact critically influenced the evolution of biological species, genes, and traits. In addition to biological genetic inheritance, however, the evolutionary process can also be treated as an analogy to the processes of learning [43] and the cultural transmission and change of memes, behaviors, and norms in human societies [27, 17]. In this context, strategies do not correspond to different genes or species, but behaviors. The concept of frequency-dependent fitness undoubtedly plays a crucial role in the evolution of human behaviors and norms, since the fitness of a behavior of any one individual depends on the abundance and types of others' behaviors and norms in the population.

Thus, following its original conception to study problems in evolutionary biology, EGT has become an increasingly common tool used to model, explain, and predict phenomena of human behavior and societies. EGT provides the theoretical and practical computational tools to study the general, dynamic, and emergent properties of social systems and phenomena, including agent behaviors, culture, and norms. EGT studies the effects of evolutionary pressures on populations of agents: a population of agents making choices according to an assigned strategy in a game that models a situation of interest. After agents have played the game (or games), they reproduce into the next generation according to a reproduction function or *population dynamic* that, generally speaking, increases the frequency of the types of agents that were successful in the current generation. In this manner EGT adds an important dynamical aspect to classical game theory for studying human behavior.

Often this evolutionary process can lead to complex population dynamics that affect the change and prominence of agent behaviors over time. EGT generally aims to analyze populations under such dynamics and to identify and characterize solution concepts such as *evolutionary stable states* - states in which the population has stabilized and generally speaking cannot be taken over by an arbitrarily small number of invading new agents. Such analysis can aid the prediction of agent behaviors in given environments and improve the understanding of behaviors and norms by illuminating the (evolutionary) reasons for their existence. In addition, it is relatively easy to validate, complement, and aid EGT analysis using computer simulation.

The social science literature is filled with examples of empirical studies showing that humans violate the strong rationality assumptions of classical game theory. EGT is thus an attractive framework under which to model human behavior in such domains, because the assumption of perfectly rational agents - in the classical game theoretic sense - is not required. The focus of EGT approaches is generally not to define internal preferences and whether or not agents satisfy these. Rather, the focus is to allow for all of a viable range of preferences or strategies and to evaluate these according to their evolutionary fitness, which is external. The dynamic relative fitness of existing preferences and strategies then determines which strategies become norms, which of them become extinct, or if a variety of strategies remain in the population cycling in a predator-and-prey-type dynamic. To this date, evolutionary game theoretic approaches have been used to study a great variety of social and cultural phenomena. Examples of such phenomena studied through evolutionary games include cooperation, altruism, and reciprocity [6, 4, 8, 78, 93, 101, 25, 24, 48, 114, 74, 83, 15, 97, 97, 19, 80], trust and reputation [34, 71, 20, 61, 44], fairness and empathy [9, 77, 87], punishment [85, 16, 18, 49, 51, 20, 21, 103, 52], and social learning [100].

Despite the fact that EGT is generally not concerned with defining internal preferences of agents *a priori*, there exists a need for closer coupling of EGT work with empirical evidence and theories from social sciences. Many evolutionary environments of concern are too complex to allow for the feasible consideration of all possible combinations of agent strategies and game parameters. Empirical evidence from the social sciences can aid in limiting the complexity of the evolutionary game environment of interest, by constraining or specifying the possible strategies, model of interactions and social learning process. Furthermore, comparing results of EGT models to empirical data enables researchers to enhance and refine these models to make their predictions more specific and more accurate. Conversely, EGT models can identify aspects of the game environment crucial to evolutionary outcomes, thus providing information that can inspire additional empirical studies aimed at characterizing these aspects. EGT models can also provide predictive models that can be tested through actual empirical experiments.

1.3 Phenomena Explored

The phenomena I explore are i) risk preferences and implications on the evolution of cooperation and ii) the willingness of humans to punish others for norm violating behavior. The following two paragraphs give a brief motivation for each, along with an overview of the models employed and results acquired in this thesis.

My work on risk preferences is inspired by an abundance of empirical evidence that human decision-making under risk does not coincide with expected value maximization and is state-dependent.A considerable 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 determined by these descriptive models could have been learned or arisen evolutionarily. I contend 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. I introduce a framework of simple lottery games in which agents make sequential choices between risky and safe lotteries to study this problem. This framework allows for the study of risk behavior of agents in evolutionary environments through mathematical analysis and computer simulation. Investigations of this lottery game framework in conjunction with known population dynamics provide the novel insight that for a large range of population dynamics which model imitation, the interplay between risk-taking and sequentiality of choices leads to the emergence of state-dependent risk preferences that do not maximize expected value. The specific risk preferences that emerge resemble well-known descriptive models of decision-making. Aside from evolutionary lottery games I also explore a sequential evolutionary version of the well-known stag hunt game, demonstrating how the principal discoveries from my evolutionary lottery game framework can facilitate the evolution of cooperation in classic game theoretic games where cooperation entails risk.

My work on punishment is inspired by empirical social science that shows that there are striking differences across cultural groups in their willingness to punish norm-violating behavior. While punishment has been a prevalent research subject in psychology and EGT, the conditions under which different propensities of punishment are more or less adaptive for humans groups has received less attention. Using Public Goods Game models in which I operationalize several forms of societal threat, I show how these cross-cultural differences in the use of punishment are at least partially determined by differences in cultures' exposure to these threats. This illuminates the evolutionary basis for the wide variation in punishment rates that exists around the globe and helps promote cross-cultural understanding by showing how cultural differences in punishment propensities, which may appear puzzling, are generally adaptive to the society's ecological and historical context.

1.4 Thesis Outline

The rest of this proposal is organized as follows: Chapter 2 describes in detail my work on the evolution of state-dependent risk preferences in an evolutionary lottery game framework. Chapter 3 describes how the results on risk preferences in the lottery game framework apply to games of social interaction. Chapter 4 describes in detail my work on cultural differences in punishment norms. These Chapters provide motivation, relevant background, and related work on the phenomena explored, describe the evolutionary game theoretic models developed to study these phenomena, and provide results from mathematical analysis and multi-agent system simulations. Chapter 5 discusses how evolutionary game theoretic approaches fit into the study of culture in psychology. Chapter 6 concludes with summary of the results and contributions in this thesis.

Chapter 2

Evolution of State-Dependent Risk Preferences

Empirical evidence shows that human decision-making, rather than conforming to the decision-theoretic notion of expected-value maximization, is *state-dependent*: the decisions are sometimes risk-averse and sometimes risk-seeking, depending on the decision maker's circumstances. Much effort has been invested into describing and modeling such behavior, but these efforts have largely lacked an explicit investigation of what evolutionary pressures might have influenced the behavior's spread. Thus an important open question is why state-dependent risk behavior is so prevalent. I contend 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.

To demonstrate this, I use tools from evolutionary game theory to investigate how agents' risk behavior relates to different *population dynamics* (i.e., rules governing changes in the number of agents of each kind). The particular population dynamics I explore are *imitation dynamics*, which model cultural evolution as a product of social learning by imitation.

The probably best-known imitation dynamics are the replicator dynamic and the imitate-the-better dynamic, but there also are many others and empirical evidence on which one corresponds to human imitation is unclear. Hence I consider a parameterized class of *imitation dynamics* in which the parameter $0 \leq \alpha \leq 1$ yields the *replicator dynamic* with $\alpha = 1$ and the *imitate-the-better* dynamic with $\alpha = 0$. This study includes (1) a detailed mathematical analysis of how different imitation dynamics can affect risk behavior when agents make sequential choices, and (2) simulations, using several different imitation dynamics, of evolutionary lottery games in which agents make sequential choices among lotteries that have equal expected value but different risks. Results demonstrate that for every population dynamic in this class except for the replicator dynamic, the interplay between risk-taking and sequentiality of choices allows state-dependent risk behavior to have an evolutionary advantage over expected-value maximization.

My investigations also consider a sequential choice evolutionary version of the well-known stag hunt game. While the lottery game 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 expected values of the outcomes, the stag hunt game results show how this can facilitate the evolution of cooperation in situations where cooperation entails risk.

This work provides a starting point for further investigation of how population dynamics influence risk behavior in evolutionary game environments. I anticipate that state-dependent risk behavior will outperform expected-value-maximizing strategies in a large variety of evolutionary game environments involving sequential choices of different risks.

The next section provides background on human decision making under risk and evolutionary game theory. The following sections describe imitation dynamics, the lottery game framework, and analysis and simulation results on the evolution of state-dependent risk behavior.

2.1 Human Decision Making Under Risk

Human decision making under risk is the subject of much research effort in the social sciences. In most of the existing literature on models of human decision making under risk, the construction of such models is approached primarily through the analysis of a decision maker's choices among lotteries that have different payoff distributions, and thus potentially different risks. Under the most traditional model of decision making, expected utility theory, a rational agent's preferences can be modeled by assigning to each possible outcome a number called the outcome's *utility*; and a rational choice is one that maximizes the expected utility of the outcomes [120]. 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 [60]. Expected utility theory can account for different attitudes towards risk through certain von Neumann-Morgenstern utility functions (e.g. [38]). 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 under the expected-utility model (e.g. [38, 58, 92]). Related efforts in economics have aimed to describe the preferences of humans over inter-temporal lotteries, recognizing the effects of temporally successive lotteries on risk preferences [32, 33]. Other studies define utility functions that take into account inter-personal or population comparisons [1]. Yet other researchers have constructed alternative descriptive theories of decision making that claim to correspond more closely to how humans make decisions involving risk. Among the most popular of these models are prospect theory [60, 118], regret theory [64], and SP/A (Security-Potential/Aspiration) theory [65, 66, 67]. 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. A common theme of both Prospect theory and SP/A theory is that agents are risk-averse when they have done well relative to some reference point, and risk-seeking when they have not done well relative to the reference point.

Several recent works speculate about the relation of risk-related behavior and biological evolutionary factors [54, 94, 107]. 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.

To study risk behavior in the framework of EGT, I model the situation of interest as a game in which agents are faced with choices among lotteries of different risks. As described in [2], these lotteries dispense resources that are considered to be an objective quantity of which 1) agents always want more than less and 2) interpersonal comparisons are meaningful. The reproduction function defining the dynamics of strategies in the population then acts directly on these resources.

2.2 Imitation Dynamics

Imitation dynamics are a class of population dynamics commonly used to model the evolution of behaviors in societies [53, 76, 55, 75, 31]. The general framework for imitation dynamics is stated by Hofbauer and Sigmund [53] as follows:

> We shall suppose that occasionally a player is picked out of the population and afforded the opportunity to change his strategy. He samples another player at random, and adopts his strategy with a certain probability.

In what follows, I refer to these players as the *observer* and the *observed agent*, respectively.

Important theoretical studies have been done of two specific imitation dynamics. One of these is the *replicator dynamic* [111, 98, 99, 53, 40], in which the probability that the observer adopts the strategy of the observed agent is proportional to how much more successful the observer was than the observed. The other is the *imitate-the-better* dynamic [10, 119, 109, 93, 46, 26, 112, 116],[1] in which the observer always adopts the observed agent's strategy if it was more successful than the observer's strategy.[2]

---

[1] The imitate-the-better dynamic is sometimes called tournament selection [93].

[2] Vega-Redondo generalizes the imitate-the-better dynamic by allowing the observer to observe a collection of agents and adopt the strategy of the most successful agent [119].

Several experimental studies investigating social learning through imitation between humans have found experimental support for Vega-Redondo's model [56, 57, 81]. Experiments on human imitation reported by Apesteguia et al. [3] indicate that the difference in observed payoff to an agent's own payoff does affect imitation (the higher the difference, the more likely imitation occurs). This is more in line with Schlag's model. Due to this evidence for both imitation models, in I explore a parameterized range of imitation dynamics based on a definition in [53], that includes the replicator dynamic, the imitate-the-better dynamic, and a spectrum of other dynamics in between those two.

2.3 Evolutionary Lottery Game Model

Here I describe the *sequential lottery game*, a class of games that I use to investigate risk behavior under evolutionary pressures. I also describe the particular range of imitation dynamics under which I explore the evolution of risk behavior.

2.3.1 Sequential Lottery Game

We shall consider a game in which agents acquire payoffs dispensed by lotteries. In each generation, each agent must play an n-lottery game defined as follows:

Definition 1. *An n-lottery game is a game in which an agent make a sequence of n choices, where each choice is between two lotteries*[3]*: a* ***safe*** *lottery, whose payoff*

---

[3] The choice of payoffs here is arbitrary. The results in this chapter would be roughly qualitatively similar if we had used any fixed value v for the safe lottery's payoff and $v + \delta$ and $v - \delta$ (for fixed δ) for the risky lottery's payoffs.

is always 4, and a ***risky*** *lottery, which one can* win *(a payoff of 8) with probability p, or* lose *(a payoff of 0) with probability* $1 - p$.

Note that if $p = 0.5$, both lotteries have expected value 4.

Our population consists of agents that follow strategies chosen from the set $\mathcal{S} = \{s_1, \ldots, s_k\}$ of all possible pure strategies for the sequence of lottery choices. In any generation, a vector $\mathbf{x} = (x_1, \ldots, x_k)$ gives the state of the population, where each x_i is the proportion of agents in the population using strategy s_i. Let $\pi(i)$ denote the payoff accumulated in a generation from the n lottery choices by agents of type i (i.e. agents following strategy s_i).

2.3.2 Population Dynamics

As discussed in Section 2.2, we want to explore a range of population dynamics that includes the replicator dynamic, imitate-the-better dynamic, and dynamics intermediate between these two extremes. Hofbauer and Sigmund [53] give the following parameterized formula for these population dynamics:

$$\dot{x_i} = x_i \sum_j x_j |\pi(i) - \pi(j)|^\alpha \; \mathsf{sign}(\pi(i) - \pi(j)), \tag{2.1}$$

where x_i is the current proportion in the population of agents of type i, $\dot{x_i}$ is the change in x_i over time,[4] and $\alpha \geq 0$ is a parameter that determines the particular

---

[4] As is common in the evolutionary game theory literature, Hofbauer and Sigmund approximate the current population as a real-valued function $x_i(t)$ where t is the current time, so that $\dot{x_i} = dx_i/dt$.

imitation dynamic. My formulation is based on theirs, but incorporates the following changes:

- We are interested in population dynamics based on payoff comparisons among individuals, as in [10, 109, 93, 46, 112, 116, 26]. To model payoff comparisons among individuals, we must take into account the stochastic variability in the payoffs to individual agents. We can do this by treating $\pi(i)$ and $\pi(j)$ as discrete random variables representing the distributions of payoffs that an agent of type i and an agent of type j receive from their lottery choices.
- The imitate-the-better dynamic and the replicator dynamic correspond to $\alpha = 0$ and $\alpha = 1$, respectively. Since the imitation dynamics that interest us are these two and and the ones that are intermediate between them, we only consider $0 \leq \alpha \leq 1$.

If we let r and s be any possible payoff values acquired by agents of type i and j, and let $p(r, s)$ the probability of obtaining this pair of values, then our modified version of Eq. (2.1) is

$$\dot{x_i} = x_i \sum_j x_j \nabla_\alpha(i, j), \tag{2.2}$$

where $\nabla_\alpha(i, j) = \sum_r \sum_s |(r - s)|^\alpha \, \mathsf{sign}(r - s) \cdot p(r, s)$ is the *switching rate* between two agent types i and j. This *switching rate* determines the effect that a pairing for imitation between agents of type i and j has, on average, on type i's growth rate in the population. For example, if $\nabla_\alpha(i, j)$ is positive, a pairing for imitation between agents of type i and j (which happens with non-zero probability if agents of type i

Table 2.1: All of the possible pure strategies in our lottery game 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

and j exist in the population) on average has a positive effect on i agents' growth rate. If this is the case, we say i has an *evolutionary advantage* over j.

2.4 Analysis and Theoretical Predictions

In the two-choice sequential lottery game, there are six possible pure strategies. These are listed in Table 2.1.[5]

Table 2.2 gives each each strategy, its possible numeric payoffs, and the probabilities of these payoffs for the case where p, the probability of wining the risky lottery, is 0.5. Note that in this case, all six strategies have the same expected

---

[5] For simplicity, I have restricted this study to pure strategies. For lottery games like the ones we are considering, the reproductive fitness of a mixed strategy is intermediate among the pure strategies in the mixed strategy's support, hence the inclusion of mixed strategies should not have a substantial affect on our results.

Table 2.2: Payoff distributions of the six pure strategies in the sequential lottery game with $n = 2$ and $p = 0.5$. All six strategies have the same expected payoff, 8, but differing distributions. In this table and others in this section, we use boldface numbers to denote payoff values, and non-boldface to denote probabilities.

	RwS			RwR			SR		RS		SS	RR		
Payoff $\boldsymbol{V}$	**12**	**8**	**0**	**16**	**8**	**4**	**12**	**4**	**12**	**4**	**8**	**16**	**8**	**0**
Prob. $P(V)$	.5	.25	.25	.25	.25	.5	.5	.5	.5	.5	1	.25	.5	.25

payoff value of 8, but they have differing probabilities of being above or below 8. For example, $P[\pi(\text{RwS}) > 8] = 0.5$ and $P[\pi(\text{RwS}) < 8] = 0.25$.

In this section, we shall examine how these strategies will perform against one another for different population dynamics (Sections 2.4.1, 2.4.2, and 2.4.4) and for different expected values of the lotteries (Section 2.4.4), culminating in an evolutionary stability result for the RwS strategy (Section 2.4.5). Finally, Section 2.4.6 briefly discusses cases where $n \neq 2$.

2.4.1 The Replicator Dynamic

I now show that under the replicator dynamic ($\alpha = 1$), the switching rate between any two agent types is equal to the difference in their expected payoffs.

Proposition 1. *Under the population dynamics given by Eq. (2.2) when* $\alpha = 1$, $\nabla_\alpha(i, j)$ *(the switching rate between any two agent types* i *and* j*) is the difference between the expected payoffs of agents of type* i *and* j *(given by the discrete random variables* $\pi(i)$ *and* $\pi(j)$*).*

Proof. Let r and s be possible payoff values acquired by agents of type i and j, and let $p(r,s)$ be the probability of obtaining this pair of values. With $\alpha = 1$,

$$\begin{aligned}\nabla_\alpha(i,j) &= \sum_r \sum_s |(r-s)|^\alpha \ \mathrm{sign}(r-s) \cdot p(r,s) \\ &= \sum_r \sum_s (r-s) \cdot p(r,s) = EV(\pi(i) - \pi(j)).\end{aligned}$$

Assuming independence between payoffs, $\nabla_\alpha(i,j) = EV(\pi(i) - \pi(j)) = EV(\pi(i)) - EV(\pi(j))$, and the proposition follows. □

Since all strategies have the same expected payoff in this environment, Proposition 1 tells us that the switching rate between any two strategies will be 0. Consequently, all six of the above strategies will perform equally well evolutionarily.

2.4.2 The Imitate-the-Better Dynamic

If we use the imitate-the-better dynamic ($\alpha = 0$), then in Eq. (2.2), only the sign of the payoff difference between two paired agents plays a role in determining the switching rate; the magnitude of the difference is irrelevant. We can compute the switching rate between two strategies by using the probabilities in Table 2.2 to calculate the probability of each pair of payoffs occurring. Table 2.3 shows these values for the RwS vs. RR pairing. We can then use these probabilities as the

Table 2.3: Payoff distribution for RwS vs. RR in the sequential lottery game with $n = 2$ and $p = 0.5$. Entries in the table show give the probabilities that each pair of payoffs occurs.

		RwS		
		12	**8**	**0**
RR	**16**	0.125	0.0625	0.0625
	8	0.25	0.125	0.125
	0	0.125	0.0625	0.0625

values of $p(r, s)$ in Eq. (2.2) to calculate the switching rate, as follows:

$$\begin{aligned}
\nabla_\alpha(\text{RwS}, \text{RR}) \quad &= \text{sign}(0-0)\cdot 0.0625 + \text{sign}(8-0)\cdot 0.0625 + \text{sign}(12-0)\cdot 0.125 \\
&\quad + \text{sign}(0-8)\cdot 0.125 + \text{sign}(8-8)\cdot 0.125 \\
&\quad + \text{sign}(12-8)\cdot 0.25 + \text{sign}(0-16)\cdot 0.0625 \\
&\quad + \text{sign}(8-16)\cdot 0.0625 + \text{sign}(12-16)\cdot 0.125 \\
&= 0.0625 + 0.125 - 0.125 + 0.25 - 0.0625 - 0.0625 - 0.125 \\
&= 0.0625.
\end{aligned}$$

Using similar calculations, we see that $\nabla_0(\text{RwS}, \text{SS}) = 0.25$, $\nabla_0(\text{RwS}, \text{RwR}) = 0.0625$, and $\nabla_0(\text{RwS}, \text{RS}) = \nabla_0(\text{RwS}, \text{SR}) = 0$. This suggests that RwS will be able to consistently win an evolutionary competition against RR, RwR, or SS and remain stable with SR and RS in this environment. The experimental results in Section 2.5 verify this prediction.

2.4.3 All Imitation Dynamics (Arbitrary α)

For $0 \leq \alpha \leq 1$, we can calculate the switching rate in a method similar to the previous section, combining the probabilities from Table 2.2 to get the values of $p(r, s)$ for Eq.(2.2). However, since both the sign and magnitude of the payoff differences are now important, the calculation is slightly more complex. For instance, the switching rate for the RwS vs. RR pairing is now as follows:

$$\begin{aligned}
\nabla_\alpha(\text{RwS}, \text{RR}) =& |0-0|^\alpha \operatorname{sign}(0-0) \cdot 0.0625 + |8-0|^\alpha \operatorname{sign}(8-0) \cdot 0.0625 \\
&+ |12-0|^\alpha \operatorname{sign}(12-0) \cdot 0.125 + |0-8|^\alpha \operatorname{sign}(0-8) \cdot 0.125 \\
&+ |8-8|^\alpha \operatorname{sign}(8-8) \cdot 0.125 + |12-8|^\alpha \operatorname{sign}(12-8) \cdot 0.25 \\
&+ |0-16|^\alpha \operatorname{sign}(0-16) \cdot 0.0625 + |8-16|^\alpha \operatorname{sign}(8-16) \cdot 0.0625 \\
&+ |12-16|^\alpha \operatorname{sign}(12-16) \cdot 0.125 \\
=& 4^\alpha(0.125) + 8^\alpha(-0.125) + 12^\alpha(0.125) + 16^\alpha(-0.0625).
\end{aligned}$$

Figure 2.1 shows how the switching rate between RwS and the other strategies varies with α. One can see that RwS has an advantage over *all* other strategies for $0 < \alpha < 1$, suggesting that RwS should be able to win any evolutionary competition in these environments. Again, this prediction is supported by the simulation results in Section 2.5.

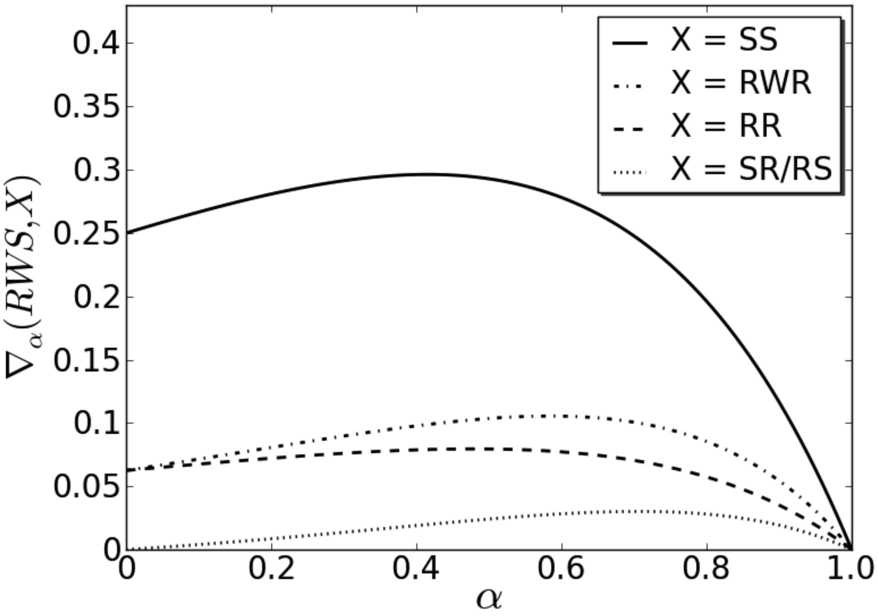

Figure 2.1: Switching rate between RwS and each of the other five pure strategies for $0 \leq \alpha \leq 1$. A positive switching rate indicates an evolutionary advantage of RwS against the other strategy (in a population made up of solely the two strategies). We see that RwS has such an advantage over all other strategies when $0 < \alpha < 1$, and all but SR and RS when $\alpha = 0$.

2.4.4 Differing Expected Values of Lotteries (Arbitrary p)

We now consider the case where p, the probability of winning the risky lottery, is any number between 0 and 1 (whence the risky lottery's expected value is between 0 and 8). Table 2.4 gives the probability distributions for each pure strategy.

We can also construct a new probability matrix for each pairing, such as the one for RwS vs. RR shown in Table 2.5. We can then compute the switching rate for our pairing as before. For example, the switching rate for RwS vs. RR is now

$$\begin{aligned}
\nabla_\alpha(\mathrm{RwR},\mathrm{RR}) =& |0-0|^\alpha \operatorname{sign}(0-0)\cdot(1-p)^4 \\
&+ |8-0|^\alpha \operatorname{sign}(8-0)\cdot(1-p)^3 p \\
&+ |12-0|^\alpha \operatorname{sign}(12-0)\cdot(1-p)^2 p \\
&+ |0-8|^\alpha \operatorname{sign}(0-8)\cdot 2(1-p)^3 p \\
&+ |8-8|^\alpha \operatorname{sign}(8-8)\cdot 2(1-p)^2 p^2 \\
&+ |12-8|^\alpha \operatorname{sign}(12-8)\cdot 2(1-p) p^2 \\
&+ |0-16|^\alpha \operatorname{sign}(0-16)\cdot (1-p)^2 p^2 \\
&+ |8-16|^\alpha \operatorname{sign}(8-16)\cdot (1-p) p^3 \\
&+ |12-16|^\alpha \operatorname{sign}(12-16)\cdot p^3 \\
=& 4^\alpha(2(1-p)p^2 - p^3) + 8^\alpha((1-p)^3 p - 2(1-p)^3 p - (1-p)p^3) \\
&+ 12^\alpha(1-p)^2 p + 16^\alpha(1-p)^2 p^2
\end{aligned}$$

Notice that when $p > 0.5$, the risky lottery has a higher expected value than the safe lottery, and the opposite is true when $p < 0.5$. Thus, RR has the highest expected value when $p > 0.5$, and SS has the highest expected value when $p < 0.5$.

Table 2.4: Payoff distributions of the six pure strategies in the sequential lottery game with $n = 2$ and $0 \leq p \leq 1$.

	RwS			RwR		
Payoff *V*	**12**	**8**	**0**	**16**	**8**	**4**
Prob. $P(V)$	p	$(1-p)p$	$(1-p)^2$	p^2	$(1-p)p$	$(1-p)$

	RS, SR		SS	RR		
Payoff *V*	**12**	**4**	**8**	**16**	**8**	**0**
Prob. $P(V)$	p	$(1-p)$	1	p^2	$2(1-p)p$	$(1-p)^2$

Table 2.5: Payoff distribution for RwS vs. RR in the sequential lottery game with $n = 2$ and arbitrary p. Entries in the table show give the probabilities that each pair of payoffs occurs.

		RwS		
		12	**8**	**0**
RR	**16**	p^3	$(1-p)p^3$	$(1-p)^2p^2$
	8	$2(1-p)p^2$	$2(1-p)^2p^2$	$2(1-p)^3p$
	0	$(1-p)^2p$	$(1-p)^3p$	$(1-p)^4$

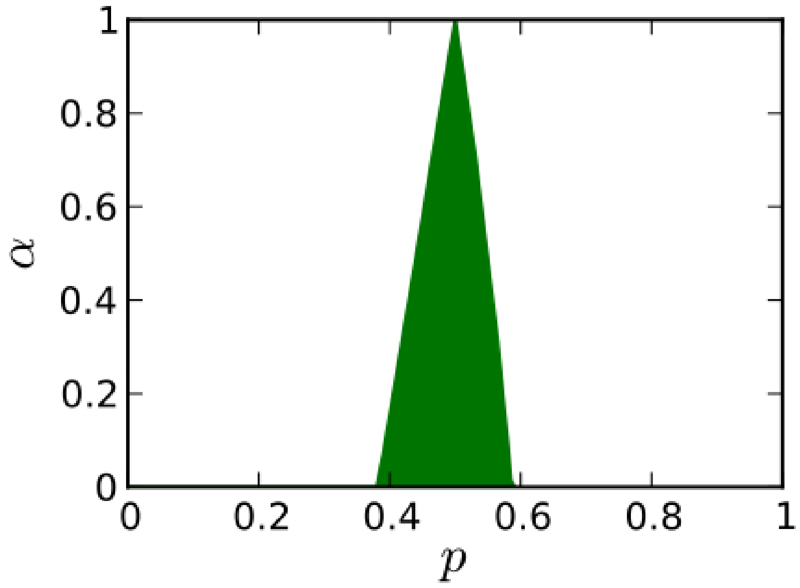

Figure 2.2: The shaded area indicates values of p and α for which the switching rate between RwS and the expected-value maximizing strategy (i.e. RR if $p \geq 0.5$, SS otherwise) is greater than 0. RwS is at a disadvantage in terms of expected utility when $p \neq 0.5$, but it still manages to retain its evolutionary advantage for a wide range of values.

Surprisingly, even though RwS has a suboptimal expected value when $p \neq 0.5$, by examining the switching rates we can see that it still has an evolutionary advantage over both SS and RR for many values of p and α. Figure 2.2 shows the values of p and α for which $\nabla_\alpha(\mathrm{RwS}, \mathrm{RR}) > 0$ and $\nabla_\alpha(\mathrm{RwS}, \mathrm{SS}) > 0$, meaning that for these values of p and α, RwS has an evolutionary advantage over expected-value maximizing strategies. Under the imitate-the-better dynamic ($\alpha = 0$), the range is surprisingly large. For example, RwS outperforms SS even when $p = 0.4$, and SS has a significantly higher expected value than RwS. As α increases, the range shrinks at a roughly linear rate, disappearing at $\alpha = 1$ (i.e., the replicator dynamic).

2.4.5 Evolutionary Stability of State-Dependent Risk Behavior

This section discusses whether RwS is an *evolutionarily stable strategy* (ESS) in the 2-lottery game when $p = 0.5$ and $0 < \alpha < 1$. I first give the classical definition of an ESS, explain why it cannot be directly applied to n-lottery games when $\alpha \neq 1$, and propose an intuitive modification to make it applicable. I then show that RwS fits our modified definition of an ESS.

2.4.5.1 Evolutionary Stability under Imitation Dynamics

In an evolutionary game, a population of agents using an ESS is resilient against an arbitrarily small number of incoming agents that use any other strategy [69]. According to Maynard Smith, strategy S is an ESS if for every strategy $T \neq S$, one of the following conditions holds:

- $E(S, S) > E(T, S)$;
- $E(S, S) = E(T, S)$ and $E(S, T) > E(T, T)$;

where $E(X, Y)$ is the expected payoff an agent receives by playing strategy X against strategy Y [69].

This definition does not apply directly to the n-lottery game, because it assumes the population dynamics are governed by the replicator equation, for which the expected payoff of a strategy pairing is all that is necessary to determine whether one strategy will grow or shrink in population proportion compared to the other. As described in the work summarized in Section 2.4.4, with imitation dynamics other than the replicator dynamic, the expected value of a pairing is *not* sufficient to determine which strategy will perform better in the n-lottery game. Instead, one must use the switching rate $\nabla_\alpha(X, Y)$ rather than $E(X, Y)$ when defining evolutionary stability for the n-lottery game under imitation dynamics. This is appropriate because if there are many agents using X and few using Y, then

- $\nabla_\alpha(X, X) > \nabla_\alpha(Y, X)$ implies that X will grow faster playing against itself than Y will grow playing against X, so Y will not be able to gain population and will eventually die off;

- $\nabla_\alpha(X, X) = \nabla_\alpha(Y, X)$ and $\nabla_\alpha(X, Y) > \nabla_\alpha(Y, Y)$ implies that X and Y grow at the same rate when playing against X, but X grows faster than Y when playing against Y, so Y will still not be able to gain population and will eventually die off.

and these scenarios correspond to the two conditions for X to be an ESS in the classical definition.

Replacing expected value with switching rates in the definition above gives us the following definition for an ESS in the n-lottery game:

Definition 2. *A strategy S is an evolutionarily stable strategy in the n-lottery game if on of the following two conditions holds:*

- $\nabla_\alpha(S, S) > \nabla_\alpha(T, S)$ *or*

- $\nabla_\alpha(S, S) = \nabla_\alpha(T, S)$ *and* $\nabla_\alpha(S, T) > \nabla_\alpha(T, T)$

for all $T \neq S$.

Combining these conditions and considering the definition of the n-lottery game (Definition 1), we get the following simpler condition for ESS:

Lemma 1. *A strategy S is an evolutionarily stable strategy (ESS) in the n-lottery game if, for any strategy $T \neq S$, $\nabla_\alpha(S,T) > 0$.*

Proof. If for any strategy $T \neq S$ $\nabla_\alpha(S,T) > 0$, then $\nabla_\alpha(S,S) > \nabla_\alpha(T,S)$ holds for any strategy $T \neq S$, since $\nabla_\alpha(S,S) = 0$ and $\nabla_\alpha(S,T) = -\nabla_\alpha(T,S)$ for all S and T. Thus the first condition of Definition 2 holds, which is sufficient for S to be an ESS. □

2.4.5.2 RWS is Evolutionarily Stable

This section shows that RwS is an ESS by Lemma 1 for the 2-lottery game with $p = 0.5$ and $0 < \alpha < 1$. To do this, one must show that it has a positive switching rate with an arbitrary strategy. Therefore, the first step must be to devise a method for representing an arbitrary strategy for the 2-lottery game. One can find that, if the decisions an agent makes and the possible lottery outcomes are arranged into a game tree as shown in Figure 2.3, then any strategy can be expressed as $S_{a,b,c,d}$, where a, b, c, and d give the agent's probability of choosing the risky lottery at each of the four decision nodes indicated in the figure. For instance, the pure strategies we have been dealing with thus far can be represented as follows (here, a "-" in place of one of the four probabilities indicates that any value is acceptable, since the decision corresponding to that probability is never reached):

- SS is $S_{0,-,-,0}$

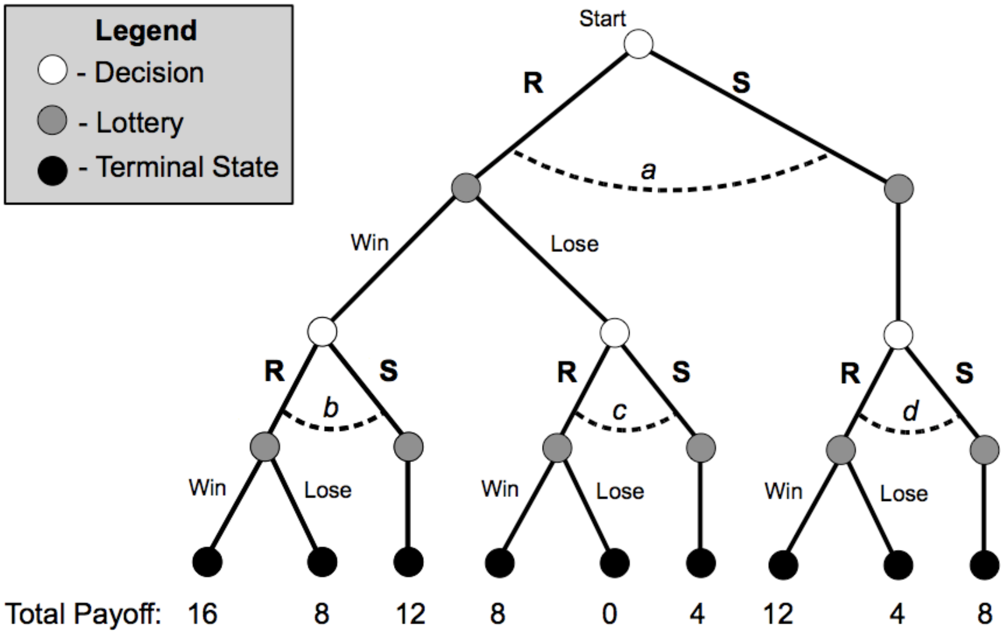

Figure 2.3: Representation of the 2-lottery game as a game tree consisting of decision nodes, in which the agent chooses between a risky (**R**) and safe (**S**) lottery, lottery nodes, and terminal nodes, which assign the agent its total payoff. Any strategy in this game can be represented as $S_{a,b,c,d}$, where a, b, c, and d give the probabilities of choosing the risky lottery at each of the four corresponding decision nodes.

- RR is $S_{1,1,1,-}$
- RS is $S_{1,0,0,-}$
- SR is $S_{0,-,-,1}$
- RwS is $S_{1,0,1,-}$
- RwR is $S_{1,1,0,-}$

We can now calculate the switching rate between RwS and an arbitrary strategy $S_{a,b,c,d}$ in terms of a, b, c, d, and α. This comes out to:

$$\begin{aligned}\nabla_\alpha(\text{RwS}, S_{a,b,c,d}) = & \left[\tfrac{1}{8}(2 * 8^\alpha - 12^\alpha - 4^\alpha)\right](a(1-c) + (1-a)d) \\ + & \left[\tfrac{1}{4}(2 * 4^\alpha - 8^\alpha)\right](1-a)(1-d) \\ + & \left[\tfrac{1}{16}(2 * 12^\alpha + 2 * 4^\alpha - 2 * 8^\alpha - 16^\alpha)\right] ab.\end{aligned} \tag{2.3}$$

Appendix A gives a full derivation. Intuitively, this says that the switching rate for the arbitrary strategy is just the probability that the strategy follows the pure strategies SS, SR, RS, and RR, times the switching rate between RwS and each of those strategies. Given the above derivation for $\nabla_\alpha(\text{RwS}, S_{a,b,c,d})$, we get the following evolutionary stability result:

Theorem 1. RwS *is an evolutionarily stable strategy in the n-lottery game with* $p = 0.5$ *and* $0 < \alpha < 1$.

Proof. By Lemma 1, we know that RwS is an evolutionarily stable strategy if $\nabla_\alpha(\text{RwS}, S_{a,b,c,d}) > 0$ holds. Eqn. 2.3 gives this switching rate between RwS and any other arbitrary strategy $S_{a,b,c,d}$ in the 2-lottery game. The three bracketed terms are all strictly greater than 0 for $0 < \alpha < 1$; in fact, they are identical to the curves in Figure 2.1 for SR/RS, SS, and RR, respectively. Given that the bracketed terms are strictly positive, $\nabla_\alpha(\text{RwR}, S_{a,b,c,d}) > 0$ unless $a = 1$, $b = 0$, and $c = 1$. Since $S_{1,0,1,-}$ is equivalent to RwS, this means that $\nabla_\alpha(\text{RwR}, S_{a,b,c,d}) > 0$ for all $S_{a,b,c,d} \neq$ RwS. Thus, RwS is an evolutionarily stable strategy by Lemma 1 and Definition 2. □

2.4.6 Higher Number of Sequential Choices ($n > 2$)

With the exception of Section 2.4.5.1, the presented theoretical development has been largely restricted to the case $n = 2$. This Section desceibes briefly what happens for other values of n.

The case $n = 1$ is relatively trivial: there are only two pure strategies, both are unconditional, and both perform equally well (for more details, see [95]).

The case $n > 2$ is very hard to analyze, because the number of pure strategies is super-exponential in n. However, intuition suggests that the behavior pattern exhibited by RwS for $n = 2$, namely to play safe when having done well and risky otherwise, should also have an advantage when $n > 2$. I discuss some simulation experiments that support this intuition in Section 2.5.2.1.

2.5 Simulations

For further investigations of the dynamics in a population consisting of all pure strategies, the next section describes results of computer simulations of agent-based models playing the two-choice evolutionary lottery game.

2.5.1 Setup and Implementation

Simulations for the two-choice lottery game environment in this Section explore population evolution under a variety of parameter combinations of α (the imitation parameter) and p (the probability of winning the risky lottery). The types of agents included were the six pure strategies for the two-choice game described earlier. All simulations started with an initial population of 1000 agents for each agent type.

To correspond with the imitation dynamics given by Eq. (2.2) in the finite population agent-based model, I used a pairwise comparison process [112] to model the transmission of strategies among agents. Each generation, after all agents have received payoffs from chosen lotteries, each agent i compares its (individual) payoff $\rho(i)$ to that of a randomly drawn agent from the population j (with payoff $\rho(j)$) and adopts the strategy of this agent with a probability q if $\rho(j) > \rho(i)$. In order to achieve the parameterized dynamics given by Eq. (2.1) in the agent-based model, I use

$$q = [|\rho(j) - \rho(i)|/\Delta\rho]^{\alpha}, \quad 0 \le \alpha \le 1,$$

where $\Delta\rho$ is the highest possible difference in payoff. Figure 2.4 provides pseudocode on how this pairwise comparison imitation process was implemented.

```
LOOP for each agent in population
    Select current agent (imitator) from population
    Select random agent (observed) from pop
    Let $\rho_{im}$ = payoff achieved by imitator
    Let $\rho_{obs}$ = payoff achieved by observed
    Let $\Delta\rho$ = highest possible payoff difference
    Let $r = \text{random number in } [0, 1)$
    Let $q = [|\rho_{im} - \rho_{obs}|/\Delta\rho]^{\alpha}$
    # Imitate observed agent with probability q based on payoff comparison
    IF $\rho_{im} < \rho_{obs}$ AND $r <= q$THEN
        Place an agent following the strategy of observed into next generation
    ELSE
        Place an agent following the strategy of imitator into next generation
    END IF
END LOOP
```

Figure 2.4: Pseudo-code for the pairwise comparison imitation process used to reproduce agents into the next generation in the evolutionary simulations. Takes parameter $0 \leq \alpha \leq 1$.

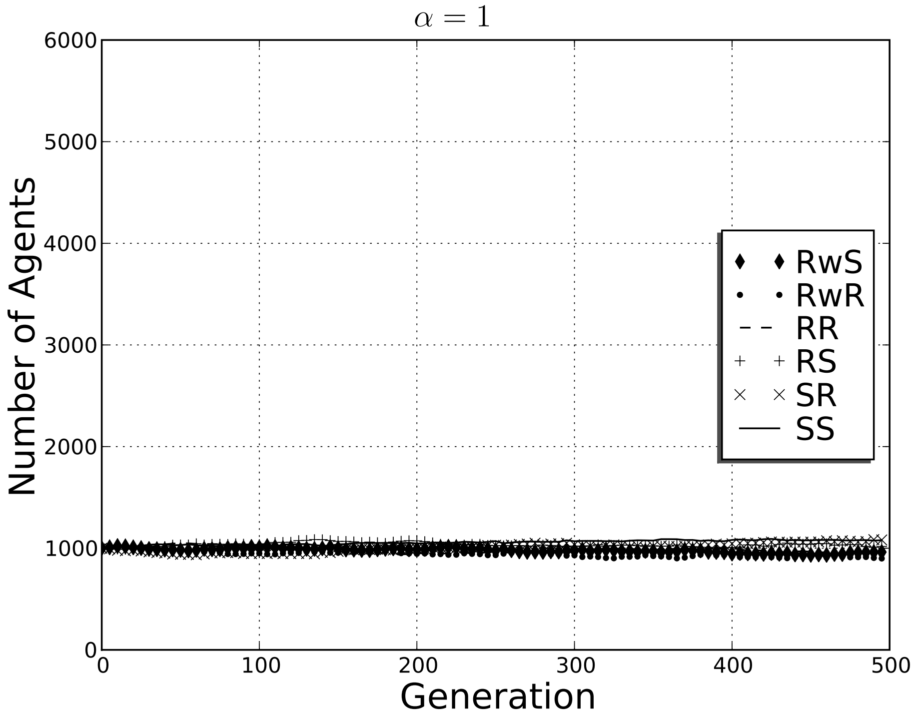

Figure 2.5: Results of simulation for $\alpha = 0$ in the population dynamic and $p = 0.5$. Plot shows the number of agents of each type over the course of evolution for 500 generations.

2.5.2 Results

Figures 2.5, 2.6, and 2.7 show the results for simulations with $p = 0.5$ for each $\alpha = 1$ (the replicator dynamic), $\alpha = 0$ (the imitate-the-better dynamic), and $\alpha = 0.5$. Each plot is an average over 20 simulation runs (the amount of variation from one run to another was quite small). These experiments confirm my analysis from Section 2.4.5.2, which shows that RwS has an evolutionarily advantageous risk behavior under any $0 \leq \alpha < 0$.

As predicted by my analysis, RwS outperformed the other strategies evolutionarily except when $\alpha = 1$:

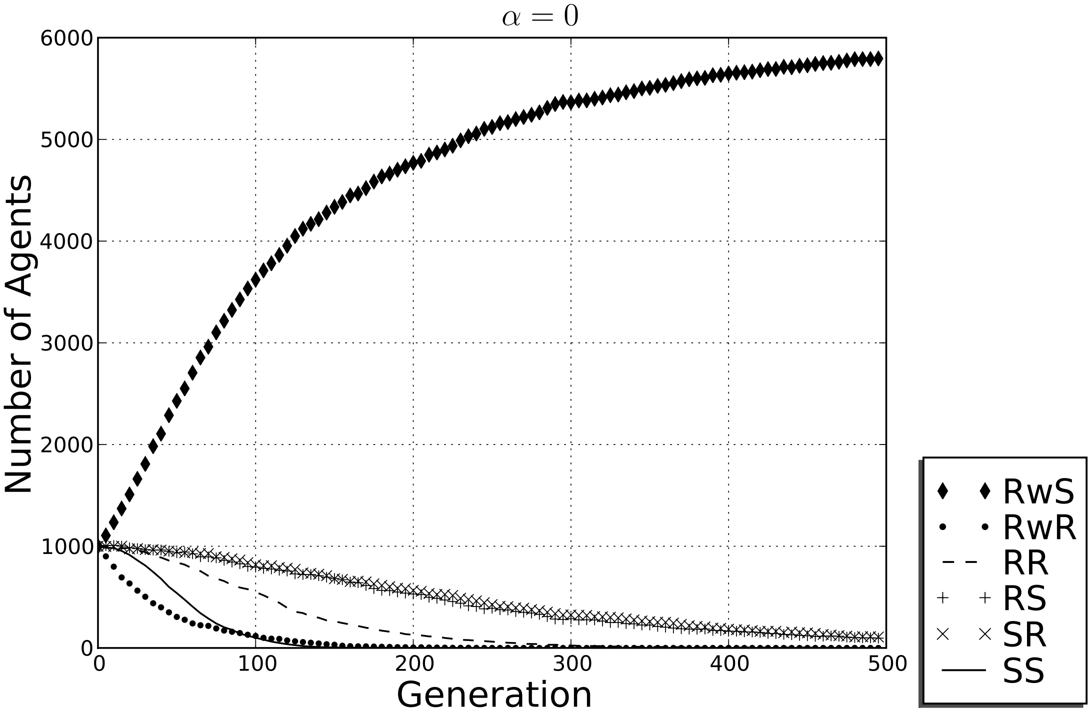

Figure 2.6: Results of simulation for $\alpha = 0$ in the population dynamic and $p = 0.5$. Plot shows the number of agents of each type over the course of evolution for 500 generations.

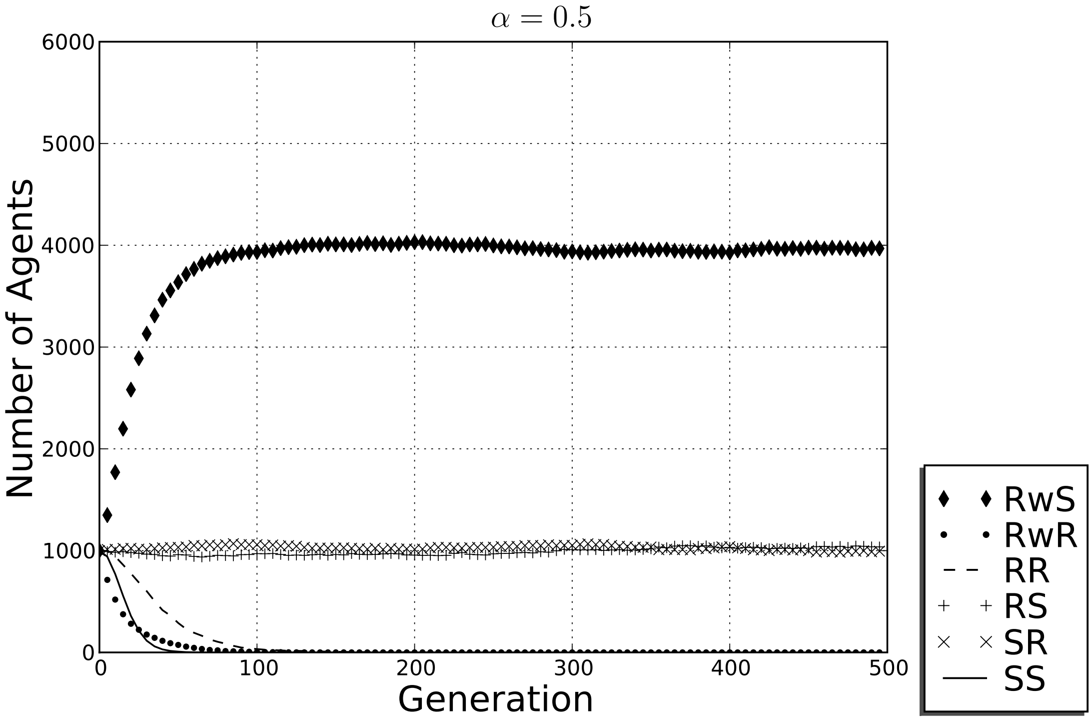

Figure 2.7: Results of simulation for $\alpha = 0.5$ in the population dynamic and $p = 0.5$. Plot shows the number of agents of each type over the course of evolution for 500 generations. Simulation runs for $\alpha = 0.2, 0.4, 0.6$, and 0.8 produced qualitatively identical results to the case of $\alpha = 0.5$.

- For $\alpha = 1$, all of the strategies performed equally well and remained at their initial population counts.

- For $\alpha = 0$, the state-dependent strategy RwS outperformed the other strategies. RwS rose in population proportion relatively quickly to comprise the majority ($> 2/3$) of the population and remained there throughout subsequent generations. Furthermore, the two unconditional strategies SR and RS remained, comprising the proportion of the population not taken over by RwS.

- For $\alpha = 0.5$, the RwS agent population grew similarly as for $\alpha = 0$, but here RwS also had an advantage against SR and RS (as indicated by Figure 2.1) and thus continued to grow to comprise 100% of the population.

I also ran simulations with $\alpha = 0.2, 0.4, 0.6$, and 0.8. The results for these α values are all essentially equal to the case of $\alpha = 0.5$. The only difference is that the rate at which RwS grows to take over the population is inversely related to α (i.e. for larger α values, it takes longer for RwS to take over the population).

In order to explore lottery games in which the risky lottery has a different expected value than the safe lottery, I also ran experiments with $p = 0.3$, 0.4, 0.55, and 0.7. These values were chosen because for $\alpha = 0$, Section 2.4.4 showed that for two of them ($p = 0.4, 0.55$) RwS has an evolutionary advantage over the expected-value-maximizing strategy, and for the other two ($p = 0.3, 0.7$) the expected-value-maximizing strategy has an evolutionary advantage over RwS.

As shown in Figure 2.8, 2.9, 2.10, and 2.11, the simulation results confirm the theoretical predictions. More specifically:

- For $p = 0.2$ (Figure 2.8), SS is the expected-value maximizing strategy and it takes over the population.

- For $p = 0.7$ (Figure 2.11), RR is the expected-value maximizing strategy and it takes over the population.

- Even though SS is the expected-value-maximizing strategy for $p = 0.4$ (Figure 2.9) and RR is for $p = 0.55$ (Figure 2.10), in both cases RwS has an evolutionary advantage and takes over the population.

In Figures 2.9, 2.10, and 2.11, some fluctuations occur before stabilization. These occur because of the differing amounts of evolutionary advantage that different strategies have over others. For example, a strategy a may grow in number temporarily because it has an advantage over another strategy b. But once b becomes extinct (or sufficiently small in number), a will diminish because some other strategy c has an advantage over a.

2.5.2.1 Simulations for $n > 2$

In Section 2.4.6, I hypothesized that RwS's behavior pattern, namely to be risk-averse when it has done well and risk-seeking when it has done badly, may be advantageous in lottery games with $n > 2$. To test this hypothesis, I ran experiments for the case $n = 4$, using the six strategies shown in Table 2.6.

For my simulations I used an initial population of 1000 agents of each type, and the parameters $p = 0.5$ for the risky lottery and $\alpha = 0$ for the population dynamic. The results were qualitatively the same as the ones in Figure 2.7: the

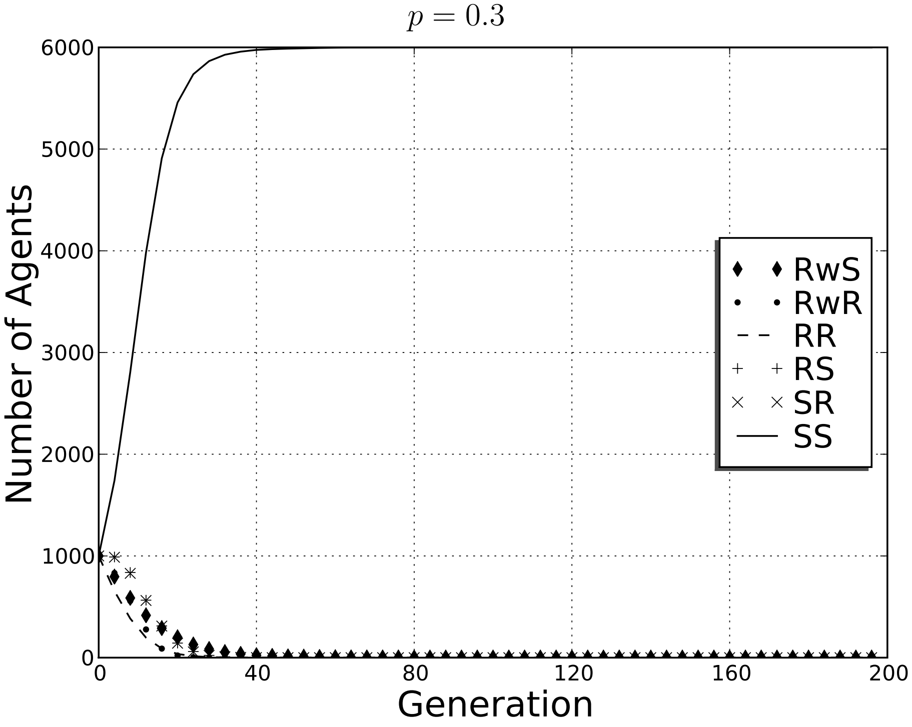

Figure 2.8: Results of simulation with $p = 0.3$ for the risky lottery and $\alpha = 0$. Plot shows the number of agents of each type over the course of evolution for 500 generations.

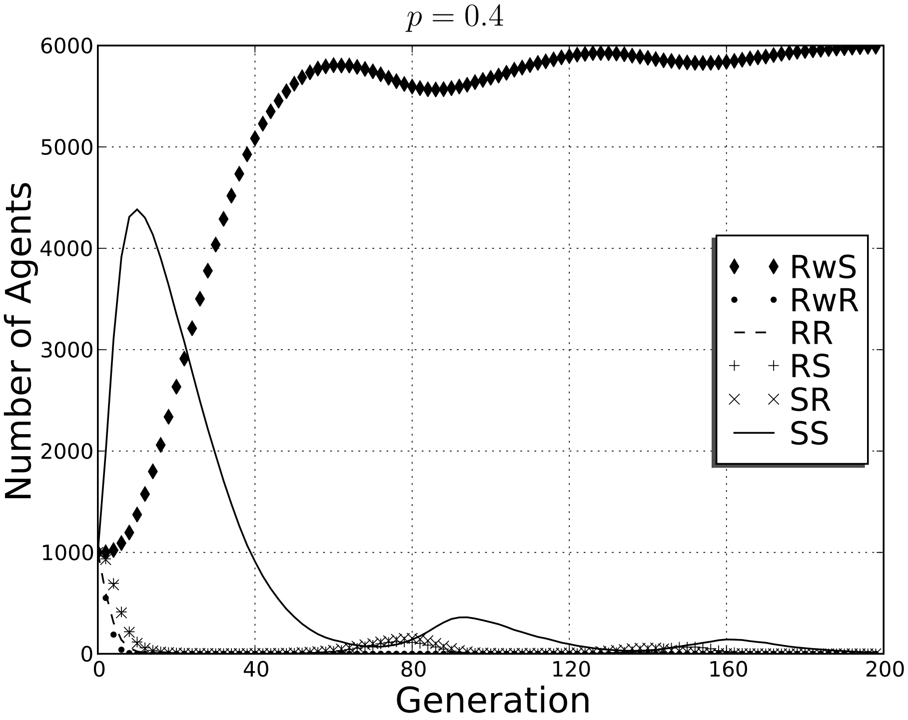

Figure 2.9: Results of simulation with $p = 0.4$ for the risky lottery and $\alpha = 0$. Plot shows the number of agents of each type over the course of evolution for 500 generations.

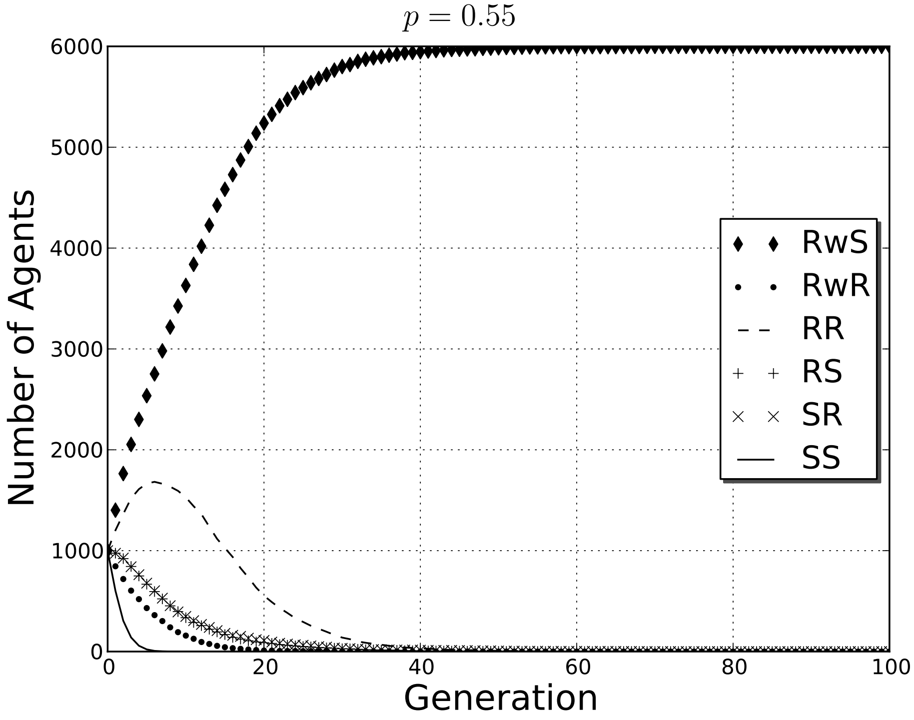

Figure 2.10: Results of simulation with $p = 0.55$ for the risky lottery and $\alpha = 0$. Plot shows the number of agents of each type over the course of evolution for 500 generations.

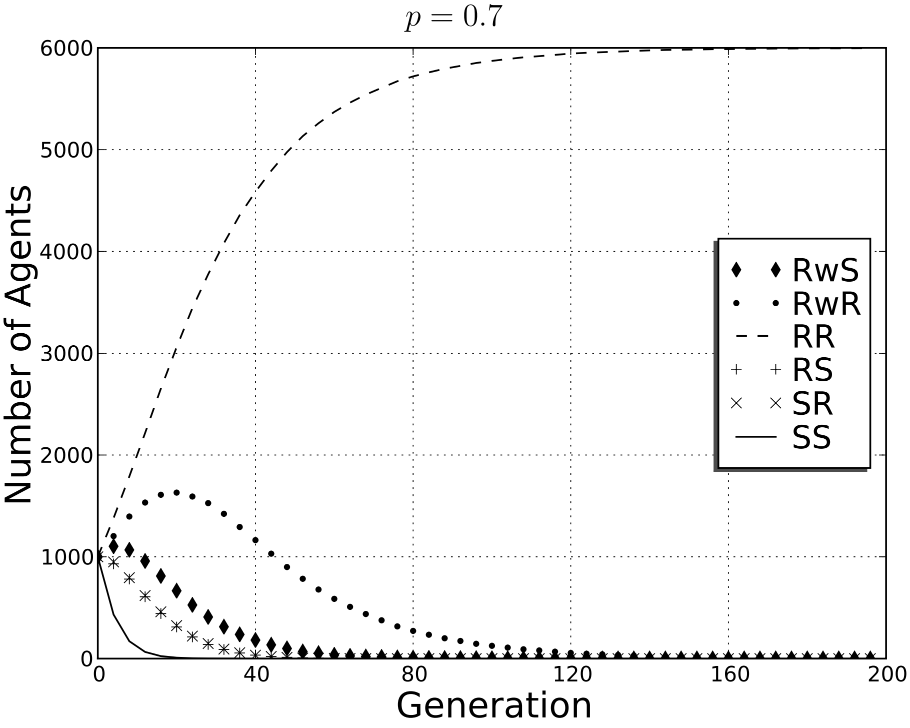

Figure 2.11: Results of simulation with $p = 0.7$ for the risky lottery and $\alpha = 0$. Plot shows the number of agents of each type over the course of evolution for 500 generations.

Table 2.6: Pure strategies in the lottery game simulations for $n = 4$.

Strategy	Behavior
SSSS	always choose safe
RRRR	always choose risky
SSRR	choose safe in lotteries 1 & 2, then choose risky in 3 & 4
RRSS	choose risky in lotteries 1 & 2, then choose safe in 3 & 4
RRwSS	choose risky in lotteries 1 & 2,
	then choose safe in 3 & 4 only if 1 & 2 were won, else choose risky
RRwRR	choose risky in lotteries 1 & 2,
	then choose risky in 3 & 4 only if 1 & 2 were lost, else choose safe

RRwSS strategy dominated the other strategies and grew to comprise 100% of the population. This would seem to confirm my hypothesis—but since there are hundreds of pure strategies when $n = 4$ and I only looked at six of the simpler ones, the result should be regarded as preliminary.

2.6 Relations to Alternative Decision Making Models

The manner in which the *RwS* strategy deviates from expected value maximization in the lottery games 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 [60, 118] 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 [67], 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 [54]. 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, [30] 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 [88] 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 the presented simulations.

The sequentiality of choices in my 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 2.4. 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, I believe that a tendency towards modeling games in normal form often leads people to overlook the impact of sequentiality on risk-related behavior.

I believe my 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),[6] 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.

2.7 Discussion

This chapter explored risk behavior of agents through analysis and simulation of evolutionary lottery games. The results demonstrated how the interplay between sequentiality of choice and population dynamics can affect decision making under

---

[6] 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. [79]) is implemented as a payoff-dependent threshold function, one 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.

risk. Specifically, for any imitation dynamics other than the pure replicator dynamic, there are evolutionary game environments in which the RwS strategy has an evolutionary advantage over expected-value maximization. Since RwS's risk-taking behavior is similar to the risk preferences captured in several prominent models of human decision making, this suggests that population dynamics other than the replicator dynamic may model an important mechanism for the emergence of those risk preferences. The *RwS* strategy exhibits behavior that is 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.

2.7.1 Limitations and Avenues for 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. The vast majority of current literature of evolutionary game-theoretic approaches considers situations of non-sequential choice, even though it seems clear that in many domains of interest sequential choice and resulting diversity in strategies exist. The presented lottery game simulations and evolutionary game analyses are a first step in exploring evolutionary mechanisms which can induce behavioral traits resembling those described in popular descriptive models of decision making. In general, there is much more opportunity for future work to use evolutionary game approaches 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.

Some specific ways in which this work could be extended are listed below:

- It is important to examine other population dynamics in which a strategy's reproductive success is not always proportional to its expected payoff. For example, if a death rate (e.g. [79]) is implemented as a payoff-dependent threshold function, one might expect risk propensities to differ depending on whether an agent is above or below that threshold, in a manner similar to behavior above or below an aspiration level in SP/A theory. 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.
- My study focused primarily on the case $n = 2$, i.e., in each generation the agents made two decisions. It should hold that state-dependent risk prefer-

ences like those of RwS should also have an advantage when $n > 2$, and Section 2.5.2.1 discussed some pilot experiments that support this intuition. Conducting more extensive studies may be an worthwhile topic for future work. Also, empirical studies that seek to estimate the rate (in terms of number of choices made) at which humans tend to update their strategy would seem important in determining what specific type of state-dependent risk behavior is to be expected in what certain environments or contexts.

- This work generally assumed a well-mixed population in which every agent was able and equally likely to imitate any other. It would be interesting to explore the possible effects of social or physical structures (that may guide or constrain imitation) on the evolution of risk behavior. Along these lines, see Section 3.5.1.

- Finally, this work has highlighted the need conduct more specific empirical studies of which type of imitation dynamics best models human imitation propensities under which conditions. The insights of this paper combined with such knowledge have potential application in any domain where human decision making under risk is of interest.

Chapter 3

Risk Behavior in Games of Interaction

The evolutionary lottery games investigated in Chapter 2 did not involve interactions between agents in the stage game, which allowed for the analysis of risk behavior in an isolated manner. This Chapter shows how the principle observations from the lottery game investigations apply to a popular social dilemma game of safety and cooperation. I consider an evolutionary game in which agents play two sequential stag hunt games in a generation. Like the prisoner's dilemma [5], the stag hunt is a game that models a dilemma between cooperation and noncooperation. I demonstrate how a strategy essentially equivalent to the RwS strategy in the 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. [108].)

3.1 Stag Hunt Environment

The stag hunt environment I consider is equivalent to the 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 I use for the stag hunt game is shown in Table 3.1.[1] 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, I will call this an evolutionary *double stag hunt* game.

Table 3.1: Payoff matrix used in the stag hunt game. The payoff values are chosen as to coincide with the lottery games, but keep the relevant payoff relations of the stag hunt.

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

3.2 Risk and Strategies

A significant difference between the stag hunt environment and the lottery game environment is that in the former payoffs are not stochastic due to probabilities on the payoffs themselves, but due to the probabilities of playing against a *stag* agent

---

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

(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. One 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 the lottery game results in this Chapter, 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 the two-choice lottery game.

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

3.3 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 3.2 lists these payoff distributions achieved by agents in a population consisting of *hare* and *stag* agents in the double stag hunt game environment.

Table 3.2: 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	*hare*	*stag*		
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 3.2, we get:

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

Since we are 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 3.1 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

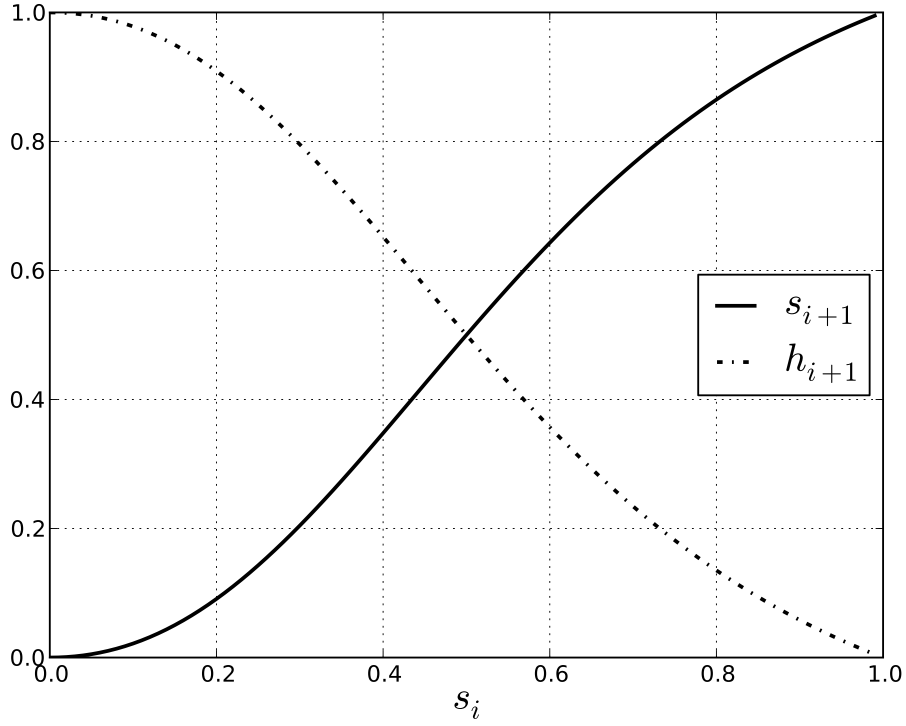

Figure 3.1: Plot of s_{i+1}, the proportion of *stag* agents in generation $i + 1$, and h_{i+1}, the proportion of *hare* agents in generation $i + 1$ against s_i, the proportion of *stag* agents in generation i, under replicator dynamics.

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* agents.

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

P(*stag* vs. *stag*) = s^2, whence *stag* reproduces.

P(*hare* vs. *hare*) = $(1-s)^2$, whence *hare* reproduces.

P(*stag* vs. *hare*) = $2s(1-s)^2$, whence the agent with the higher payoff reproduces, or a random agent if payoffs are equal.

We can calculate s_{i+1} under imitate-the-better by combining these pairing probabilities and the payoff distribution information from Table 3.2. Doing so gives:

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

Figure 3.2 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.

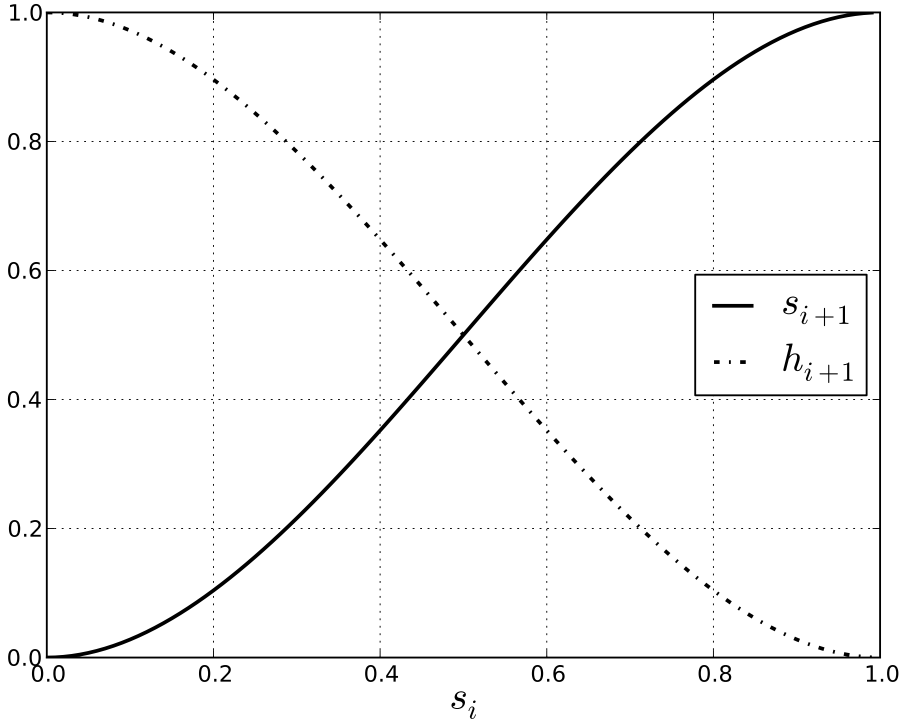

Figure 3.2: Plot of s_{i+1}, the proportion of *stag* agents in generation $i+1$, and h_{i+1}, the proportion of *hare* agents in generation $i+1$ against s_i, the proportion of *stag* agents in generation i, under imitate-the-better.

Hence I 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.

3.4 Simulations and Results

My first set of stag hunt simulation experiments above serve 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, I run a large number of simulation runs and count the amount of times the population is entirely taken over by either agent type. Figure 3.3 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. Observe 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*).

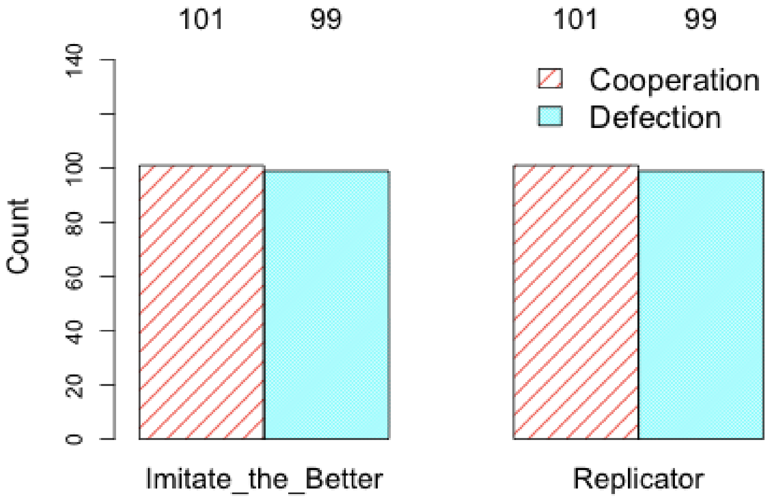

Figure 3.3: Simulation results for an initial population of 3000 *hare* and 3000 *stag* agents. The plot shows the count of simulations in which the population resulted in all *stag* agents (cooperation) and all *hare* agents (defection) for 200 simulations under each imitate-the-better and replicator dynamics.

I have hypothesized in Section 6.2 that in a population of 50% *stag* and 50%

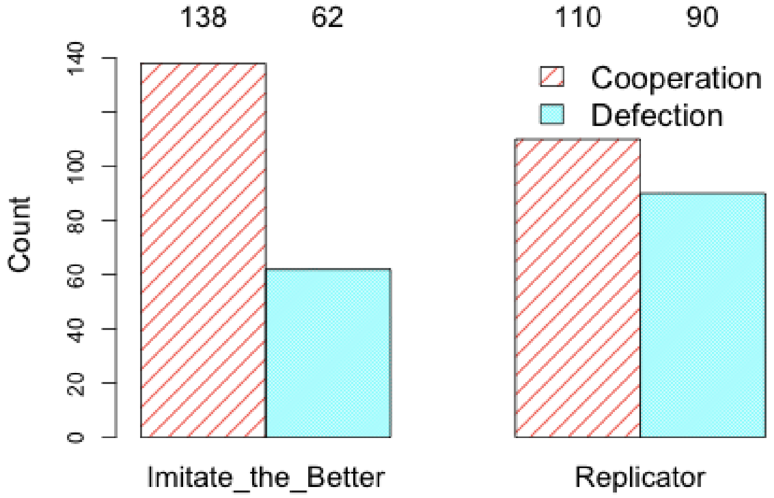

Figure 3.4: Simulation results for an initial population of 3000 *hare* and 3000 *stag* agents and 30 *RwS* agents. The plot shows the count of simulations in which the population resulted in all *stag* agents (cooperation) and all *hare* agents (defection) for 200 simulations under each imitate-the-better and replicator dynamics.

hare players, given the payoff matrix in Table 3.1, the *RwS* agent in the 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 my earlier lottery games. My second set of experiments serves to verify this hypothesis and investigate the impact it has on population evolution. For this set of simulations, I used an initial population of 3000 *stag*, 3000 *hare*, and a small amount (30) of *RwS* agents. I 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 the simulations again lead to the population evolving to complete cooperation (100% *stag*)or complete defection (100% *hare*).

Figure 3.4 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 ($X^2 = 7.7356$).

3.4.1 RWS as Catalyst for the Evolution of Cooperation

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 the 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 not 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.

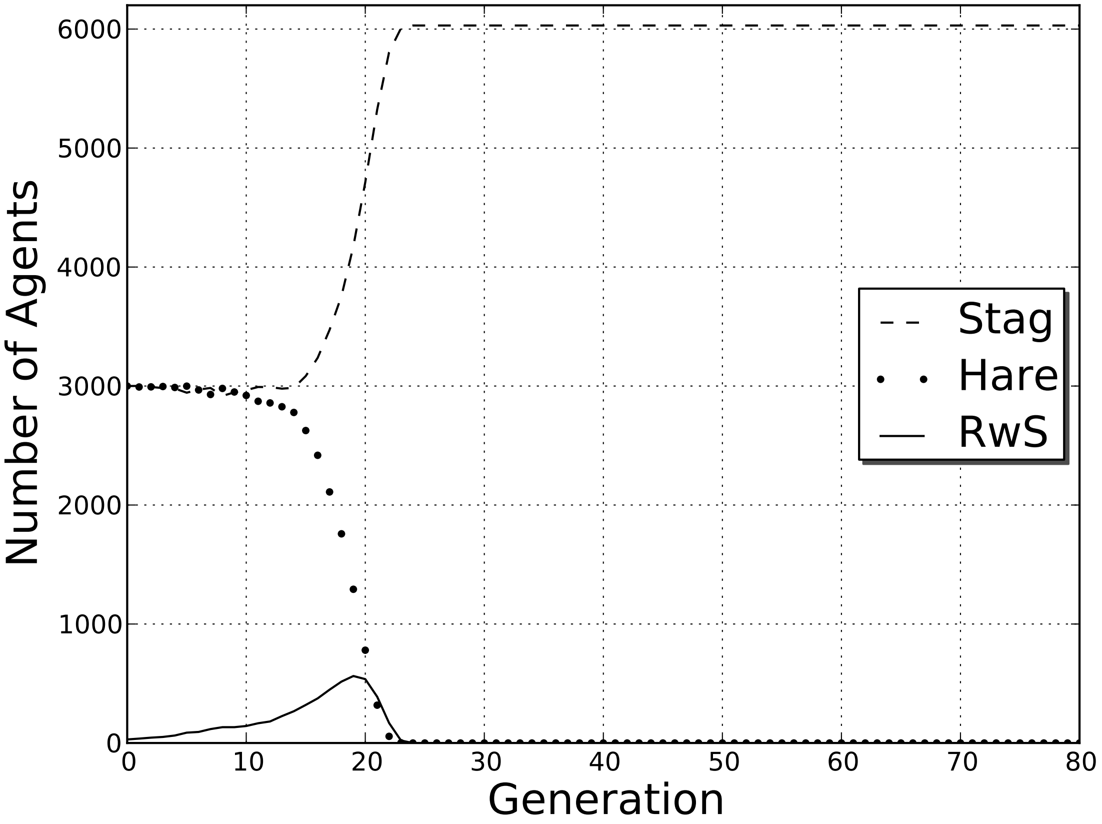

Figure 3.5: Agent type frequencies for a typical stag hunt simulation run under imitate-the-better in which *RwS* agents grew and boost stag players, leading them to take over the entire population

Figure 3.5 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. One can 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 the 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.

3.5 Discussion

The evolutionary stag hunt game investigations in this Chapter demonstrated how the results from the lottery games of Chapter 2 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). Hence, the existence of state-dependent risk preferences in conjunction with an imitate-the-better social learning dynamic increased the likelihood that cooperation emerged as a norm in the population. I 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. Simulation source code and result data used for this Chapter are made available for download online: http://www.cs.umd.edu/users/roos/materials/ACS2010.html.

3.5.1 Risk in Games on Graphs

In the evolutionary stag hunt game presented, the risk involved in a *hunt stag* or *hunt hare* action stemmed from the likelihood that the opponent would pick either action. Since to demonstrate the main points here I only considered well-mixed populations here, in which every agent was equally likely to play any other agent in the population, it may seem very unlikely that the distribution of agents would be so diverse in a population as to produce a significant variance in the risky choice and a resulting significance of RwS agents. This is because, as demonstrated, any population would rather quickly arrive at an all *hare* or all *stag* agent equilibrium, and only a relatively small proportion of different agents would be entering the population through mutation or exploration dynamics. However, in evolutionary games on graphs (e.g. [20, 82, 96, 97, 22]), where the interactions between agents is structured on a graph in which agents are nodes and play their neighbors, it is much more likely that agents would face a more diverse group of opponents. Agents generally have a much lower number of neighbors to possibly interact with, and thus rare mutations or explorations would have a much greater effect on (local) strategy diversity and resulting risky-choice variance.

3.5.2 Winner-Takes-All Interpretation of Imitate-the-Better

Both this and the previous Chapter have considered the imitate-the-better dynamic as a type of social learning process of imitation, but it is worth noting that this population dynamic has another analogy, giving it a broader scope of applicability. This alternate analogy is that of winner-take-all games [121, 29, 7]. In a winner-take-all game, the agent that acquires the higher payoff of two agents reaps all the benefits, meaning, in evolutionary terms, that the agent with higher acquired payoff is the only agent that replicates. Real life examples of winner-take-all situations are patent races or various forms of price competition, and it has been argued that many species' mating process is winner-takes-all, i.e. a "winner-takes-all game determines reproductive success" (the more well-off agent gets to mate) [121, 7]. This suggests that my results on risk-taking under the imitate-the-better dynamic has applicability to a broader spectrum of situations and environments than those in which social learning governs the reproductive dynamics.

Chapter 4

Societal Threat and the Evolution of Punishment Propensities Across Cultural Groups

In this chapter I employ evolutionary game theoretic approaches to model and investigate the evolution of human behaviors. This work specifically considers the evolution of cultural groups, with the goal to understand and explain differences in empirically observed characteristics between them. The specific cultural characteristic explored is the willingness of individuals to punish others for norm-violating or non-cooperative behavior.

There are striking differences across cultural groups in their willingness to punish norm violators. However the conditions under which punishment of norm violators is more adaptive or less for human societies and whether such differences have an evolutionary basis has received little attention. To fill this void, I propose that punishment propensities vary across groups at least partially as a function of the degree of threat to which societies are exposed, because different punishment propensities are optimal for group survival or stable under different threat conditions. Whether there exists such a causal relationship between societal threat and adapted punishment propensities is difficult, if not impossible, to test with laboratory or field studies. But this question is well-suited to evolutionary game theoretic

(EGT) models, since such models allow for the observation of the effects of varying threat levels on the evolution of punishment propensities.

A considerable number of studies have studied the evolution of punishment in conjunction with cooperation through the Public Goods Game, a popular social dilemma metaphor in which groups of individuals interact under a choice to *Cooperate* by contributing or to *Defect* by withholding contribution to a public good. To date, a primary concern of these studies is how different forms of punishment can arise and how they aid the evolution of punishment; none of these models, however, have considered the question of how different punishment propensities may evolve across groups.

In this chapter, I analyze the effects of varying degrees of different types of societal threats through EGT models based on the Public Goods Game, and show through mathematical analysis and computer simulations how differences in punishment propensities can arise from differences in the group's exposure to societal threats. I consider the evolution of punishment under two reasonable evolutionary PGG models that differ in strategy set choice and the manner in which punishment propensity is modeled. I focus on whether there is an evolutionary advantage (or disadvantage) that different punishment propensities give a population. Thus I first use a basic model which simply assumes that different groups or populations have a way of maintaining a certain probability with which Cooperating members punish Defectors. I consider this probability the group's *punishment propensity* and can examine effects on evolutionary viability of a population under different circumstances by varying it. I refer to this model as the *Basic Model.* I also seek to replicate these effects in another, more complex PGG model with punishment. Recently, [52] proposed a model that shows how responsible punishment can evolve in a population model that allows for anti-social and spiteful punishment while simultaneously avoiding the problem of higher order free-riding (cooperative outcomes being endangered by Cooperators that do not punish Defectors invading the population). Hence, I also consider this more expanded model, which includes the state-of-the art strategy set including 16 strategies and a form of reputation proposed by [52]. I refer to this model the *Hilbe and Traulsen Model*. Since in this model agents either punish or do not punish, I consider a population's *punishment propensity* to be the proportion of punishing agents in the population. The consideration of these two alternative models in this chapter also serves to illustrate the robustness of the general results presented.

Social scientists have identified a number of societal threats cultural groups might face to widely varying degrees, including external man-made threats such as invasions or warfare threatening a society's own territory and ecological threats such as natural disasters. Operationalizing these threats in the evolutionary game models and investigating their impact on evolutionary dynamics shows that increased levels of these threats lead to increased punishment propensities. In the basic model this occurs because under increased threat, increased punishment is required to maintain high cooperation rates, providing high overall group payoffs. Hence to increase chances for survival under cultural group selection [48, 13, 114], societies facing higher degrees of societal threat require higher punishment propensities. However, since punishment is costly, punishment propensities too far above the required amount can reduce overall group payoff. Thus group selection favors groups that maintain optimal propensities only slightly higher than what is needed to sustain cooperation under their particular threat conditions. Interestingly, in the Hilbe and Traulsen Model, no group selective pressures are needed for the differences in punishment propensities to arise from differences in societal threat. In this model, assuming a game parameter range under which populations can uphold cooperation, a mix of punishers and non-punishers is evolutionarily stable within the population, and this mix consists of a higher proportion of punishers under higher degrees of threat. This work provides a critical insight into the study of cultural variation in punishment by showing how populations that have higher societal threat evolve higher punishment propensities, and that this relationship is of a causal rather than merely correlational nature.

The remainder of this chapter is organized as follows. In the next Section, I provide some more background on social science work concerning punishment and motivation for this research. Section 4.2 describes the general evolutionary Public Goods Game with punishment and the general population dynamics that specify the change in strategy frequencies within a group or population used throughout the models. Section 4.3 discusses the societal threats that I will operationalize in the models. Sections 4.4 and 4.5 then respectively describe the Basic Model and the Hilbe and Traulsen Model in detail and show how in both models the relationship between societal threats and punishment support my hypothesis. Finally, Section 4.6 concludes this Chapter with a discussion of the results.

4.1 Punishment Propensities in Humans

There is large amount of evidence that a willingness to engage in (costly) punishment exists in humans [36, 35, 86, 91, 42, 28, 41, 23, 72, 51] and that punishment of Defectors can play a vital role in the evolution of cooperation [18, 12, 16, 20, 21, 104, 49, 45, 48, 47, 113, 44, 123]. Yet there are striking differences across cultural groups in their willingness to punish people who violate social norms [51, 68, 50]. Early anthropological research showed that traditional societies have differing punishment propensities for norm violations: some groups (e.g., the Hutterites and Hanno) have severe punishment, and others (e.g., the Kung Bushman, Cubeo) exhibit much greater permissiveness [89]. Recent evidence shows that modern cultures also vary widely in their punishment of norm-deviating behavior [51, 68, 50].

More recently, in research across 33 nations, [39] placed cultures around the world on a psychological dimension scale of “tight” vs“loose”. They showed that the “tightness”, i.e. strength of social norms and punishment of deviations from them, of human cultural groups or populations is related to a broad array of ecological and human-made societal threats (or lack thereof) that nations have historically encountered. They argued that a high degree of threat increases the need for strong punishment systems to facilitate the coordination necessary for survival. Nations with few ecological and human-made threats, by contrast, have a much lower need for order and social coordination, affording weaker social norms and much lower punishment of deviant behavior. Societies that have had high degrees of territorial threats (from 1918-2001), low natural resources (e.g., food supply and water resources), and high degrees of natural disasters (e.g., floods, cyclones, and droughts) and human disease (pathogen prevalence) had much stronger norms and punishment of deviance than societies that had low levels of these threats. The correlational nature of this research, however, leaves open the question of whether groups actually require stronger punishments to survive under high threat, and more generally whether differences in punishment across societies has any evolutionary basis.

4.2 Evolutionary Public Goods Game (PGG)

The Public Goods Game (PGG) is a well-established paradigm for studying cooperation and norm violation [45, 20, 21, 49, 48, 113, 44], and it has also been used to study punishment [20, 21, 113, 123]. In the PGG, N players may each either *Cooperate* (contribute some amount c) or *Defect* (contribute nothing). If more than one player contributes, the sum of all contributions is multiplied by a factor r. The resulting amount is divided evenly among all players, regardless of whether they contributed. The higher the proportion of Defectors, the less there is to share. Due to the temptation to defect, it would be easy for the entire population to fall into a state of all Defectors. However, several studies have shown how social mechanisms such as individual-based peer-punishment or institutional punishment can foster and establish cooperative behavior as a societal norm [18, 16, 20, 21, 49, 113, 123]. Under peer-punishment, if a Cooperator decides to punish a Defector, punishing reduces the Defector's payoff by ρ, at a cost λ to the Cooperator.

Following established work on PGG models [45, 113, 123], each generation, the model samples multiple disjoint game-groups from the population to play.[1] Each game-group plays a PGG, then Cooperators have a chance to punish Defectors in their game-group. After num_games samplings and PGGs played, the population changes under a combination of total-payoff-proportional imitation and random exploration of strategies. Payoff-proportional imitation can be viewed as a process of social learning in which agents imitate other agent's strategy with a probability that increases with the agent's payoff. This process is commonly modeled in infinite population models through the replicator dynamics [40, 53]. Alternatively, the process is often modeled in finite populations using the Fermi Rule [11, 110, 115, 52]: an agent a imitates (switches to) a randomly chosen other agent b's strategy with a probability $p_{a \to b} = 1/(1 + e^{-s(\pi_a - \pi_b)})$, where π_a and π_b are the total payoffs of a and b respectively, and $s \geq 0$ is the selection or imitation strength. I assume all agents update their strategy in such a way simultaneously each generation. Random exploration (i.e. exploration dynamics) of strategies are analogous to random mutation. Such random exploration of the available strategy space has recently been shown to play an important and often underestimated role in human strategy updating within social contexts [113, 117].

---

[1] I shall use the term "game-group" to refer to any group of agents playing a PGG within a population, while I more generally use the term "group" as synonymous to "population" or "society" when I refer to group selection.

4.3 Societal Threats

Social scientists have identified a number of societal threats cultural groups might face to widely varying degrees. These include ecological threats such as natural disasters and external man-made threats that threaten a society's territory (e.g. invasions or warfare) [39]. The degree of tightness (strength of social norms and punishment of deviations from them) of different cultures was found to correlate positively with the cultures' exposure to such threats. In the following subsections I discuss these categories of threat and describe how they are operationalized in the PGG models.

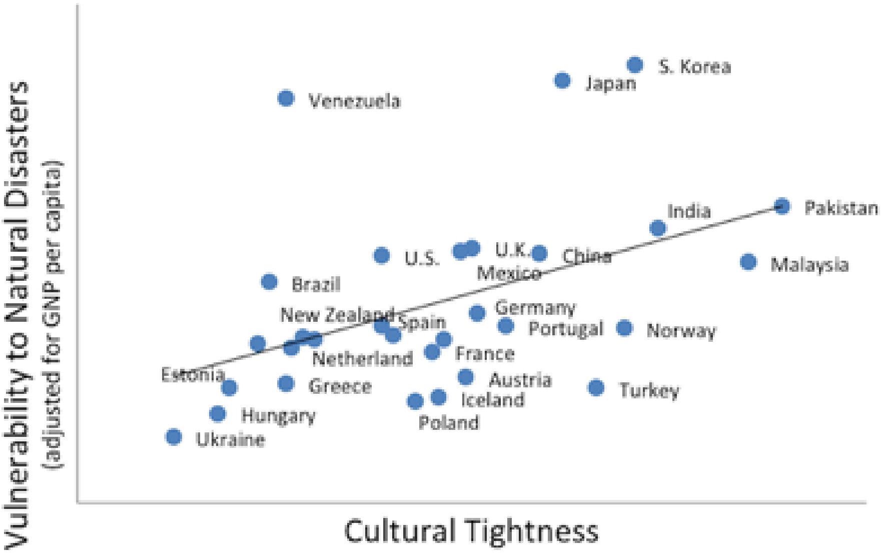

Figure 4.1: Correlation between cultures' exposure to natural disasters, a type of ecological threat, and cultural tightness.

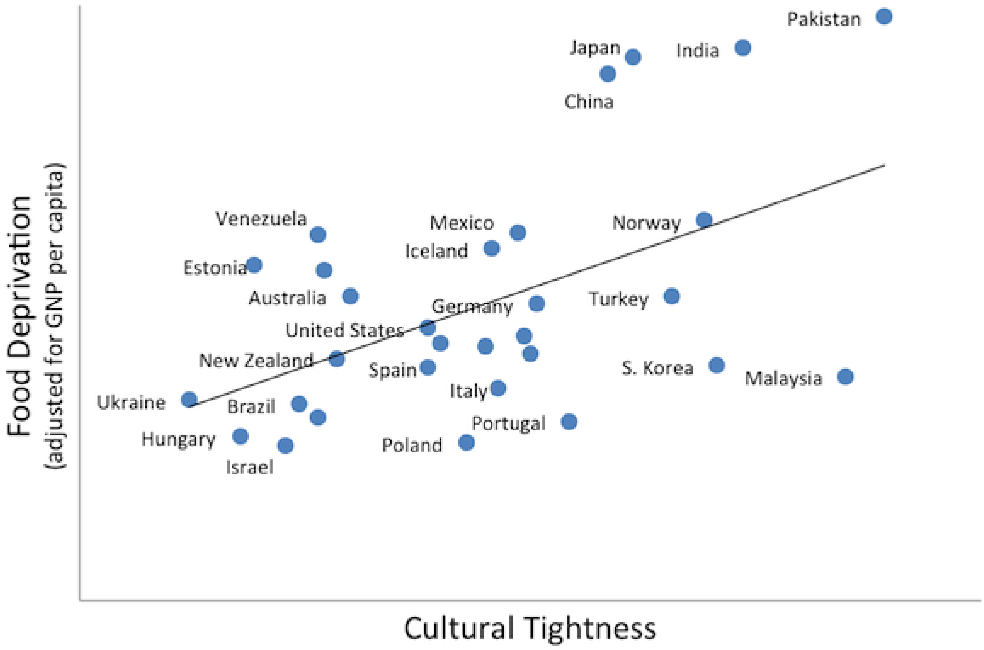

Figure 4.2: Correlation between cultures' exposure to food depravation, a type of ecological threat, and cultural tightness.

4.3.1 Ecological Threats

Nations that face ecological threats such as floods, tropical cyclones, droughts, or higher prevalence of pathogens, have been found to tend to stronger norms and punishment systems [39]. See e.g., Figures 4.1 and 4.2, which show correlations between more natural disasters ($r = .47, p = .01$) and higher food deprivation ($r = .52, p < .001$) with the degree of tightness of cultures. Presumably, having strong norms and punishment of deviants enables such groups to coordinate social action and survive in the face of such severe threats. Similarly, nations that have a high prevalence of pathogens, particularly those that are highly communicable (e.g., tuberculosis), require stricter rules in order to avoid contamination and ultimately

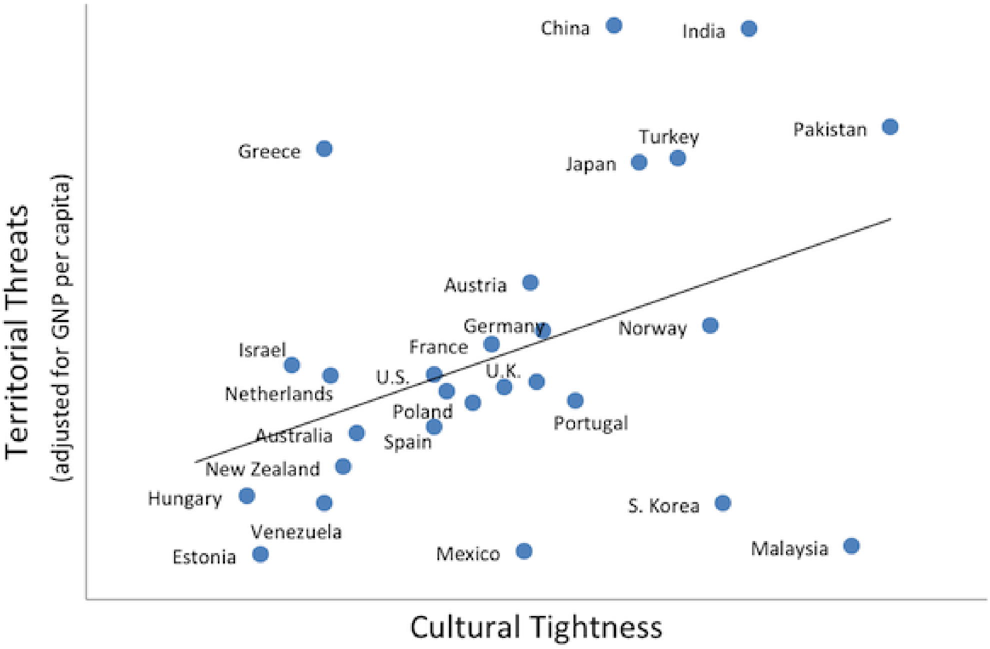

Figure 4.3: Correlation between cultures' exposure to territorial threats from their neighbors, an external man-made threat, and cultural tightness.

enhance survival. In their field research, [39] found historical prevalence of pathogens to be higher in tight nations as were the number of years of life lost to communicable diseases, the prevalence of tuberculosis, and infant and child mortality rates. By contrast, societies that do not face natural disaster threats can afford to have a fewer rules and weaker punishment systems.

All of these threats may lead to inefficiencies in production, or managing them may require the use of the population's resources. Natural disasters are also related to the availability of natural resources in that they often diminish agricultural yields and engender food shortages [90]. Hence, a straightforward way to operationalize ecological threats in the PGG models is by decreasing the payoff to the group members in general. One can do this by varying the parameter r, the multiplication factor of contributions creating the public good to be divided among agents in both the Basic Model and the Hilbe and Traulsen Model.

4.3.2 External Man-Made Threats

External man-made threats that threaten a society's territory include, e.g., migration, intentional sabotage, and territorial invasion. [39]'s field research found that societies facing potential invasions from neighboring groups (and by extension, facing challenges to their group resources) developed stronger punishment systems than societies that had few territorial threats. See Figure 4.3, which shows the correlation between cultures' exposure to territorial threats from their neighbors during the period 1918-2001 and the degree of tightness of the culture ($r = .41, p = .04$). One way to operationalize this type of threat is to reduce a population's overall payoff, as above, since resources that would otherwise go to the population must instead be used to fend off threats. Another reasonable alternative is to model this type of threat as an invasion of Defectors: by taking some of the group's payoff without contributing, Defectors in effect steal from the group: they decrease the per-capita payoff, hurting group survivability. The threat's intensity can be interpreted as the size of the invasion of Defectors.

4.4 Basic Model

This section describes the strategy sets used in the Basic Model, and show how different degrees of threats in this model relate to punishment. Since the strategy set for the Basic Model is relatively simple, I can provide mathematical analysis of the theoretical infinite population dynamics to provide a basic intuitive insight into the dynamics between cooperation, defection, and punishment in the PGG. I represent the basic population dynamics mathematically through a system of ordinary differential equations ODEs that is solvable for the effects of different model parameters on evolutionary outcomes, and I shall use this system of ODEs to explore the effects of different threats.

4.4.1 Strategy Set

The strategy set composing the Basic Model simply consists of Cooperators and Defectors. As described in the Introduction, the main objective of this work is to explore the relationships among various types of societal threat, punishment propensity, and evolutionary outcomes. Hence I model a *punishment propensity* as a probability q with which a Cooperator in any particular group punishes Defectors. Modeling punishment in this sense allows not only for the representation of different punishment propensities in different cultures, but, as I shall show, it also allows us to describe optimal punishment propensity values (in terms of overall group payoff) under different conditions of societal threat. While assuming the existence of such a general punishment propensity within a population is not standard in

EGT modeling, I believe that empirical evidence showing differences in punishment propensity across populations supports this assumption, and I note that there are various mechanisms (e.g. honor, reputation, or conformist transmission) through which a society may uphold a certain propensity to punish among its members. If a certain degree of punishment propensity is necessary for a society to reach cooperative, high-payoff outcomes and hence survive (on its own or under group selective pressures), I believe it is reasonable to assume that societies that do survive have found or developed some mechanism to uphold such a degree of punishment.

4.4.2 Analysis

For the theoretical infinite populations case of the Basic Model, the replicator dynamics with mutation described below (Eqn. 4.2) combined with the equations giving the expected payoffs for agent types (shown in Eqns. 4.4 and 4.5) give us a parameterized system of ODEs solvable numerically for the change in agent type frequencies over time under different conditions. Under replicator dynamics we consider payoffs achieved by agents as analogous to the agent's fitness, i.e., an agent's probability of reproducing is directly proportional to these payoffs [40, 53].The change in population proportions according to the replicator dynamics is thus given by the following ODE:

$$\dot{x_i} = x_i \, [\, \pi_i(x) - \theta(x) \,], \quad \theta(x) = \sum_i x_i \, \pi_i(x), \tag{4.1}$$

where x_i is the proportion of agents of type i in the population, $\pi_i(x)$ is the expected payoff an agent of type i, and $\theta(x)$ is the expected payoff of all agents in the population. An agent's type is simply the strategy it employs to make choices. I shall use x_C and x_D to denote the proportion of Cooperators and Defectors in the population. To also include random exploration of strategies in the population dynamics, we can use the replicator dynamics with mutation (i.e. replicator-mutator equation):

$$\dot{x}_i = \sum_j x_j \pi_j(x) Q_{ji} - \theta(x) x_i, \tag{4.2}$$

where the matrix Q gives the mutation probabilities from one agent type to another. With a mutation rate μ and n strategy types $Q_{ij} = (1 - \mu/n)$ if $i = j$ and $Q_{ij} = \mu/(n-1)$ if $i \neq j$.

Since in the model game-groups of size N are sampled at random from the population to play PGGs, we need to calculate the average (expected) payoffs to each agent type by considering the probabilities of group compositions. Following this approach, for any given agent sampled to play a PGG in a game-group of size N, the probability that k of its co-players are Cooperators is

$$x_C^k \, x_D^{N-1-k} \binom{N-1}{k}. \tag{4.3}$$

With k cooperating co-players and $N - k$ Defectors, the payoff to each Defector is rk/N from the common good, minus, $kq\rho$, since Defectors are punished by Cooperators in the group with probability q. Thus, considering group composition probabilities, and simplifying for $x_C + x_D = 1$, the expected payoff to Defectors is

$$\begin{aligned} \pi_D &= \sum_{k=0}^{N-1} (\frac{r}{N}k - k\,q\,\rho)\, x_C^k \, x_D^{N-1-k} \binom{N-1}{k} \\ &= \left(\frac{r}{N} - q\,\rho\right)(N-1)\,x_C. \end{aligned} \tag{4.4}$$

In a group of k Cooperators and $N-k$ Defectors as co-players, the payoff to a Cooperator is $r(k+1)/N-c$ from the common good ($k+1$ because the player itself cooperates), minus $q\,\lambda\,(N-k-1)$ due to the cost of punishing Defectors in the group with probability q (there are $N-k-1$ Defectors in the group and each is punished with likelihood q). Thus, again considering all group compositions and their probabilities, the overall expected payoff to Cooperators is

$$\begin{aligned}\pi_C &= -c+\sum_{k=0}^{N-1}\left[\frac{r}{N}(k+1)-q\,\lambda\,(N-k-1)\,\right]x_C^k\,x_D^{N-1-k}\binom{N-1}{k}\\ &= rx_C+\frac{rx_D}{N}-c-q\lambda(N-1)x_D. \qquad (4.5)\end{aligned}$$

A population under replicator-mutator dynamics and the strategies and payoffs above can maintain a high level of cooperation as long as there exists a large enough punishment propensity q among the Cooperators relative to the number of Defectors currently in or entering the population. As an example, for a population consisting of Cooperators in which Defectors are introduced by mutation $\mu=0.01$ and the game parameters $r=3, c=1, \lambda=1, \rho=1, N=5$, Cooperators must have a punishment propensity of $q\geq 0.13889$ to avoid a take-over by Defectors. More generally, if it is possible for Cooperators to withstand Defectors through punishment, there exists a value for the minimum punishment propensity required to withstand Defectors (dependent on the game parameters). I denote such a required punishment propensity q_{req}. If a population's punishment propensity $q<q_{req}$, Defectors take over the entire population, I refer to this as a *societal break-down.* However, if $q>q_{req}$, the population withstands the Defectors, maintains high cooperation rates, and hence high overall payoff. See Figure 4.4 for an illustration of the population dynamics in both cases. The population dynamics given by Eqn. 4.2, using the expected payoffs for player types (Eqn. 4.4 and 4.5), give a systems of ODEs that I can solve numerically for (q_{req}) under different game conditions (and hence different conditions of societal threat). The results in Figure 4.5 described in the Results section were derived this way.

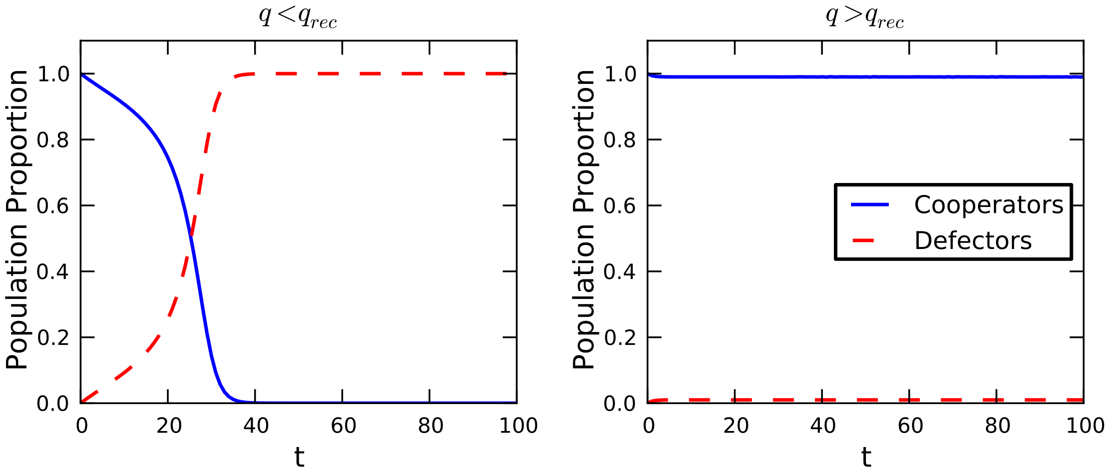

Figure 4.4: Evolution of strategy proportions of the Basic Model.The left graph shows a break-down to defection when Cooperators have a punishment propensity q lower than than that required to withstand take-over by Defectors ($q = 0.1 < q_{rec}$). The right shows maintenance of cooperation when q is high enough ($q = 0.5 > q_{rec}$). Game parameters $r = 3, c = 1, \lambda = 0.3, \rho = 0.7, N = 5$. Defectors are introduced by mutation $\mu = 0.01$.

4.4.3 Cultural Group Selection

The population dynamics described above determine the evolution of different individuals within groups or populations, but it is important to understand the implications of group selection in the interpretation of the results presented. Group selection has been argued to play an important role in cultural evolution and promotion of cooperation [106, 37, 48, 13, 114, 14]. Generally speaking, group selection is an additional level of selection (apart from selection at the individual level) that selects for groups based on their overall group payoff . The most straightforward group selection mechanism in human groups is direct group competition and empirical studies have shown that selective pressures through inter-group lethal competition (warfare) have been strong enough to account for the selection of altruistic behavior in human groups [13]. Another mechanisms of cultural group selection are selective intergroup migration, for which substantial literature exists supporting that migrants flow from societies where immigrants find their prospects poor to ones where they perceive them to be better, and most immigrant populations assimilate to the host culture within a few generations [19]. Another form of group selection recognized is inter-group cultural transmission [19]. Group selection leads groups which are able to maintain higher overall payoff (i.e. group fitness) to have an evolutionary advantage (i.e. higher likelihood to survive) over competing groups. This is important for the understanding of the results presented in the following Section because these show how higher degrees of threats require higher optimal punishment propensities to maintain high group cooperation. Because high rates of cooperation lead to higher overall group payoff, high rates of cooperation are linked directly to group survival.

4.4.4 Results

The operationalization of societal threats in the Basic Model support the hypothesis that higher threats lead to higher punishment propensities in populations. In general, the relationship between threat and punishment in the model is as follows: there is a minimum required punishment propensity q_{rec} that a population requires in order to maintain cooperation and thus to be evolutionarily viable; and q_{rec} increases monotonically with the amount of societal threat. Any $q > q_{rec}$ is neutrally stable (not considering group selection). However, I also find that there is an optimal punishment propensity q_{opt} slightly higher than q_{rec}. Punishment in excess of q_{opt} can harm overall group payoffs, hence would be selected against under group selection. Thus group selective pressures would favor groups that establish punishment propensities slightly above the minimum amount needed to prevent a societal break-down into defection, and this minimum amount of punishment is an increasing function of the degree of societal threat faced by the group:

- *Ecological Threats Increase the Required Punishment Propensity:* Solving the system of ODEs giving the population dynamics for q_{req} under different r parameters shows that a higher r (less societal threat) lessens the required punishment propensity to maintain cooperation, while a lower r (more societal threat) raises the required punishment propensity. See left graph of Figure 4.5.

- *External Man-Made Threats Increase the Required Punishment Propensity:* Solving the system of ODEs giving the population dynamics for q_{req} under different invasion sizes, i.e. different proportions of Defectors entering the

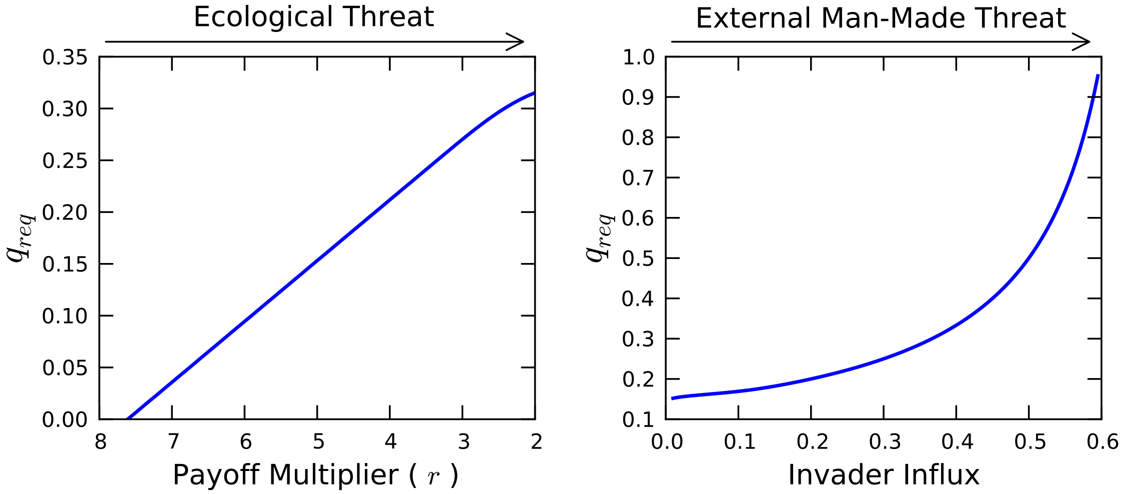

Figure 4.5: Left graph: Minimum punishment propensity required (q_{req}) to resist Defector take-over as a function of the game parameter r. Higher external threat that reduces overall payoff means a lower r value. Right graph: Minimum punishment propensity required (q_{req}) to resist Defector take-over as a function of the proportion of Defectors invading the population. Both computed by solving the system of ODEs given by the replicator dynamics of the PGG model for the minimum q needed to extinguish Defectors within $t = 100$. Game parameters used: with $r = 3, c = 1, \lambda = 0.3, \rho = 0.7, N = 5$.

population, shows that the greater the influx of Defectors, the higher the punishment propensity needed to prevent a break-down into defection. See right graph of Figure 4.5. Hence if a group is threatened in this fashion, it needs a higher punishment propensity against Defectors to maintain high rates of cooperation, high group payoff, and increased chances of group survival.

- *Excess Punishment is Not Optimal for Group Payoff:* The above results for the Basic Model show how increases in societal threats increase q_{req}, the minimum punishment propensity needed to maintain cooperation (hence high population payoffs), but as such do not show any reason for why any population would not simply keep a punishment propensity of $q = 1$ at all times. Finite population model simulations of the Basic Model also show (see Figure 4.6) that there exists an optimal punishment propensity, q_{opt}, slightly above q_{req}. These simulations were done using a straight-forward implementation of the described evolutionary PGG Model, see Appendix B for pseudo-code of the basic simulation sequence. As is evident from Figure 4.6, punishment propensities above q_{opt} decrease the overall population payoff. This is because a constant exploration rate leads to a continuous, unavoidable presence of Defectors, and punishing them is costly. This effect is especially dramatic when there is action perception noise - a nonzero probability κ that agents will misinterpret a cooperative action as a defection and hence punish mistakenly, or vice versa. Hence group selection selects against punishment propensities above q_{opt}. This principle applies to all threat types explored.

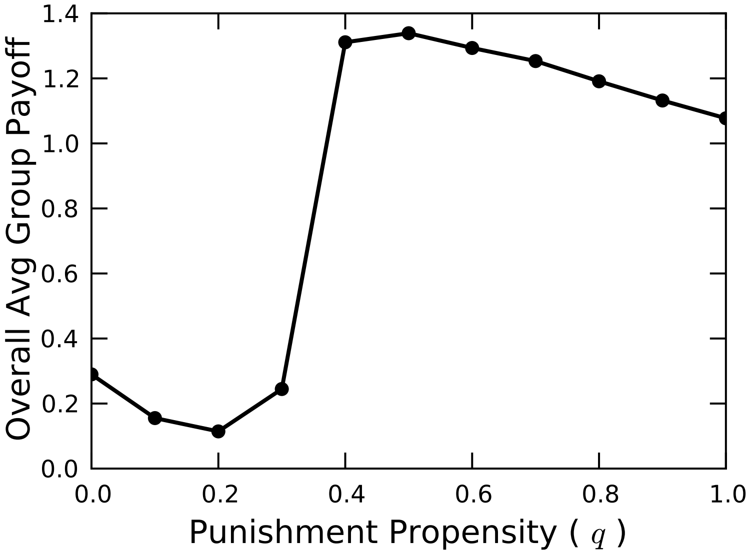

Figure 4.6: Overall group payoff θ vs. the punishment propensity q for a population of 500 agents in the Basic Model (with $r = 3, c = 1, l = 0.7, \rho = 0.7, N = 10, \mu = 0.01, \kappa = 0.05$). The punishment propensity q_{req} to withstand Defector takeover is $q_{req} \approx 0.4$. Note also that there is an optimal punishment propensity $q_{opt} \approx 0.5$ that maximizes the overall population payoff. A group that can maintain a punishment propensity closer to q_{opt} will do better under group selection than if it had a punishment propensity lower or higher than q_{opt}.

4.5 Hilbe and Traulsen Model with Reputation

In this section I describe the strategy sets used in the Hilbe and Traulsen Model, and show how different degrees of threats in this model relate to punishment. The basic qualitative interplay between cooperation, defection, and punishment of this model are the same as that described for our Basic Model. However, since the Hilbe and Traulsen Model uses a much more complex strategy set of a total of 16 strategies, mathematical analysis is of limited intuitive utility and the model lends itself more readily to be explored through agent-based simulations.

4.5.1 Strategy Set

Model B is based on the state-of the art strategy set used by Hilbe and Traulsen (2012) [52]. Using this strategy set, which includes all forms of punishment and a form of reputation, Hilbe and Traulsen showed as a first how responsible punishment can evolve in a population model that allows for anti-social and spiteful punishment while simultaneously avoiding the problem of higher order free-riding (cooperative outcomes being endangered by Cooperators that do not punish Defectors invading the population). The model includes a total of 16 possible strategies. There are four possible strategies to play in the contribution stage and four possible strategies to play in the punishment phase of the PGG, as listed in Tables 1 and 2.

Two of the contribution phase strategies may take reputation about their co-players' punishment behavior into account when deciding whether to cooperate or defect. The level of reputation available is modeled through an environment parameter i that determines the probability with which any player knows their co-players' punishment strategies. Opportunistic Cooperators cooperate unless they know that it is beneficial to defect, which is the case if they know that the number of antisocial punishers A in the group is greater than the number of responsible punishers R. Opportunistic Defectors defect unless they know that it is beneficial to cooperate, which is the case if they know that that the number of responsible punishers R in the group is greater than the number of antisocial punishers A. To allow for errors in perception, a players' known reputation is wrong (perceived as a random other punishing strategy) with probability e. The presence of the conditional strategies and punishment reputation is crucial for the evolution of cooperation and responsible punishment. Responsible punishers can "force" Opportunistic Defectors into cooperating based on their reputation to punish Defectors, and hence Responsible punishers fare better than Non-Responsible Punishers in groups with Opportunistic Defectors (or Cooperators). Thus reputation fosters an intertwined benefit for cooperators and responsible punishers.

A representative example of the population dynamics of this strategy set under game parameters that allow for cooperation to evolve is shown in Figure 4.7. For readability, the plots show the aggregated proportion of contribution and punishment strategies over time separately. Observe that the population settles at a mix of (Opportunistic and regular) Cooperators as well as Non-Punishers and Responsible Punishers. The population was initialized with 100% Non-Punishing Opportunistic Defectors (OdN), but if enough knowledge about co-players' punishment reputation exists (high enough i), Opportunistic Cooperators (both R and N punishing types)

Table 4.1: Contribution Phase Strategies

Label	*Name*	*Description*
C	Cooperator	Always contributes.
D	Defector	Never contributes.
Oc	Opportunistic Cooperator	Cooperates unless it knows that it is beneficial to defect based on punishment reputation of co-players.
Od	Opportunistic Defector	Defects unless it knows that it is beneficial to cooperate based on punishment reputation of co-players.

Table 4.2: Contribution Phase Strategies

Label	*Name*	*Description*
R	Responsible Punisher	Punishes defecting players.
S	Spiteful Punisher	Punishes everyone.
A	Antisocial Punisher	Punishes cooperating players.
N	Non-Punisher	Punishes no one.

are able to invade the population, establishing cooperation as a norm. Once cooperation is established, CN and CR agents are able to grow to significant proportions as well. Eventually, the population settles at a mix of OcN, OcR, CN, and CR agents, with the rest of the strategies remaining at very low proportions, introduced each generation through exploration dynamics. (For a more detailed analysis of these dynamics, see [52].) All simulations for this model use a population size of 1280.

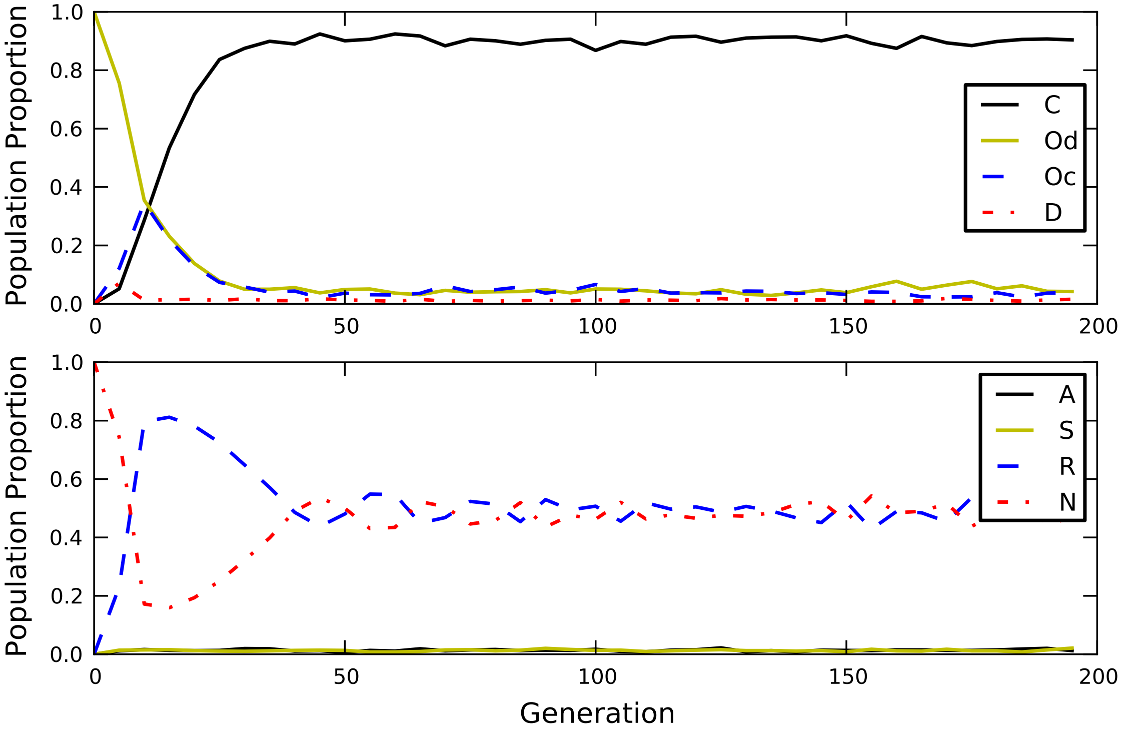

Figure 4.7: Example of evolution of strategy proportions of Model B. Model game and environment parameters are $r = 3, c = 1, \lambda = 1/2, \rho = 3/2, i = 0.7, \mu = e = s = 0.05, N = 5$.

4.5.2 Results

In this Section, I show that the operationalization of societal threats in the Hilbe and Traulsen Model support our thesis that higher threats lead to higher punishment propensities in populations. As in the Basic Model, increased threats increase punishment propensity. The manner in which this relationship comes to being however differs from the Basic Model, and does not even require group selection. In the Hilbe and Traulsen Model, as we saw in the above example, a mix of Responsible Punishers and Non-Punishers is stable within a population itself. Furthermore, the proportion of Punishers and Non-Punishers that is stable varies with societal threat in such a way that higher threat conditions lead to a higher proportion of Punishers in the stable state of the population:

- *Ecological Threats Increase Punishment Propensity:* Simulations of the Hilbe and Traulsen Model under different r parameters show that lower r (more societal threat) raises the stable amount of (Responsible) Punishers in the population. See left graph of Figure 4.8.

- *External Man-Made Threats Increase Punishment Propensity:* To measure the effect of external man-made threat, I ran simulations in which I introduce an influx of Defectors (in particular, Antisocial Punishing Defectors, since the invaders are hostile) in each generation, replacing random other agents. As before, a greater influx (proportion of Defectors added) represents higher threat. The results show again that the greater the threat, the greater the stable proportion of (Responsible) Punishers in the population. See right graph of Figure 4.8. After an influx greater than 10% of the population, the population breaks down into defection and hence I restricted the graph to this influx range.

The general dynamics in the Hilbe and Traulsen Model are both different and much more intricate than in the Basic Model. Most notably, a mix of Responsible and Non-Punishers is evolutionarily stable within a population, while in the Basic Model, punishment propensity is only neutrally stable and group selection is required to select between different punishment propensities. It is therefore even more surprising that the general relationship between societal threats and punishment, namely that higher threats lead to higher punishment, holds in this model just like in the Basic Model. This demonstrates the robustness of the causal relationship between societal threats and punishment.

4.6 Discussion

Cross-cultural social science has made great strides in understanding differences in cultural norms. This research expands upon this tradition through the use of evolutionary game theoretic models to study the evolution of differences in punishment toward norm violators across cultural groups. In this chapter I used evolutionary Public Goods Games with punishment to show how societies' optimal (in terms of overall group payoff) punishment propensities depend on the degree of societal threat that they face. In order to demonstrate the robustness of the results to various modeling choices, I implemented two models: The Basic Model, which

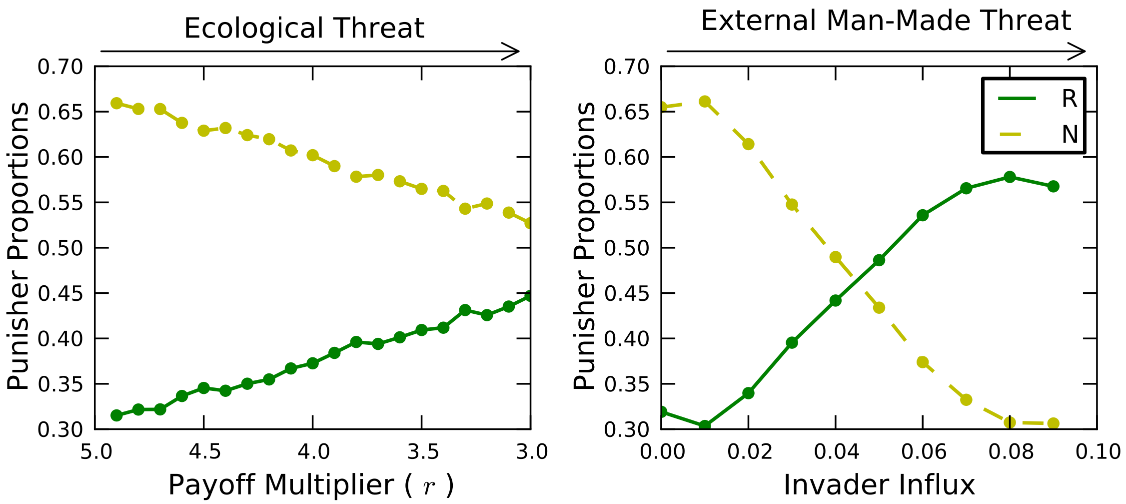

Figure 4.8: Left graph: Stable proportion of punishing types as a function of r. Lower r (higher threat) leads to more Punishers. Right graph: Stable proportion of punishing types as a function of the proportion of Defectors invading. Higher influx (higher threat) leads to more Punishers. Each point is the stable distribution (long-run average) determined by simulation. Game and environment parameters used: $r = 4.9, c = 1, \lambda = 1/2, \rho = 3/2, i = 0.7, \mu = e = s = 0.5, N = 5$.

assumes a population's ability to maintain a certain propensity to punish amongst its Cooperators, and the Hilbe and Traulsen Model, which is based on an expanded strategy set used in state of the art work on punishment and cooperation. This chapter considered two general types of threat examined by social science research in relation to punishment norms: ecological disasters and other threats to group resources, and external man-made threat. For each of these forms of societal threat and their plausible interpretations in the evolutionary game model, higher degrees of threat increased the punishment propensity in the population.

While I have only explored a form of peer-punishment in this work, I should note that recent research has demonstrated the importance and revived the exploration of institutional or "co-operative" [sic] punishment as an aspect of the evolution of cooperation [103, 59]. A general assumption is that such a punishing institution is publicly funded through a portion of agents' payoffs. Since institutional punishment acts as a replacement for individual punishment, strong (highly funded) institutions result in a decreased need for peer-punishment while weak (sparsely funded) institutions result in an increased need for peer-punishment. Most relevant to the current research is that societal threats are likely to weaken institutions by leading to a decrease in the overall payoff created by a society and hence a decrease in the relative size of the institutional punishment funds available. With a decrease in funding for (or effectiveness of) punishing institutions due to societal threats, there is again a greater need for individual-based punishment to maintain the same degree of cooperation. Therefore the existence of institutional or "co-operative" punishment mechanisms in a population would not change the general relationship between societal threat and punishment propensity illustrated in this paper.

Most existing research on punishment in evolutionary games has focused on whether the existence of punishment is evolutionarily viable and how it can aid the evolution of cooperation [45, 20, 21, 49, 48, 113, 44, 59, 123]. This work makes a significant contribution to this line of research by investigating relationships among differing degrees of societal threats, punishment propensities, and group survival.

Grounded in relevant social science data, the presented results show that the levels of societal threat to which different cultures are exposed can play an integral role in determining differences in cultures' evolved propensities to punish deviations from social norms. This illuminates the evolutionary basis for the wide variation in punishment rates that exists around the globe, may help predict changes in punishment propensities in different cultural groups, and helps promote cross-cultural understanding by showing how cultural differences in punishment propensities, which may appear puzzling, are generally adaptive to the society's ecological and historical context.

One limitation of the models presented in this chapter is that they investigated punishment in societies under threat that is constant over time, or under threat at one particular moment in time. An interesting question for future research is to investigate how societies respond to degrees of threat that vary over time. Figure 4.6 suggests that for a society to thrive under group selection, there is an evolutionary pressure for the society to adjust its punishment propensity toward q_{opt}. But the figure also suggests that for punishment propensities below q_{opt}, the possible consequence (societal disintegration) is much worse than the consequence (a slightly lower overall payoff) of too high a punishment propensity. This suggests that successful societies may raise their punishment propensities very quickly when external threats arise, and may be much slower to lower those punishment propensities when such threats abate.

In conclusion, in a world of increasing interdependence, it is critical to understand the mechanisms that drive cultural differences in norms. This work illustrates that evolutionary game theoretical models can be fruitfully integrated into cross-cultural social science to illuminate new insights into the nature of culture.

Chapter 5

EGT Approaches and the Study of Culture in Psychology

To this date there exists an incredibly rich literature on cultural differences in psychological dimensions, behaviors, norms, and related social measures between populations throughout the world. There are three main traditional theoretical perspectives in the study of culture and psychology: the cultural psychology approach, the indigenous culture approach, and the cross-cultural approach. All of these three perspectives recognize culture as a crucial behavioral influence, and thus culture must be considered in understanding human behavior and cognitive processes. However, each approach differs in its view on the extent to which universality of the human mind and psychological processes exists, and how culture relates to this issue. Hence each approach differs in the methods employed to study human psychology in relation to culture. My goal in this Chapter is not to give a comprehensive overview of these approaches, their contributions (which are plentiful), and their respective shortcomings. Rather, I aim to touch on the relevant aspects of these perspectives and recent trends in the science of psychology and culture that make evolutionary game theoretic approaches to culture a valuable and complementary approach.

5.1 Theoretical Approaches to Culture in Psychology

Cross-cultural psychology studies human behavior and mental processes across different cultures, commonly seeking to find universals and discovering relationships between psychological antecedents created by culture and the consequences of these antecedents on behavior [63]. A different theoretical perspective is that of cultural psychology, which, according to Schweder (1990), is the study of "the way cultural traditions and social practices regulate, express, and permute the human psyche, resulting less in psychic unity for humankind than in ethnic divergences of mind, self, and emotion [102]." In the view of cultural psychology, human beings and cultural environment, because they are so intertwined, cannot be separated analytically into independent and dependent variables, which is considered a flaw of the cross-cultural perspective. According to cultural psychology, humans and culture develop jointly within "intentional" worlds conceived by humans in a particular culture. A common criticism against cross-cultural approaches is that traditionally, cross-cultural approaches do not recognize this relationship between mind and culture; and thus instead of investigating how cultural practices shape psychological processes, cross-cultural studies are often too focused on testing the universality of psychological processes, which, allegedly, is often incorrectly assumed. Similar criticisms are sometimes made by proponents of the indigenous psychology perspective, which emphasizes the extent to which concepts and knowledge are specific to particular cultures, and stresses that "foreign theories and categories cannot necessarily be applied to understand behavior of a particular culture [105]."

In more recent years, the evolutionary approach to culture and psychology has gained more traction. The cultural evolution approach, argues that cultural diversity develops as a result of cultural transmission [73]. Humankind's capacity to acquire cultural information is unique in the animal world. Humans are agents that learn from their social surroundings, occasionally make errors in acquiring information or simply exploring new behavior. These actions are analogous to mutations or exploration dynamics in evolutionary game theoretic models and result in cultural variance. Under different environmental, social, or cultural conditions, different cultural variants may be stable, reinforced, or disappear. The existing cultural variant in turn affect the social, cultural, and possibly environmental conditions of that culture. In the cultural evolution view, this process is what drives the emergence of diverse cultures and resulting behaviors and norms around the world.

I believe that in many ways, the evolutionary approach to culture is the most precise and explicit in addressing the concerns about how culture affects behavior. The evolutionary approach does not assume culture informing behavior or vice versa, but through its circular process, recognizes the interdependence of the two, which is emphasized in the work of Shweder [102], and also represented in the ecocultural framework by Lonner and Adamopoulos [63]. As Newson, Richerson, and Boyd point out, the evolutionary approach alleviates the most common criticism of cultural behavior studies: namely they are not explanatory [73]. The lack of explanatory results is a criticism that probably applies most heavily to typical cross-cultural psychology studies, and this criticism could explain why cultural and indigenous psychologists are adamant about maintaining their approaches. The cultural and indigenous approaches much more aim to "recreate" the different "worlds" within which different cultures exist, or that different cultures have created, in order to understand different cultural groups. I hold that evolutionary game and multi-agent system models, while of course highly abstract and simplified, are some of the most explicit and transparent tools at our disposal for recreating such different worlds.

5.2 Structural Approaches to Understanding Culture

Until recently, most psychological studies took the *subjective* approach to culture, which generally solely considers culture as a result of individuals' internal mental representations. *Structural* approaches on the other hand emphasize the importance of external factors that affect individuals' psychology and behaviors [70]. Since EGT and multi-agent system models make it possible to operationalize such external factors and test their effects on individuals and their behaviors, structural approaches provide additional support for the use of these methods to study culture. In the past few years, there has been an increase in psychological studies that explore how external structural factors affect culture, individuals' minds, and their behavior. Examples of such structural approaches are the works of Yamagishi, Oishi, and Gelfand et al. Yamagishi presents a niche construction approach to culture, where a culture is a collectively created and maintained set of constraints and incentives and agents/humans are *cultural game players* that behave in ways to pursue goals [122]. Oishi considers residential mobility as a specific structural factor and explores how it affects psychology and culture [84]. Gelfand et al., as I have discussed in Chapter 3, relate various structural factors to the degree of "tightness" or "looseness" of different cultures [39]. From an evolutionary game theoretic approach, such structural studies are interesting and useful, as they present various possibilities of structural factors and their effects that can be tested or used to inform EGT and multi-agent system models of societies and cultures.

In general, evolutionary game theory is a useful framework to explore and understand how different external factors determine the nature of human interactions, and how this affects the evolution and dynamics of different behaviors. Of course, human behavior and psychology is extremely complex and consist of a multitude of possibly interacting factors and behaviors. Thus, in line with Richerson and Boyd's approach of "sample theory" [17], it is crucial to begin with simple-as-possible evolutionary game theoretic models that seek to understand fully the fundamental dynamics and evolutionary relation between a characteristics of interest and basic structural factors before moving to more complex models. A complete understanding of the basic relationships between cultural characteristics and certain structural factors form the necessary foundation to build and understand fuller, more complex, or ultimately complete models

Chapter 6

Conclusion

Even with the considerable surge of evolutionary game theoretic literature in the past decade, evolutionary game theory is still a relatively young field, especially as applied to the social sciences. Thus there still exists tremendous opportunity for it to be more fully integrated and applied to its potential in the social sciences to aid our understanding of human behaviors, culture, and societies. This thesis presents a step in this direction, solidifying evolutionary game theoretic approaches as a complementary approach to those common in cross-cultural social sciences, particularly psychology. The following section concludes with a summary of the contributions of this thesis.

6.1 Summary of Contributions

This thesis has presented evolutionary game theoretic models of the dynamics and evolution of human decision-making, specifically state-dependent risk preferences, and aspects of culture, specifically punishment norms. These models have been constructed in close consideration of social science data and collaboration with cultural psychologists. The inclusion of social science evidence in the construction of models presented in this thesis enabled this work to discover important new insights into the dynamics of human behaviors and culture that otherwise may have remained unexplored. In my work on risk preferences, I considered a range of imitation dynamics for the reproduction of strategies that, according to empirical evidence on social learning, arguably model the ways in which humans adapt their behavior more accurately than the commonly used replicator dynamics in biological or social applications of evolutionary game theory. In my work on punishment, empirical evidence from cultural psychology helped in identifying the environmental factors that may be the cause of different evolved punishment propensities, and I was thus able to explore the effects of these factors in evolutionary game models.

The main contributions of this thesis are 1) a simple sequential lottery game framework to study the evolution of human risk preferences and 2) demonstrations of how the principles observed in our lottery game studies affect classic cooperation games, and 3) game theoretic and multi-agent system-based PGG models incorporating various interpretations of societal threat to study punishment, demonstrating how societal threat plays an integral role in determining cultures' evolved punishment propensities. More specifically:

1. Studying risk preferences under the presented lottery game framework advances the understanding of human decision-making by showing how in evolutionary game environments with sequential choices, a large range of imitation dynamics can lead to state-dependent risk behavior that does not maximize expected payoff. In this range of dynamics, agents that are sometimes risk-prone and sometimes risk-averse in a manner that reflects descriptive models

of observed human decision-making are evolutionarily stable and hence proliferate.

2. The demonstrations of how the principles observed in the lottery game models affect classic cooperation games illustrates how state-dependent risk preferences under imitation dynamics can facilitate the evolution of cooperation in situations where cooperating entails risk, increasing the likelihood of cooperation emerging as the norm.
3. Incorporating societal threats identified by the social science literature into PGG models allowed for the study of their effects on the dynamics of punishment and cooperation. Increased threat led to the evolution of increased punishment propensities. The results from these models illuminate the evolutionary basis for the wide variation in punishment propensities that exists around the globe and provide support for a causal relationship between otherwise purely correlational data between societal threat and punishment behavior. This helps promote cross-cultural understanding by showing how the tightness or looseness of a society's cultural norms is adaptive to the society's ecological and historical context.

In general, the work in this thesis increases understanding of human behaviors, cultures, and their evolution, and how to model these through evolutionary game and multi-agent system approaches. Through the identification of important factors in cultural change and the characterization of their effects, these studies aid our understanding of socio-cultural processes that may lead to stability, instability, or general normative changes in different regions, environments, or populations. This work provides foundational knowledge likely required for predictive tools that can aid in making decisions about where and in what ways to invest resources in order to achieve desired societal outcomes. While the models presented are still highly abstract, understanding the relationships presented throughout these models are an integral part of the fundamental understanding required for more complex, detailed, and elaborate social modeling systems.

Appendix A

Derivations

This Section presents the derivation of $\nabla_\alpha(\text{RwS}, S_{a,b,c,d})$ used in Section 2.4.5. Using Table 2.4 and Figure 2.3, one can determine the probability of each pair of payoffs occurring and use them for the values of $p(r, s)$ as follows:

$$\begin{aligned}
\nabla_\alpha(\text{RwS}, S_{a,b,c,d}) =& |12-8|^\alpha(p)(apb(1-p)+apc(1-p)+(1-a)(1-d)) \\
&+|12-4|^\alpha(p)(a(1-p)(1-c)+(1-a)d(1-p)) \\
&+|12-0|^\alpha(p)(a(1-p)c(1-p)) \\
&+|8-4|^\alpha(p(1-p))(a(1-p)(1-c)+(1-a)d(1-p)) \\
&+|8-0|^\alpha(p(1-p))(a(1-p)c(1-p)) \\
&-|16-12|^\alpha p(apbp) \\
&-|16-8|^\alpha(1-p)p(apbp) \\
&-|16-0|^\alpha(1-p)^2(apbp) \\
&-|12-8|^\alpha(1-p)p(ap(1-b)+(1-a)dp) \\
&-|12-0|^\alpha(1-p)^2(ap(1-b)+(1-a)dp) \\
&-|8-0|^\alpha(1-p)^2(apb(1-p)+a(1-p)cp+(1-a)(1-d)) \\
&-|4-0|^\alpha(1-p)^2(a(1-p)(1-c)+(1-a)d(1-p))
\end{aligned}$$

Since we are considering the case where $p = 0.5$, we can collect terms as though $p = (1-p)$ (for ease of exposition we will wait to substitute 0.5 for p):

$$\begin{aligned}\nabla_\alpha(\mathrm{RwS}, S_{a,b,c,d}) =& |16|^\alpha(0 - abp^4) + |12|^\alpha(acp^3 - a(1-b)p^3 - (1-a)dp^3) \\ &+ (|4|^\alpha p - |8|^\alpha p^2)(abp^2 + acp^2 + (1-a)(1-d)) \\ &+ (|8|^\alpha p + |4|^\alpha p^2 - |4|^\alpha p^2)(a(1-c)p + (1-a)dp) \\ &+ |8|^\alpha(acp^4 + abp^4) \\ &+ |4|^\alpha(0 - abp^3 - a(1-b)p^3 - (1-a)dp^3),\end{aligned}$$

which yields

$$\begin{aligned}\nabla_\alpha(\mathrm{RwS}, S_{a,b,c,d}) =& -|16|^\alpha(abp^4) + |12|^\alpha(acp^3 - a(1-b)p^3 - (1-a)dp^3) \\ &+|8|^\alpha(acp^4 + a(1-c)p^2 + (1-a)dp^2 \\ &\quad - abp^4 - abp^4 - acp^4 - (1-a)(1-d)p^2) \\ &+|4|^\alpha(abp^3 + acp^3 + (1-a)(1-d)p - abp^3 - a(1-b)p^3 - (1-a)dp^3),\end{aligned}$$

which yields

$$\begin{aligned}\nabla_\alpha(\mathrm{RwS}, S_{a,b,c,d}) =& -|16|^\alpha(abp^4) + |12|^\alpha(acp^3 - a(1-b)p^3 - (1-a)dp^3) \\ &+|8|^\alpha(a(1-c)p^2 + (1-a)dp^2 - abp^4 - abp^4 - (1-a)(1-d)p^2) \\ &+|4|^\alpha(acp^3 + (1-a)(1-d)p - a(1-b)p^3 - (1-a)dp^3).\end{aligned}$$

Recollecting terms gives us

$$\begin{aligned}\nabla_\alpha(\mathrm{RwS}, S_{a,b,c,d}) =&(12^\alpha + 4^\alpha)acp^3 + 8^\alpha a(1-c)p^2\\ &+4^\alpha(1-a)(1-d)p + 8^\alpha(1-a)dp^2\\ &-(4^\alpha + 12^\alpha)(1-a)dp^3 - 8^\alpha(1-a)(1-d)p^2\\ &-(16^\alpha + 2 * 8^\alpha)abp^4 - (12^\alpha + 4^\alpha)a(1-b)p^3.\end{aligned}$$

Substituting $p = \frac{1}{2}$ and expanding the final term, we get

$$\begin{aligned}\nabla_\alpha(\mathrm{RwS}, S_{a,b,c,d}) =&\frac{1}{8}(12^\alpha + 4^\alpha)ac + \frac{1}{4}8^\alpha a(1-c)\\ &+\frac{1}{4}(2 * 4^\alpha - 8^\alpha)(1-a)(1-d) + \frac{1}{8}(2 * 8^\alpha - 4^\alpha - 12^\alpha)(1-a)d\\ &-\frac{1}{16}(16^\alpha + 2 * 8^\alpha)ab - \frac{1}{8}(12^\alpha + 4^\alpha)a + \frac{1}{8}(12^\alpha + 4^\alpha)ab,\end{aligned}$$

which yields

$$\begin{aligned}\nabla_\alpha(\mathrm{RwS}, S_{a,b,c,d}) =&\frac{1}{8}(12^\alpha + 4^\alpha)(ac-a) + \frac{1}{4}8^\alpha a(1-c)\\ &+\frac{1}{4}(2 * 4^\alpha - 8^\alpha)(1-a)(1-d) + \frac{1}{8}(2 * 8^\alpha - 4^\alpha - 12^\alpha)(1-a)d\\ &+\frac{1}{16}(2 * 12^\alpha + 2 * 4^\alpha - 2 * 8^\alpha - 16^\alpha)ab.\end{aligned}$$

Since $(ac - a) = -a(1-c)$, this yields

$$\begin{aligned}\nabla_\alpha(\mathrm{RwS}, S_{a,b,c,d}) =&\frac{1}{8}(2 * 8^\alpha - 4^\alpha - 12^\alpha)a(1-c)\\ &+\frac{1}{4}(2 * 4^\alpha - 8^\alpha)(1-a)(1-d) + \frac{1}{8}(2 * 8^\alpha - 4^\alpha - 12^\alpha)(1-a)d\\ &+\frac{1}{16}(2 * 12^\alpha + 2 * 4^\alpha - 2 * 8^\alpha - 16^\alpha)ab.\end{aligned}$$

Finally, recollecting terms gives us

$$\begin{aligned}\nabla_{\alpha}(\mathrm{RwS}, S_{a,b,c,d}) =& \frac{1}{8}(2 * 8^{\alpha} - 12^{\alpha} - 4^{\alpha})(a(1-c) + (1-a)d) \\ &+ \frac{1}{4}(2 * 4^{\alpha} - 8^{\alpha})(1-a)(1-d) \\ &+ \frac{1}{16}(2 * 12^{\alpha} + 2 * 4^{\alpha} - 2 * 8^{\alpha} - 16^{\alpha})ab,\end{aligned}$$

which matches Equation 2.3.

Appendix B

Simulation Pseudo-Code

Basic Model evolutionary PGG of Chapter 4:

```
- create initial population of agents
LOOP for each generation:
    # play games
    LOOP NumGames times
        - sample PopSize/M random disjoint sets of M agents from population
        LOOP for each set of agents
            - all agents in set play their type's strategy in the PGG
            - all agents in set receive payoffs from PGG
        END LOOP
    END LOOP
    - set fitness of all agents equal to their accumulated payoff
    # replicator dynamics
    - create new population according to the discrete replicator dynamics
    # exploration dynamics
    LOOP for each agent in new population
        - agent switches to a random strategy with probability $\mu$
    END LOOP
END LOOP
```

Bibliography

[1] Abel, A. B. Asset prices under habit formation and catching up with the Joneses. *The American Economic Review 80*, 2 (1990), 38–42.

[2] Alexander, J. M. Evolutionary game theory. In *The Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed., fall 2009 ed. 2009.

[3] Apesteguia, J., Huck, S., and Oechssler, J. Imitation–theory and experimental evidence. *Journal of Economic Theory 136*, 1 (Sept. 2007), 217–235.

[4] Axelrod, R., and Dion, D. The Further Evolution of Cooperation. *Science 242*, 4884 (Dec. 1988), 1385–1390.

[5] Axelrod, R., and Hamilton, W. D. The evolution of cooperation. *Science 211*, 4489 (March 1981), 1390–1396.

[6] Axelrod, R., and Hamilton, W. D. The evolution of cooperation. *Science 211*, 4489 (Mar. 1981), 1390–1396.

[7] Baye, M. R., and Morgan, J. Winner-take-all price competition. *Economic Theory 19*, 2 (2002), 271–282.

[8] Bendor, J., and Swistak, P. Types of evolutionary stability and the problem of cooperation. *Proceedings of the National Academy of Sciences 92* (Apr. 1995), 3596–3600,.

[9] Binmore, K. G. The evolution of fairness norms. *Rationality and Society 10* (1998), 275–302.

[10] Blume, L. The statistical mechanics of strategic interaction. *Games and Economic Behavior 5*, 3 (1993), 387–424.

[11] Blume, L., et al. The statistical mechanics of strategic interaction. *Games and economic behavior 5*, 3 (1993), 387–424.

[12] Boehm, C., Barclay, H. B., Dentan, R. K., Dupre, M. C., Hill, J. D., Kent, S., Knauft, B. M., Otterbein, K. F., and Rayner, S. Egalitarian behavior and reverse dominance hierarchy [and comments and reply]. *Current Anthropology 34*, 3 (1993), 227–254.

[13] Bowles, S. Group competition, reproductive leveling, and the evolution of human altruism. *Science (New York, N Y) 314*, 5805 (Dec. 2006), 1569–72.

[14] Bowles, S. Did warfare among ancestral hunter-gatherers affect the evolution of human social behaviors? *Science 324*, 5932 (June 2009), 1293–8.

[15] Bowles, S., and Gintis, H. The evolution of strong reciprocity: cooperation in heterogeneous populations. *Theoretical Population Biology 65*, 1 (Feb. 2004), 17–28.

[16] Boyd, R., Gintis, H., Bowles, S., and Richerson, P. J. The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences of the United States of America 100*, 6 (Mar. 2003), 3531–5.

[17] Boyd, R., and Richerson, P. J. *Culture and the evolutionary process.* University of Chicago Press, 1988.

[18] Boyd, R., and Richerson, P. J. Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology 13*, 3 (May 1992), 171–195.

[19] Boyd, R., and Richerson, P. J. Culture and the evolution of human cooperation. *Philosophical transactions of the Royal Society of London Series B, Biological sciences 364*, 1533 (Nov. 2009), 3281–8.

[20] Brandt, H., Hauert, C., and Sigmund, K. Punishment and reputation in spatial public goods games. *Proceedings Biological sciences / The Royal Society 270*, 1519 (May 2003), 1099–104.

[21] Brandt, H., Hauert, C., and Sigmund, K. Punishing and abstaining for public goods. *Proceedings of the National Academy of Sciences of the United States of America 103*, 2 (Jan. 2006), 495–7.

[22] Broom, M., Hadjichrysanthou, C., and Rychtar, J. Evolutionary games on graphs and the speed of the evolutionary process. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 466*, 2117 (Dec. 2009), 1327–1346.

[23] Camerer, C. F., and Fehr, E. When does" economic man" dominate social behavior? *Science 311*, 5757 (2006), 47.

[24] Colman, A. M. Cooperation, psychological game theory, and limitations of rationality in social interaction. *Behavioral and Brain Sciences 26* (2003), 139–198.

[25] Darwne, P. J., and Yao, X. Co-Evolutions in Iterated Prisoner's Dilemma with intermediate Levels of Cooperation: Application to Missile Defense. *International Journal of Computational Intelligence and Applications 2*, 1 (2002), 83–107.

[26] Dawid, H. Evolutionary game dynamics and the analysis of agent-based imitation models: The long run, the medium run and the importance of global analysis. *Journal of Economic Dynamics and Control 31*, 6 (June 2007), 2108–2133.

[27] Dawkins, R. *The Selfish Gene.* New York: Oxford University Press, 1976.

[28] De Quervain, D. J. F., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., and Fehr, E. The neural basis of altruistic punishment. *Science 305*, 5688 (2004), 1254.

[29] Dekel, E. On the Evolution of Attitudes towards Risk in Winner-Take-All Games. *Journal of Economic Theory 87*, 1 (July 1999), 125–143.

[30] Diecidue, E., and Van De Ven, J. Aspiration level, probability of success and failure, and expected utility. *International Economic Review 49*, 2 (May 2008), 683–700.

[31] Eguluz, V. M., Zimmermann, M. G., J Cela-Conde, C., and San Miguel, M. Cooperation and emergence of role differentiation in the dynamics of social networks. *American Journal of Sociology 110*, 977 (Feb. 2005).

[32] Epstein, L. G., and Zin, S. E. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica: Journal of the Econometric Society* (1989), 937–969.

[33] Epstein, L. G., and Zin, S. E. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis. *Journal of Political Economy* (1991), 263–286.

[34] Fang, C., Kimbrough, S. O., Pace, S., Valluri, A., and Zheng, Z. On adaptive emergence of trust behavior in the game of stag hunt. *Group Decision and Negotiation 11*, 6 (Nov. 2002), 449–467.

[35] Fehr, E., and Fischbacher, U. The nature of human altruism. *Nature 425*, 6960 (2003), 785–791.

[36] Fehr, E., and Gächter, S. Altruistic punishment in humans. *Nature 415*, 6868 (2002), 137–140.

[37] Fletcher, J. A., and Zwick, M. Strong altruism can evolve in randomly formed groups. *Journal of Theoretical Biology 228*, 3 (June 2004), 303–13.

[38] Friedman, M., and Savage, L. J. The utility analysis of choices involving risk. *The Journal of Political Economy 56*, 4 (Aug. 1948), 279–304.

[39] Gelfand, M. J., Raver, J. L., Nishii, L., Leslie, L. M., Lun, J., Lim, B. C., Duan, L., Almaliach, A., Ang, S., Arnadottir, J., et al. Differences between tight and loose cultures: A 33-nation study. *science 332*, 6033 (2011), 1100.

[40] Gintis, H. *Game Theory Evolving: A Problem-centered Introduction to Modeling Strategic Behavior.* Princeton University Press, 2000.

[41] Gintis, H. *Moral sentiments and material interests: The foundations of cooperation in economic life*, vol. 6. The MIT Press, 2005.

[42] Hammerstein, P. *Genetic and cultural evolution of cooperation.* The MIT Press, 2003.

[43] Harley, C. B. Learning the evolutionarily stable strategy. *Journal of Theoretical Biology 89* (1981), 611–633.

[44] Hauert, C. Replicator dynamics of reward & reputation in public goods games. *Journal of Theoretical Biology 267*, 1 (2010), 22–28.

[45] Hauert, C., De Monte, S., Hofbauer, J., and Sigmund, K. Replicator dynamics for optional public good games. *Brain 218*, 2 (2002), 187–194.

[46] Hauert, C., and Szabo, G. Game theory and physics. *American Journal of Physics 73*, 5 (2005), 405–414.

[47] Hauert, C., Traulsen, A., Brandt, H., Nowak, M. A., and Sigmund, K. Via freedom to coercion: the emergence of costly punishment. *Science 316*, 5833 (2007), 1905.

[48] Henrich, J. Cultural group selection, coevolutionary processes and large-scale cooperation. *Journal of Economic Behavior & Organization 53*, 1 (Jan. 2004), 3–35.

[49] Henrich, J., and Boyd, R. Why people punish defectors Weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas. *Journal of theoretical biology 208*, 1 (Jan. 2001), 79–89.

[50] Henrich, J., Ensminger, J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., Cardenas, J. C., Gurven, M., Gwako, E., Henrich, N., et al. Markets, religion, community size, and the evolution of fairness and punishment. *Science 327*, 5972 (2010), 1480.

[51] Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cardenas, J. C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D., and Ziker, J. Costly punishment across human societies. *Science (New York, N Y) 312*, 5781 (June 2006), 1767–70.

[52] Hilbe, C., and Traulsen, A. Emergence of responsible sanctions without second order free riders, antisocial punishment or spite. *Scientific Reports 2* (2012).

[53] Hofbauer, J., and Sigmund, K. Evolutionary game dynamics. *Bulletin of the American Mathematical Society 40*, 4 (July 2003), 479–519.

[54] Houston, A. L. H., McNamara, J. 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.

[55] Huberman, B. A., and Glance, N. S. Evolutionary games and computer simulations. *Proceedings of the National Academy of Sciences of the United States of America 90*, 16 (Aug. 1993), 7716–7718.

[56] Huck, S., Normann, H. T., and Oechssler, J. Learning in cournot oligopoly: An experiment. *SSRN eLibrary* (Sept. 1997).

[57] Huck, S., Normann, H. T., and Oechssler, J. Does information about competitors' actions increase or decrease competition in experimental oligopoly markets? *International Journal of Industrial Organization 18*, 1 (2000), 39–57.

[58] J, A. K. *Essays in the theory of risk-bearing.* Markham, Chicago, 1971.

[59] Jaffe, K., and Zaballa, L. Co-operative punishment cements social cohesion. *Journal of Artificial Societies and Social Simulation 13*, 3 (2010), 4.

[60] Kahneman, D., and Tversky, A. Prospect theory: An analysis of decision under risk. *Econometrica 47* (1979), 263–291.

[61] Kimbrough, S. Foraging for trust: Exploring rationality and the stag hunt game. *Trust Management* (2005), 151–156.

[62] Lewontin, R. C. Evolution and the theory of games. *Journal of theoretical biology 1*, 3 (1961), 382–403.

[63] Lonner, W., and Adamopolous, J. *Culture as an antecedent to behavior.* Needham Heights, MA: Allyn and Bacon, 1997, pp. 44–67.

[64] Loomes, G., and Sugden, R. Regret theory: An alternative theory of rational choice under uncertainty. *The Economic Journal 92* (December 1982), 805–824.

[65] Lopes, L. L. Between hope and fear: The psychology of risk. *Advances in Experimental Social Psychology 20* (1987), 255–295.

[66] Lopes, L. L. Re-modeling risk aversion. In *Acting under uncertainty: Multidisciplinary conceptions*, G. M. von Furstenberg, Ed. Boston: Kluwer, 1990, pp. 267–299.

[67] 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.

[68] Marlowe, F. W., Berbesque, J. C., Barr, A., Barrett, C., Bolyanatz, A., Cardenas, J. C., Ensminger, J., Gurven, M., Gwako, E., Henrich, J., et al. More 'altruistic' punishment in larger societies. *Proceedings of the Royal Society B: Biological Sciences 275*, 1634 (2008), 587.

[69] Maynard Smith, J. *Evolution and the theory of games.* Cambridge University Press, Cambridge, United Kingdom, 1982.

[70] Morris, M, P. J., and Ariel, S. *Missing relations: Incorporating relational constructs into models of culture.* Thousand Oaks, Calif : SAGE, 2000, pp. 52–90.

[71] Mui, L., Mohtashemi, M., and Halberstadt, A. Notions of reputation in multi-agents systems. In *Proceedings of the first international joint conference on Autonomous agents and multiagent systems part 1 - AAMAS '02* (New York, New York, USA, July 2002), ACM Press, p. 280.

[72] Nakamaru, M., and Iwasa, Y. The coevolution of altruism and punishment: Role of the selfish punisher. *Journal of theoretical biology 240*, 3 (2006), 475–488.

[73] Newson, L, R. P. J., and Boyd, R. *Cultural evolution and the shaping of cultural diversity.* New York: Guilford Press, 2007, pp. 454–476.

[74] Nowak, M. A. Five Rules for the Evolution of Cooperation. *Science 314*, 5805 (Dec. 2006), 1560–1563.

[75] Nowak, M. A., Bonhoeffer, S., and May, R. M. Spatial games and the maintenance of cooperation. *Proceedings of the National Academy of Sciences of the United States of America 91*, 11 (May 1994), 4877–4881.

[76] Nowak, M. A., and May, R. M. Evolutionary games and spatial chaos. *Nature 359*, 6398 (Oct. 1992), 826–829.

[77] Nowak, M. A., Page, K. M., and Sigmund, K. Fairness Versus Reason in the Ultimatum Game. *Science 289*, 5485 (Sept. 2000), 1773–1775.

[78] Nowak, M. A., and Sigmund, K. Tit for tat in heterogeneous populations. *Nature 355*, 6357 (1992), 250–253.

[79] 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.

[80] Nowak, M. A., Tarnita, C. E., and Wilson, E. O. The evolution of eusociality. *Nature 466*, 7310 (Aug. 2010), 1057–62.

[81] Offerman, T., Potters, J., and Sonnemans, J. Imitation and belief learning in an oligopoly experiment. *The Review of Economic Studies 69*, 4 (Oct. 2002), 973–997.

[82] Ohtsuki, H., Hauert, C., Lieberman, E., and Nowak, M. A. A simple rule for the evolution of cooperation on graphs and social networks. *Nature 441*, 7092 (May 2006), 502–505.

[83] Ohtsuki, H., and Nowak, M. A. Direct reciprocity on graphs. *Journal of theoretical biology 247*, 3 (2007), 462–470.

[84] Oishi, S. The psychology of residential mobility implications for the self, social relationships, and well-being. *Perspectives on Psychological Science 5*, 1 (2010), 5–21.

[85] Okamoto, K., and Matsumura, S. The evolution of punishment and apology: an iterated prisoner's dilemma model. *Evolutionary Ecology 14*, 8 (Nov. 2000), 703–720.

[86] Ostrom, E., Gardner, R., and Walker, J. *Rules, games, and common-pool resources.* Univ of Michigan Pr, 1994.

[87] Page, K., and Nowak, M. Empathy leads to fairness. *Bulletin of Mathematical Biology 64*, 6 (Nov. 2002), 1101–1116.

[88] 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.

[89] Pelto, P. J. The difference between 'tight' and 'loose' societies. *Transaction* (1968), 37–40.

[90] Popp, A. The effects of natural disasters on long run growth. *Major Themes in Economics 1* (2006), 61–81.

[91] Price, M. E., Cosmides, L., and Tooby, J. Punitive sentiment as an anti-free rider psychological device. *Evolution and Human Behavior 23*, 3 (2002), 203–231.

[92] Rabin, M. Risk aversion and Expected-Utility theory: A calibration theorem. *Econometrica 68*, 5 (Sept. 2000), 1281–1292.

[93] Riolo, R. L., Cohen, M. D., and Axelrod, R. Evolution of cooperation without reciprocity. *Nature 411* (2001), 441–443.

[94] Rode, C., and Wang, X. Risk-Sensitive Decision making Examined within an Evolutionary Framework. *American Behavioral Scientist 43*, 6 (March 2000), 926–939.

[95] Roos, P., and Nau, D. S. Risk preference and sequential choice in evolutionary games. *Advances in Complex Systems* (2010). To appear.

[96] Santos, F. C., Pacheco, J. M., and Lenaerts, T. Evolutionary dynamics of social dilemmas in structured heterogeneous populations. *Proceedings of the National Academy of Sciences of the United States of America 103*, 9 (Feb. 2006), 3490–4.

[97] Santos, F. C., Santos, M. D., and Pacheco, J. M. Social diversity promotes the emergence of cooperation in public goods games. *Nature 454*, 7201 (July 2008), 213–6.

[98] Schlag, K. H. Why imitate, and if so, how? a bounded rational approach to Multi-Armed bandits. *Journal of Economic Theory 78*, 1 (1998), 130–156.

[99] Schlag, K. H. Which one should I imitate? *Journal of Mathematical Economics 31*, 4 (1999), 493–522.

[100] Schotter, A., and Sopher, B. Social learning and coordination conventions in intergenerational games: An experimental study. *Journal of Political Economy 111*, 3 (2003).

[101] Sethi, R. Preference Evolution and Reciprocity. *Journal of Economic Theory 97*, 2 (Apr. 2001), 273–297.

[102] Shweder, R. A. *Cultural psychology: What is it?* Cambridge, MA: Cambridge University Press, 1990, pp. 1–43.

[103] Sigmund, K., De Silva, H., Traulsen, A., and Hauert, C. Social learning promotes institutions for governing the commons. *Nature 466*, 7308 (July 2010), 861–863.

[104] Sigmund, K., Hauert, C., and Nowak, M. A. Reward and punishment. *Proceedings of the National Academy of Sciences 98*, 19 (2001), 10757.

[105] Sinha, D. *Indigenizing psychology.* Needham Heights, MA: Allyn and Bacon, 1997, pp. 129–169.

[106] Soltis, J., Boyd, R., and Richerson, P. J. Can Group-Functional Behaviors Evolve by Cultural Group Selection?: An Empirical Test. *Current Anthropology 36*, 3 (1995).

[107] Stevens, J. R. Rational decision making in primates: the bounded and the ecological. *Primate neuroethology Oxford University Press, Oxford* (2010), 98–116.

[108] Syrms, B. *The Stag Hunt and the Evolution of Social Structure.* Cambridge University Press, Cambridge, U K, 2003.

[109] Szabó, G., and Tőke, C. Evolutionary prisoner's dilemma game on a square lattice. *Phys Rev E 58*, 1 (Jul 1998), 69–73.

[110] Szabó, G., and Tőke, C. Evolutionary prisoner's dilemma game on a square lattice. *Physical Review E 58*, 1 (1998), 69.

[111] Taylor, P. D., and Jonker, L. B. Evolutionary stable strategies and game dynamics. *Mathematical Biosciences 40*, 1-2 (1978), 145 – 156.

[112] Traulsen, A., and Hauert, C. Stochastic evolutionary game dynamics. In *Reviews of Nonlinear Dynamics and Complexity* (2009), H. G. Shuster, Ed., Wiley-VHC, New York.

[113] Traulsen, A., Hauert, C., De Silva, H., Nowak, M. A., and Sigmund, K. Exploration dynamics in evolutionary games. *Proceedings of the National Academy of Sciences 106*, 3 (2009), 709.

[114] Traulsen, A., and Nowak, M. A. Evolution of cooperation by multilevel selection. *Proceedings of the National Academy of Sciences of the United States of America 103*, 29 (July 2006), 10952–5.

[115] Traulsen, A., Nowak, M. A., and Pacheco, J. M. Stochastic dynamics of invasion and fixation. *Physical Review E 74*, 1 (2006), 011909.

[116] Traulsen, A., Pacheco, J. M., and Nowak, M. A. Pairwise comparison and selection temperature in evolutionary game dynamics. *Journal of Theoretical Biology 246*, 3 (2007), 522–529.

[117] Traulsen, A., Semmann, D., Sommerfeld, R. D., Krambeck, H.-J., and Milinski, M. Human strategy updating in evolutionary games. *Proceedings of the National Academy of Sciences of the United States of America 107*, 7 (Feb. 2010), 2962–6.

[118] Tversky, A., and Kahneman, D. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty 5* (1992), 297–323.

[119] Vega-Redondo, F. The evolution of walrasian behavior. *Econometrica 65*, 2 (Mar. 1997), 375–384. ArticleType: primary_article / Full publication date: Mar , 1997 / Copyright 1997 The Econometric Society.

[120] von Neumann, J., and Morgenstern, O. *Theories of games and economic behavior*. Princeton University Press Princeton, NJ, 1944.

[121] Warneryd, K. Rent, risk, and replicationPreference adaptation in winner-take-all markets. *Games and Economic Behavior 41*, 2 (Nov. 2002), 344–364.

[122] Yamagishi, T. Micro–macro dynamics of the cultural construction of reality. *Advances in Culture and Psychology 1* (2010), 5–51.

[123] Ye, H., Tan, F., Ding, M., Jia, Y., and Chen, Y. Sympathy and punishment: Evolution of cooperation in public goods game. *Journal of Artificial Societies and Social Simulation 14*, 4 (2011), 20.