universität wien

DISSERTATION

Titel der Dissertation

Evolutionary games and equilibrium selection

Verfasser

Boyu Zhang

angestrebter akademischer Grad

Doktor der Naturwissenschaften (Dr.rer.nat.)

Wien, im August 2012

Studienkennzahl lt. Studienblatt:	A0948612
Dissertationsgebiet lt. Studienblatt:	Mathematik
Betreuer:	Univ. Prof. Josef Hofbauer

Acknowledgements

Foremost, I would like to thank my PhD supervisors Prof. Josef Hofbauer and Prof. Karl Sigmund. I have been extremely lucky to learn from two of the most famous experts in the field of evolutionary game theory. Prof. Josef Hofbauer provided me with the opportunity to study at the University of Vienna. I am very grateful for his patience, motivation, enthusiasm, and immense knowledge in evolutionary dynamics. He introduced me to the subject of equilibrium selection and has always been available to advise me. Prof. Karl Sigmund deeply impressed me from his passion, sense to research ideas and insightful views towards many scientific branches. His support and guidance made the experiment (in Chapter 5) possible.

Second, I would like to thank the Chinese Scholarship Council (CSC) for three years scholarship. The support from the Austrian Science Funds and the European Science Foundation (TECT I-104 G15) for running the experiment is acknowledged. WWTF (project number MA09-017) supported the presentation of the thesis at the 4th World Congress of the Game Theory Society. Many thanks to the Chinese Embassy (especially Prof. Hangzhu Chen) and the Chinese Student Union for help during the study abroad.

I am also indebted to Ross Cressman, Daisuke Oyama, William Sandholm, Yi Tao, Jorgen Weibull, Xiaoguang Yang, my colleagues and collaborators Peter Bednarik, Zhigang Cao, Christian Hilbe, De Silva Hannelore, Jun Honda, Cong Li, Mathias Staudigl, and two reviewers for valuable discussions and suggestions.

Moreover, I am grateful to all my friends in Vienna and Beijing for helping me get through the difficult times, and for all the emotional support and entertainment.

Finally, I would like to give my deepest gratitude to my parents, who made me who I am. I dedicate this thesis to them.

Abstract

The object of this thesis is to study several equilibrium selection methods for certain classes of games and compare to what extent these selection methods lead to similar or different results. The thesis consists of five chapters.

Chapter 1 describes a theoretical framework for equilibrium selection by tracing the graph of the quantal response equilibrium (QRE) correspondence.

Chapter 2 analyzes the quantal response methods for equilibrium selection in detail for 2×2 bimatrix games.

Chapter 3 investigates the ultimatum game by a learning-mutation process related to the quantal response equilibrium.

Chapter 4 studies two equilibrium selection methods based on the replicator dynamics.

Chapter 5 provides a economic experiment to show that social learning can lead to a spontaneously emerging social contract, based on a sanctioning institution to overcome the free rider problem.

Zusammenfassung

In dieser Dissertation werden mehrere Methoden zur Gleichgewichtsselektion für gewisse Klassen von Spielen studiert. Es wird untersucht, inwiefern diese Methoden zu ähnlichen oder verschiedenen Resultaten führen. Die Dissertation besteht aus fünf Kapiteln.

In Kapitel 1 werden die theoretischen Grundlagen einer Homotopiemethode entlang des Graphen der quantal response Gleichgewichte beschrieben.

In Kapitel 2 wird diese Methodik im Detail auf 2×2 Bimatrixspiele angewendet.

Kapitel 3 untersucht das Ultimatumspiel mittels eines Lern- und Mutationsprozesses.

Kapitel 4 widmet sich zwei weiteren Methoden der Gleichgewichtsauswahl, die auf der Replikatorgleichung basieren.

Kapitel 5 stellt ein ökonomisches Experiment vor, das zeigt, wie eine strafende Institution dem Problem der Trittbrettfahrer Herr werden kann.

Contents

Preface

"In general, a given game may have several equilibria. Yet uniqueness is crucial to the foregoing argument. Nash equilibrium makes sense only if each player knows which strategies the others are playing; if the equilibrium recommended by the theory is not unique, the players will not have this knowledge. Thus it is essential that for each game, the theory selects one unique equilibrium from the set of all Nash equilibria." -Robert Aumann (foreword to Harsanyi and Selten, 1988)

In a game, if each player has chosen a strategy and no player can benefit by changing his or her strategy while the other players keep theirs unchanged, then the set of strategy choices is called a Nash equilibrium. Every game has at least one Nash equilibrium (Nash, 1950) but in general there are many. Trying to select the "best" equilibrium for each game is a difficult problem. Methods to do this have been suggested by Harsanyi and Selten (1988), inventing the *risk dominant equilibrium*, and by many other researchers.

This thesis studies several equilibrium selection models. These models could be roughly classified into two categories. Evolutionary game theory consider the behavior of large populations, where individuals choose which actions to play genetically or using simple myopic rules (e.g., best response, imitation). In contrast, learning models focus on the behavior of small groups in repeated games. Individuals make decisions according to explicit learning rules, which could be simple myopic rules (called heuristic learning or adaptive learning) or more complicated Bayesian rules (called coordinated Bayesian learning or rational learning). The heuristic learning is close to the spirit of evolutionary approach. In the Bayesian learning, individuals play the best response to their beliefs about other individuals' strategies and update the beliefs over rounds.

One representative class of Bayesian learning methods consist of homotopy approaches, such as the tracing procedure of Harsanyi and Selten (1988; Harsanyi, 1975) or the (one parameter family of) quantal response equilibria of McKelvey and Palfrey (1995, 1998; Turocy, 2005). In these models, individuals are usually considered boundedly rational that may make mistakes in estimating the utilities of their strategies. As players gain experience from repeated observations, they can be expected to make more precise estimations and finally reach a Nash equilibrium. The tracing procedure always leads to the risk dominant equilibrium but quantal response equilibria do not.

On the other hand, from the point of evolution, a simple idea is to choose the equilibrium with the largest basin of attraction (for the replicator dynamics or some other deterministic evolutionary dynamics). This implies that a population with uncertain initial state is more likely to evolve to the dominant equilibrium in the long run. For symmetric 2×2 games, the risk dominant equilibrium has the largest basin of attraction, but this is not true for more general situations.

The object of the thesis is to study these equilibrium selection methods for certain classes of games and compare to what extent these selection methods lead to similar or different results. The thesis consist of five chapters. Chapter 1 describes a theoretical framework for equilibrium selection by tracing the graph of the quantal response equilibrium (QRE) correspondence. Chapter 2 analyzes the quantal response methods for equilibrium selection in detail for 2×2 bimatrix games. Chapter 3 investigates the ultimatum game by a learning-mutation process related to the quantal response equilibrium. Chapter 4 studies two equilibrium selection methods based on the replicator dynamics. Chapter 5 provides an economic experiment which is a follow-up on a theoretical paper by Sigmund et al. (2010). Figure 1 summarizes the interactions among the chapters.

Chapters 1-4 are written under the guidance of Prof. Josef Hofbauer. Chapter 5 is a joint work with Cong Li, Dr. Hannelore De Silva, Peter Bednarik and Prof. Karl Sigmund.

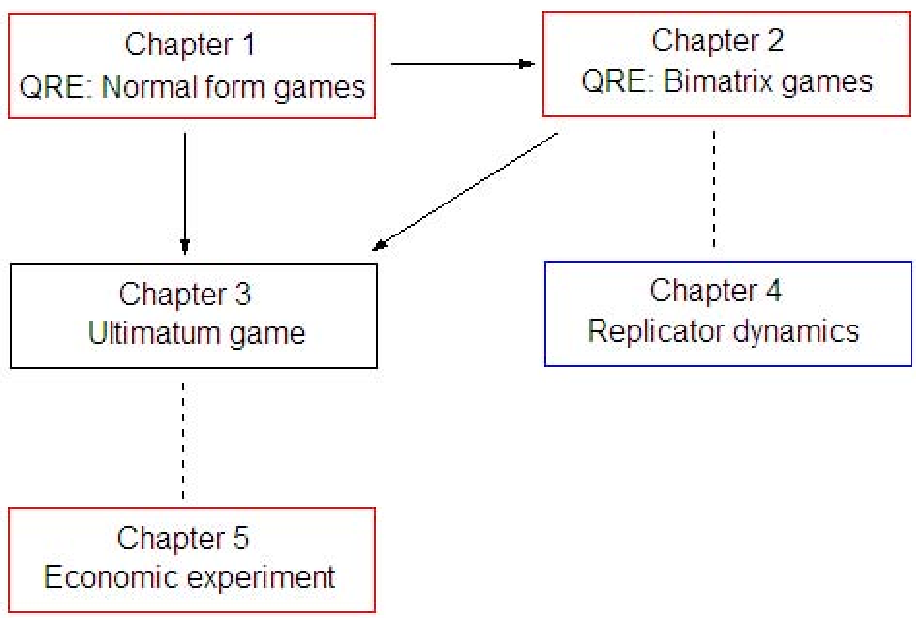


Figure 1: The main interactions between the chapters. The colors of the boxes indicates the category of equilibrium selection models appeared in the chapters: Red means learning approach, blue means evolutionary approach and black means a combination of learning and evolution. A solid arrow connecting two boxes indicates that one chapter depends on the other. A dash line connecting two boxes indicates that two chapters studies the same class of games.

Chapter 1

Quantal response methods for equilibrium selection: Normal form games

Abstract

This chapter describes a theoretical framework for equilibrium selection by tracing the graph of the quantal response equilibrium (QRE) correspondence as a function of the estimation error. If a quantal response function satisfies C^2 continuity, monotonicity and cumulativity, the graph of QRE correspondence generically includes a unique branch that starts at the centroid of the strategy simplex and converges to a unique Nash equilibrium as noises vanish. This equilibrium is called the limiting QRE of the game. We show that the limiting QRE of a symmetric game must be a symmetric Nash equilibrium, and provide a sufficient condition for the limiting QRE in two-person symmetric games.

Key words

Quantal response equilibrium; equilibrium selection; symmetric game; role game

1.1 Introduction

Quantal response equilibrium (QRE) was introduced by McKelvey and Palfrey (1995) in the context of bounded rationality. In a QRE, players do not always choose best responses. Instead, they make decisions based on a probabilistic choice model (called the quantal response or the perturbed best response) and assume other players do so as well. A general interpretation of this model is that players observe random perturbations on the payoffs of strategies and choose optimally according to those noisy observations (McKelvey and Palfrey, 1995, 1998; Goeree et al., 2005; Turocy, 2005; Sandholm, 2010). For a given error structure, QRE is defined as a fixed point of this process. [1]

The most common specification of QRE is the logit equilibrium, where the noises follow the extreme value distribution (Luce, 1959; McFadden, 1976; Blume, 1993, 1995; McKelvey and Palfrey, 1995, 1998; Anderson et al., 2004; Turocy, 2005; Hofbauer and Sandholm, 2002, 2007; Sandholm, 2010). The logistic response function has one free parameter λ, whose inverse $\frac{1}{\lambda}$ has been interpreted as the temperature, or the intensity of noise. At $\lambda = 0$, players have no information about the game and each strategy is chosen with equal probability. As λ approaches infinity, players achieve full information about the game and choose the best responses. McKelvey and Palfrey (1995) then defined an equilibrium selection from the set of Nash equilibria by "tracing" the branch of the logit equilibrium correspondence starting at the centroid of the strategy simplex (the only QRE when $\lambda = 0$) and continuing for larger and larger values of λ. For almost all normal form games, this branch converges to a unique Nash equilibrium as λ goes to infinity. This Nash equilibrium is called the *limiting logit equilibrium* (LLE) of the game. Later, McKelvey and Palfrey (1998) extended the original notion of QRE to extensive-form games (AQRE), and they found that the logit-AQRE also implies a unique selection from the set of sequential equilibria in generic extensive form games.

QRE allows every strategy to be played with non-zero probability, therefore can be applied to explain data from laboratory experiments which Nash equilibrium analysis can not. In McKelvey and Palfrey's original paper (1995), they analyzed data from four past experiments on two-person normal form games, where participants displayed non-equilibrium behaviors that are anomalous with respect to standard game theory. [2]

[1] The model is equivalent to an incomplete information game where the actual payoff is the sum of payoffs of some fixed game and independent random terms, and each players private signal is his own payoffs. A QRE is a probability distribution of action profiles in a Bayesian Nash equilibrium (Ui, 2006). Ui (2002) also provided an evolutionary interpretation for QRE. In an n-population game, if a stochastic best response process satisfies the detailed balance condition then the support of the stationary distribution converges to the set of quantal response equilibria as the population size goes to infinity.

[2] These experiments include 3 by 3 zero sum game (Lieberman, 1960), 4 by 4 zero sum game (O'Neill, 1987), 5 by 5 zero sum game (Rapoport and Boebel, 1992) and other bimatrix games with unique mixed equilibria (Ochs, 1993).

For each experiment, they compared subjects' choices period by period with the logit equilibrium and calculated the maximum likelihood estimate of the noise parameter λ. They found that the QRE model is surprisingly successful in fitting the data. Subsequent studies include auctions (Anderson et al., 1998; McKelvey and Palfrey, 1998; Goeree et al., 2002), bargaining (Goeree and Holt, 2000; Yi, 2005), social dilemmas (Capra et al., 1999; Goeree and Holt, 2001), coordination games (Anderson et al., 2001) and games with network structures (Choi et al., 2009). In these experiments, estimates of λ usually increased as the game progresses. [3] This then provides an empirical evidence of the equilibrium selection above. As players gain experience from repeated observations, they can be expected to make more precise estimates of the expected payoffs of different strategies.

Formally, a quantal response function maps the vector of expected payoffs into a vector of choice probabilities. Haile et al. (2008) pointed out that without further restrictions on the error structures, QRE can be constructed to match any choice probabilities in any normal form game. Therefore, sensible empirical assumptions on the distributions of payoff perturbations are necessary. Haile et al. (2008) then suggested two promising restrictions: exchangeability and invariance. Responding to an earlier draft of this paper (Haile et al., 2004), Goeree et al. (2005) proposed a "reduced form" definition of QRE. Rather than restricting payoff disturbances explicitly, they define a regular QRE by restricting quantal response functions to satisfy four axioms: continuity, interiority, responsiveness, and monotonicity. They showed that exchangeability is a sufficient condition for monotonicity and invariance is a sufficient condition for responsiveness. Hence, payoff perturbations that satisfy exchangeability and invariance generate regular QRE. More generally, the reduced form approach does not require that quantal response functions are derived from some underlying choice models of stochastic utility maximization, therefore allows for a richer set of models for data estimation.

In this chapter, we describe a theoretical framework for equilibrium selection by quantal response methods in normal form games. Following the logit equilibrium, define a QRE at noise level λ as a fixed point of quantal response functions where payoffs are multiplied by the factor λ. The set of QRE can be viewed as a correspondence from λ to the set of mixed strategy profiles. Similarly as Goeree et al. (2005), we impose three restrictions on quantal response functions: C^2 continuity, monotonicity and cumulativity. Continuity is a technical property, and both monotonicity and cumulativity have significant economic content. Monotonicity is a weak form of rational choice, meaning that strategies with higher expected payoffs are used more frequently. Cumulativity ensures that players choose best responses as λ goes to infinity. Intuitively, quantal response functions that satisfy the three axioms are smooth generalizations of best response functions.

[3] Although there is a tendency for λ to increase with experience, estimates of λ from different experiments can vary significantly. See the effect of payoff magnitude on λ in McKelvey et al., 2000.

We show that for almost all normal form games, there is a unique equilibrium selection by tracing the graph of the QRE correspondence. The selected Nash equilibrium is called the *limiting QRE* of the game.

The rest of this chapter is organized as follows. Section 1.2 defines QRE at noise level λ and introduces some properties. Section 1.3 studies the topological structure of the graph of QRE correspondence. If a quantal response function satisfies (C^0) continuity, monotonicity and cumulativity, the graph contains a component that connects the centroid of the strategy simplex and a Nash equilibrium. If the quantal response function is further C^2 continuous, for almost all normal form games, this component is diffeomorphic to a C^1 segment, which implies a unique equilibrium selection. Section 1.4 indicates that the limiting QRE of a symmetric game must be a symmetric Nash equilibrium. Section 1.5 provides a sufficient condition for the limiting QRE in two-person symmetric games and compares the limiting QRE to other equilibrium notions. Section 1.6 shows that there is a one-to-one mapping between the logit equilibria of a bimatrix game and the corresponding symmetric role game.

1.2 Quantal response equilibrium

Consider an n-person normal-form game $\Gamma = (N, S, u)$, where $N = \{1, ..., n\}$ is the set of *players*. For each player $i \in N$, there is a *strategy set* $S_i = \{s_{i1}, ..., s_{iJ_i}\}$ consisting of J_i pure strategies and a *payoff function*, $u_i : S \to \mathbb{R}$, where $S = \prod_{i \in N} S_i$ is the set of strategy profiles.

Let Δ_i be the set of probability distributions on S_i. Elements of Δ_i are of the form $p_i : S_i \to \mathbb{R}$, where $\sum_{s_{ij} \in S_i} p_i(s_{ij}) = 1$ and $p_i(s_{ij}) \geq 0$ for all $s_{ij} \in S_i$. For convenience, use the notation $p_{ij} = p_i(s_{ij})$. We write the set of mixed strategy profiles by $\Delta = \prod_{i \in N} \Delta_i$ and denote points in Δ by $p = (p_1, ..., p_n)$. Therefore, given a mixed strategy profile p, player i's expected payoff is $u_i(p) = \sum_{s \in S} p(s)u_i(s)$, where $p(s) = \prod_{i \in N} p_i(s_i)$, where $s_i \in S_i$ denotes the ith element of s. For convenience, for each $i \in N$ and $j \in \{1, ..., J_i\}$, denote by $u_{ij}(p)$ the expected payoff to player i adopting pure strategy s_{ij} when the other players adopt their components of p. The space of payoff vectors of player i's pure strategies is $\mathbb{R}^{J_i}$, and write $\mathbb{R}^{\sum J_i} = \prod_{i \in N} \mathbb{R}^{J_i}$. Define the function $\bar{u} : \Delta \to \mathbb{R}^{\sum J_i}$ by $\bar{u}(p) = (\bar{u}_1(p), ..., \bar{u}_n(p))$, where $\bar{u}_i(p) = (u_{i1}(p), ..., u_{iJ_i}(p))$.

It is assumed that for each pure strategy s_{ij}, there is an additional payoff disturbance ε_{ij}, and we denote the noisy payoff by

$$\tilde{u}_{ij}(p) = u_{ij}(p) + \varepsilon_{ij} \tag{1.2.1}$$

Player i's noise vector, $\varepsilon_i = (\varepsilon_{i1}, ..., \varepsilon_{iJ_i})$, is distributed according to a joint distribution with density function $f_i(\varepsilon_i)$. $f = (f_1, ..., f_n)$ is called *admissible* (McKelvey and Palfrey, 1995; Goeree et al., 2005) if

(a) the marginal distribution of f_i exists for each ε_{ij},
(b) disturbances are independent across players (not necessarily across strategies),
(c) $E(\varepsilon_i) = 0$ for all $i \in N$.

Define $B_{ij}(\bar{u}_i)$ to be the set of ε_i such that strategy s_{ij} has the highest disturbed payoff, i.e.,

$$B_{ij}(\bar{u}_i) = \{\varepsilon_i \in \mathbb{R}^{J_i} | u_{ij} + \varepsilon_{ij} \geq u_{ik} + \varepsilon_{ik}, \forall k = 1, ..., J_i\} \tag{1.2.2}$$

Therefore, for given $\bar{u}_i$, player i selects s_{ij} with probability

$$\sigma_{ij}(\bar{u}_i) = \int_{B_{ij}(\bar{u}_i)} f(\varepsilon) d\varepsilon \tag{1.2.3}$$

$\sigma_i : \mathbb{R}^{J_i} \to \Delta_i$ defined by Eq.(1.2.3) is called the *structural quantal response function* of player i (Goeree et al., 2005). For any admissible $f(\varepsilon)$ with a full support condition [4], σ_i satisfies

(i) Interiority: $\sigma_{ij}(\bar{u}_i) > 0$ for all $j \in \{1, ..., J_i\}$ and $\bar{u}_i \in \mathbb{R}^{J_i}$.
(ii) Continuity: $\sigma_{ij}(\bar{u}_i)$ is a continuous and differentiable function for all $\bar{u}_i \in \mathbb{R}^{J_i}$.
(iii) Responsiveness: $\frac{\partial \sigma_{ij}(\bar{u}_i)}{\partial u_{ij}} > 0$ for all $j \in \{1, ..., J_i\}$ and $\bar{u}_i \in \mathbb{R}^{J_i}$.

If the payoff disturbances are interchangeable [5], i.e., $f_i(\varepsilon_{i1}, ..., \varepsilon_{iJ_i}) = f_i(\varepsilon_{i\psi(1)}, ..., \varepsilon_{i\psi(J_i)})$ for any permutation ψ, σ_i also satisfies

(iv) Monotonicity: $u_{ij} > u_{ik} \Rightarrow \sigma_{ij}(\bar{u}_i) > \sigma_{ik}(\bar{u}_i)$ for all $j, k \in \{1, ..., J_i\}$.

On the other hand, any function $\sigma_i : \mathbb{R}^{J_i} \to \Delta_i$ that satisfies (i)-(iv) is called a *regular quantal response function* of player i (Goeree et al., 2005). One well known example is the logistic response function

$$\sigma_{ij}(\bar{u}_i) = \frac{e^{\lambda u_{ij}}}{\sum_{k=1}^{J_i} e^{\lambda u_{ik}}} \tag{1.2.4}$$

where $\frac{1}{\lambda}$ has been interpreted as the intensity of noises (McKelvey and Palfrey, 1995; Hofbauer and Sandholm, 2002, 2007; Turocy, 2005). Eq.(1.2.4) arises from Eq.(1.2.3) if all the noises follow the extreme value distribution with cumulative distribution function $\exp(-\exp(-\lambda\varepsilon_{ij} - \gamma)))$, where γ is Euler's constant. There are also many regular quantal response functions that cannot be derived by the structural approach. For instance, see Eq.(6.1), Eq.(6.2) and Proposition 6 in Goeree et al., 2005.

Following the logistic response function, consider the quantal response function as a function of the noise level λ

$$\bar{\sigma} : \mathbb{R}^{\sum J_i} \times [0, +\infty) \to \Delta \tag{1.2.5}$$

[4] Full support condition says that $f(\varepsilon) > 0$ for any $\varepsilon \in R^{\sum J_i}$. Without full support, e.g., uniformly distributed disturbances, the inequalities in (i) and (iii) hold only weakly (Goeree et al., 2005).

[5] A special case of interchangeable random variables is i.i.d.

with $\bar{\sigma}(\bar{u}, \lambda) = \sigma(\lambda\bar{u})$, where $\lambda = 0$ means full noise and $\lambda = +\infty$ means no noise (Goeree et al., 2005). For convenience, we use the abusive notation σ to denote $\bar{\sigma}$. For given $\lambda \geq 0$, a *quantal response equilibrium* (QRE) is any $p \in \Delta$ such that for each $i \in N$ and $j \in \{1, ..., J_i\}$,

$$p_{ij} = \sigma_{ij}(\lambda\bar{u}_i(p)) \tag{1.2.6}$$

Denote the set of QRE at noise level λ by $\pi_\lambda = \{p \in \Delta | p_{ij} = \sigma_{ij}(\lambda\bar{u}_i(p))\}$.

In the rest of this section, we focus on the quantal response function Eq.(1.2.5) and investigate the properties of π_λ. Theorem 1.1 indicates that π_λ is nonempty for any continuous σ. If σ is Lipschitz continuous in a neighborhood of $\mathbf{0}$, Theorem 1.2 asserts that for sufficiently small λ, not only the existence but also the uniqueness of QRE can be guaranteed. If σ is monotonic, Theorem 1.3 claims that π_0 consists of only the centroid of Δ when $\lambda = 0$. Finally, Theorem 1.4 says that QRE approach Nash equilibria of the game when $\lambda \to +\infty$ if σ has cumulativity.

Theorem 1.1

If σ is continuous, there exists a QRE for any $\lambda \geq 0$. [6]

Proof

This result follows from Brouwer's fixed point theorem, since $\sigma \circ \bar{u}$ is continuous. □

For given λ, Theorem 1.1 says that a QRE exists for any continuous random disturbance, but the maximum number of QRE is unclear. Surprisingly, even for two-person games, π_λ may include infinite number of QRE. Consider a 2×2 bimatrix game $\begin{pmatrix} a_1, b_1 & 0, 0 \\ 0, 0 & a_2, b_2 \end{pmatrix}$ with (Lipschitz) continuous quantal response function

$$\sigma_{i1}(u_{i1}, u_{i2}, \lambda) = \begin{cases} 0 & \lambda(u_{i1} - u_{i2}) \leq -\frac{1}{2} \\ \frac{1}{2} + \lambda(u_{i1} - u_{i2}) & -\frac{1}{2} \leq \lambda(u_{i1} - u_{i2}) \leq \frac{1}{2} \\ 1 & \frac{1}{2} \leq \lambda(u_{i1} - u_{i2}) \end{cases} \tag{1.2.7}$$

For small λ such that $|\lambda(u_{i1} - u_{i2})| \leq \frac{1}{2}$, QRE are the solutions of

$$\begin{aligned} p_{11} &= \frac{1}{2} + \lambda(b_1 p_{21} - b_2(1 - p_{21})) \\ p_{21} &= \frac{1}{2} + \lambda(a_1 p_{11} - a_2(1 - p_{11})) \end{aligned} \tag{1.2.8}$$

[6]By applying Brouwer's fixed point theorem, McKelvey and Palfrey (1995) pointed out that a sufficient condition for the existence of a QRE is admissibility. Similar, Goeree et al. (2005) proved the existence of a QRE for regular quantal response functions.

Substituting p_{21} in the first equation of Eq.(1.2.8) by the second equation and p_{11} in the second equation of Eq.(1.2.8) by the first equation,

$$\begin{aligned} p_{11}(1-\lambda^2(b_1+b_2)(a_1+a_2)) &= \frac{1}{2}+\lambda\frac{b_1+b_2}{2}-\lambda b_2-\lambda^2(b_1+b_2)a_2 \\ p_{21}(1-\lambda^2(b_1+b_2)(a_1+a_2)) &= \frac{1}{2}+\lambda\frac{a_1+a_2}{2}-\lambda a_2-\lambda^2(a_1+a_2)b_2 \end{aligned} \quad (1.2.9)$$

Thus, for given $\lambda > 0$, if the payoff matrix satisfies

$$\begin{aligned} &\lambda^2(b_1+b_2)(a_1+a_2)=1 \\ &\frac{1}{2}+\lambda\frac{b_1+b_2}{2}=\lambda b_2+\lambda^2(b_1+b_2)a_2 \\ &\frac{1}{2}+\lambda\frac{a_1+a_2}{2}=\lambda a_2+\lambda^2(a_1+a_2)b_2 \end{aligned} \quad (1.2.10)$$

any $(p_{11}, p_{21}) \in \Delta$ is a solution of Eq.(1.2.8). This implies that $\pi_\lambda = \Delta$. (See Example 1.1)

The quantal response function (1.2.7) is not regular, but one can easily regularize it by adding small perturbations. In a similar way, it is possible to construct a quantal response function such that π_λ is countably infinite.

Example 1.3

Consider a 2×2 bimatrix coordination game $\begin{pmatrix} \frac{1}{3},\frac{2}{3} & 0,0 \\ 0,0 & \frac{2}{3},\frac{1}{3} \end{pmatrix}$ with quantal response function (1.2.7). If $0 \le \lambda < 1$, the game has a unique QRE, $p_{11} = \frac{2\lambda+3}{6(\lambda+1)}$, $p_{21} = \frac{4\lambda+3}{6(\lambda+1)}$. If $\lambda > 1$, the game has three QRE, $p_{11} = p_{21} = 0$, $p_{11} = p_{21} = 1$ and $p_{11} = \frac{2\lambda+3}{6(\lambda+1)}$, $p_{21} = \frac{4\lambda+3}{6(\lambda+1)}$. If $\lambda = 1$, from Eq.(1.2.10), any $(p_{11}, p_{21}) \in \Delta$ is a QRE. (See Figure 1.2.1)

Although the sets of QRE can be very complicated, next three theorems indicate that π_λ has good properties for limit cases $\lambda \to 0$ and $\lambda \to +\infty$.

Theorem 1.2

If σ is Lipschitz continuous in a neighborhood $B_\delta(\mathbf{0})$ of $\mathbf{0}$, π_λ is a singleton for sufficiently small λ. [7]

Proof

For given λ, define

$$\sigma \circ \bar{u}(p) = \sigma(\lambda \bar{u}(p)) \quad (1.2.11)$$

[7]McKelvey and Palfrey (1995) proved this theorem for the logit equilibrium by the same technique. Ui (2006) provided a sufficient condition for the uniqueness of QRE in 2×2 symmetric games with Gaussian noises.

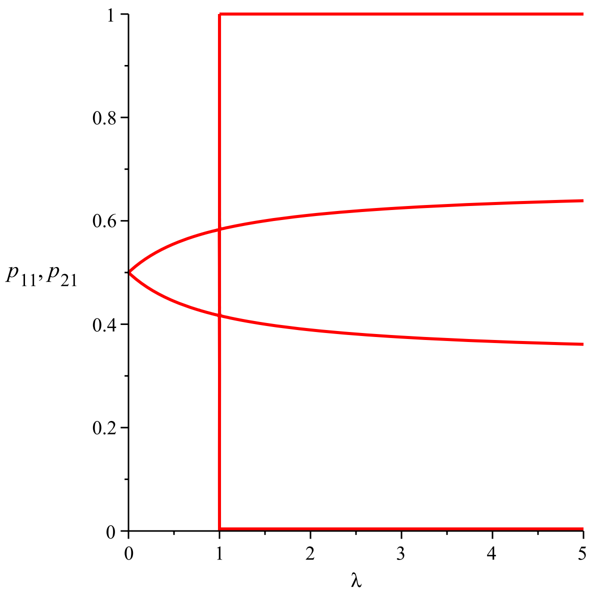


Figure 1.2.1: The graph of the QRE correspondence for Example 1.1.

From the definition of QRE, $p \in \pi_\lambda$ if and only if p is a fixed point of $\sigma \circ \bar{u}$. We will show that for sufficiently small λ, $\sigma \circ \bar{u}$ has a unique fixed point. Notice that σ is Lipschitz continuous in $B_\delta(0)$ and $u_{ij}(p)$ is smooth, there are $S > 0$ and $T > 0$ such that

$$\begin{aligned} &\|\sigma \circ \bar{u}(p) - \sigma \circ \bar{u}(q)\| = \max_{ij} |\sigma_{ij}(\bar{\lambda}\bar{u}_i(p)) - \sigma_{ij}(\bar{\lambda}\bar{u}_i(q))| \\ &\leq \bar{\lambda} S \max_{ij} |u_{ij}(p) - u_{ij}(q)| \leq \bar{\lambda} ST \max_{ij} |p_{ij} - q_{ij}| = \bar{\lambda} ST \|p - q\| \end{aligned} \tag{1.2.12}$$

for any $p, q \in \Delta$, where $\|\cdot\|$ represents the sup norm, and $\bar{\lambda}$ is picked to satisfy $\bar{\lambda} ST \leq 1$ and $\bar{\lambda}\|\bar{u}(p)\|, \bar{\lambda}\|\bar{u}(q)\| < \delta$ for any $p, q \in \Delta$. This implies that $\sigma \circ \bar{u}$ is a contraction mapping for $\lambda \leq \bar{\lambda}$. From the Banach fixed-point theorem, it has a unique fixed point. □

Theorem 1.2 extends the existence of a QRE to the case where σ may not be (globally) continuous. (See Example 1.2)

Example 1.2

Consider a 2×2 zero-sum game $\begin{pmatrix} -1, 1 & 0, 0 \\ 0, 0 & -2, 2 \end{pmatrix}$ with discretely distributed noises $\Pr(\varepsilon = -1) = \Pr(\varepsilon = 1) = \frac{1}{2}$. The quantal response function is written as

$$\sigma_{i1}(u_{i1}, u_{i2}, \lambda) = \begin{cases} 0 & \lambda(u_{i1} - u_{i2}) \leq -1 \\ \frac{1}{2} & -1 \leq \lambda(u_{i1} - u_{i2}) \leq 1 \\ 1 & 1 \leq \lambda(u_{i1} - u_{i2}) \end{cases} \tag{1.2.13}$$

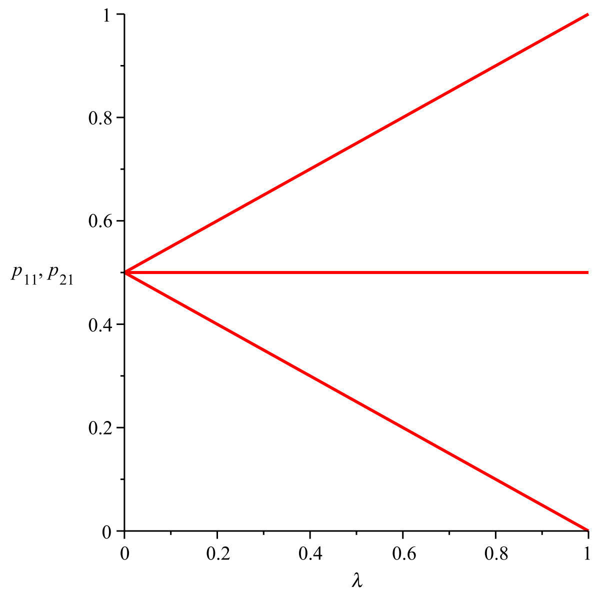


Figure 1.2.2: The graph of the QRE correspondence for Example 1.3.

It is easy to verify that the game has a unique QRE, $p_{11} = p_{21} = \frac{1}{2}$ if $\lambda < 2$, but no QRE if $\lambda \geq 2$.

Notice that the differentiability in (ii) implies that σ_i is absolutely continuous for each u_{ij}, it is natural to ask whether the Lipschitz continuity condition in Theorem 1.2 can be relaxed. However, Example 1.3 shows that it cannot be replaced by the absolute continuity.

Example 1.3

Consider a 2×2 symmetric game $\begin{pmatrix} 1,1 & 0,0 \\ 0,0 & 1,1 \end{pmatrix}$ with the absolutely continuous quantal response function

$$\sigma_{i1}(u_{i1}, u_{i2}, \lambda) = \begin{cases} 0 & \lambda(u_{i1} - u_{i2}) \leq -1 \\ \frac{1}{2} - \frac{\sqrt{-\lambda(u_{i1}-u_{i2})}}{2} & -1 \leq \lambda(u_{i1} - u_{i2}) \leq 0 \\ \frac{1}{2} + \frac{\sqrt{\lambda(u_{i1}-u_{i2})}}{2} & 0 \leq \lambda(u_{i1} - u_{i2}) \leq 1 \\ 1 & 1 \leq \lambda(u_{i1} - u_{i2}) \end{cases} \tag{1.2.14}$$

For any $0 \leq \lambda \leq 1$, π_λ includes three QRE, $p_{11} = p_{21} = \frac{1-\lambda}{2}$, $p_{11} = p_{21} = \frac{1}{2}$ and $p_{11} = p_{21} = \frac{1+\lambda}{2}$. (See Figure 1.2.2)

Theorem 1.3

If σ is monotonic, π_0 consists of only the centroid of Δ, i.e., $p_{ij} = \frac{1}{J_i}$ for all $i \in N$ and $j \in \{1, ..., J_i\}$.

Proof

Monotonicity and continuity implies that $\sigma_{ij}(\mathbf{0}) = \sigma_{ik}(\mathbf{0})$ for all $i \in N$ and $j, k \in \{1, ..., J_i\}$. $\square$

In particular, if σ is structural, monotonicity can be relaxed to interchangeability (Goeree et al., 2005, Proposition 5).

Theorem 1.4a

Let $p^\lambda \in \pi_\lambda$. If σ is structural and $\lim\limits_{\lambda \to +\infty} p^\lambda = p^*$, p^* must be a Nash equilibrium. [8]

Proof

If p^* is not a Nash equilibrium, there are $i \in N$ and $j, k \in \{1, ..., J_i\}$ such that $p^*_{ij} > 0$ and $u_{ik}(p^*) > u_{ij}(p^*)$. Since $\bar{u}_i$ is continuous, it follows that for sufficiently small ϵ, there is a Λ such that for $\lambda > \Lambda$, $u_{ik}(p^\lambda) > u_{ij}(p^\lambda) + \epsilon$. As $\lambda \to +\infty$, we have $p^\lambda_{ij} = \sigma_{ij}(\lambda \bar{u}_i(p^\lambda)) \leq \int_{\lambda u_{ij}(p^\lambda) + \varepsilon_{ij} > \lambda u_{ik}(p^\lambda) + \varepsilon_{ik}} f(\varepsilon) d\varepsilon \to 0$. This contradicts $p^*_{ij} > 0$. $\square$

Theorem 1.4a says that the limit set of QRE as $\lambda \to +\infty$ includes a Nash equilibrium for any structural quantal response function. [9] However, a surprising fact is that the limit set may not contain any Nash equilibrium even if σ is regular. For instance, suppose that σ is the logistic response function and define $\theta_i = \frac{\sigma_i}{2} + \frac{1}{2J_i}$ for all $i \in N$. It is obvious that θ_i is regular but $\theta_{ij} > \frac{1}{2J_i}$ for all $j \in \{1, ..., J_i\}$. Therefore, if the unique Nash equilibrium of the game has a component $p_{ij} < \frac{1}{2J_i}$, it can not be included in the limit set. In order to provide a sufficient condition of Theorem 1.4a for non-structural quantal response functions, we introduce a new property *cumulativity* (the name is borrowed from the cumulative distribution function).

(v) Cumulativity: $u_{ij} > u_{ik} \Rightarrow \lim\limits_{\lambda \to \infty} \frac{\sigma_{ik}(\lambda \bar{u}_i)}{\sigma_{ij}(\lambda \bar{u}_i)} = 0$ for all $i \in N$ and $j, k \in \{1, ..., J_i\}$.
The intuition is that strategies with lower payoffs will not be used as noises go to zero.

Theorem 1.4b

Let $p^\lambda \in \pi_\lambda$. If σ is cumulative and $\lim\limits_{\lambda \to +\infty} p^\lambda = p^*$, p^* must be a Nash equilibrium.

Proof

If p^* is not a Nash equilibrium, there are $i \in N$ and $j, k \in \{1, ..., J_i\}$ such that

[8] McKelvey and Palfrey (1995) proved this theorem for the logit equilibrium.

[9] The limit set may not include all Nash equilibria of the game. For an example, see subsection 2.5.2 in Chapter 2.

$p^*_{ij} > 0$ and $u_{ik}(p^*) > u_{ij}(p^*)$. But from cumulativity and continuity, $\lim_{\lambda\to\infty} \frac{\sigma_{ij}(\lambda \bar{u}_i(p^\lambda))}{\sigma_{ik}(\lambda \bar{u}_i(p^\lambda))} = \lim_{\lambda\to\infty} \frac{\sigma_{ij}(\lambda \bar{u}_i(p^*))}{\sigma_{ik}(\lambda \bar{u}_i(p^*))} = 0$. This contradicts $p^*_{ij} > 0$. □

Notice that all structural quantal response functions are cumulative, Theorem 1.4a is a special case of Theorem 1.4b.

1.3 Equilibrium selection in normal form games

In this section, we study a particular class of quantal response functions that satisfy Theorems 1.1, 1.3 and 1.4b, i.e., $\sigma : \mathbb{R}^{\sum J_i} \times [0, +\infty) \to \Delta$ is continuous, monotonic and cumulative. Our purpose is to define an equilibrium selection by "tracing" the graph of the QRE correspondence.

Denote the graph of the QRE correspondence by $\pi = \{(\lambda, p) | \lambda \geq 0, p \in \pi_\lambda\}$. Theorem 1.5 shows that for all normal form games, the QRE at $\lambda = 0$ is connected by a component of π to at least one Nash equilibrium. If the quantal response function is C^2 continuous, Theorem 1.6 indicates that for almost all games, this component is diffeomorphic to a C^1 segment. This implies that the graph of the QRE correspondence contains a unique branch which starts for $\lambda = 0$ at the centroid and converges to a unique Nash equilibrium as λ goes to infinity.

Theorem 1.5

π includes a component T_π that connects the centroid and a Nash equilibrium. [10]

Proof

Let us make the transformation $\lambda = \frac{\gamma}{1-\gamma}$ and define the mapping

$$\sigma \circ \bar{u} : \Delta \times [0, 1) \to \Delta \tag{1.3.1}$$

with $\sigma \circ \bar{u}(p, \gamma) = \sigma(\frac{\gamma}{1-\gamma}\bar{u}(p))$. For given γ, denote the set of QRE by $\tilde{\pi}_\gamma = \{p | p_{ij} = \sigma_{ij}(\frac{\gamma}{1-\gamma}\bar{u}_i(p))\}$ and the graph of QRE by $\tilde{\pi}$. Clearly, $(\gamma, p) \in \tilde{\pi}$ if and only if $\sigma(\frac{\gamma}{1-\gamma}\bar{u}(p)) = p$. From Browder's Theorem (e.g., Mas-Colell, 1974, Theorem 1), for any given $0 < \gamma < 1$, there is a component T of $\tilde{\pi}$ such that $T \cap \Delta \times \{0\} \neq \emptyset$ and $T \cap \Delta \times \{\gamma\} \neq \emptyset$. For $n \in \mathbb{N}$, denote the component for $\gamma = 1 - \frac{1}{n}$ by T_n. By Mas-Colell (1990, Theorem A.5.1.(ii) page 10, see also Jean-Jacques Herings, 2002, Theorem 4.3), the closed limit of the sequence T_n, denoted by T_π, is compact and connected. From Theorem 1.4b, T_π must include a Nash equilibrium. □

Theorem 1.6

[10] Jean-Jacques Herings (2002) proved this theorem for the logit equilibrium by the same technique.

If σ is C^2, for almost all games, π includes a unique branch that starts at the centroid as $\lambda = 0$ and converges to a Nash equilibrium as $\lambda \to +\infty$. [11]

Proof

Define

$$F(p, \lambda, u) = \sigma(\lambda \bar{u}(p)) - p \tag{1.3.2}$$

where $u \in \mathbb{R}^{n \prod J_i}$ denotes the payoff matrix. For given u, write $F_u(p, \lambda) = F(p, \lambda, u)$. Clearly, $(\lambda, p) \in \pi$ if and only if $F_u(p, \lambda) = 0$.

The Transversality Theorem (Mas-Colell, 1990, Proposition 8.3.1 page 320) says that if F is C^2 (the factor 2 comes from $1 + \dim(\lambda)$) and $DF(p, \lambda, u)$ has rank $\sum_{i=1}^{n} J_i - n$ whenever $F(p, \lambda, u) = 0$, then for almost all u, $DF_u(p, \lambda)$ has rank $\sum_{i=1}^{n} J_i - n$ whenever $F_u(p, \lambda) = 0$. This implies that 0 is a regular value of $F_u(p, \lambda)$ for almost all u.

We next calculate the rank of $DF(p, \lambda, u)$.

$$\begin{aligned} DF(p, \lambda, u) &= (\frac{\partial F}{\partial p}, \frac{\partial F}{\partial \lambda}, \frac{\partial F}{\partial u}) \\ &= (\frac{\partial \sigma(\lambda \bar{u}(p))}{\partial p} - I, \frac{\partial \sigma(\lambda \bar{u}(p))}{\partial \lambda}, \frac{\partial \sigma(\lambda \bar{u}(p))}{\partial u}) \\ &= (-I, \frac{\partial \sigma(\lambda \bar{u}(p))}{\partial \lambda}, 0) \\ &\quad + \lambda \sum_{i \in N} \sum_{j=1}^{J_i} (\frac{\partial \sigma(\lambda \bar{u})}{\partial u_{ij}} \frac{\partial u_{ij}}{\partial p}, 0, \frac{\partial \sigma(\lambda \bar{u})}{\partial u_{ij}} \frac{\partial u_{ij}}{\partial u}) \end{aligned} \tag{1.3.3}$$

where I is the $(\sum_{i=1}^{n} J_i - n) \times (\sum_{i=1}^{n} J_i - n)$ unit matrix.

We use the notation (s_{ij}, s_{-i}) to represent the pure strategy profile that player i adopts the strategy s_{ij} and all other players adopt their component of s_{-i}, where $s_{-i} = (s_1, ..., s_{i-1}, s_{i+1}, ..., s_n) \in \prod_{j \neq i} S_j = S_{-i}$. u_{ij} is then written as

$$u_{ij} = \sum_{s_{-i} \in S_{-i}} u_i(s_{ij}, s_{-i}) \prod_{t \neq i} p_t(s_t) \tag{1.3.4}$$

Since $p_{kJ_k} = 1 - \sum_{l=1}^{J_k - 1} p_{kl}$,

$$\begin{aligned} \frac{\partial u_{ij}}{\partial p_{kl}} &= \sum_{s_k = s_{kl}} u_i(s_{ij}, s_{-i}) \prod_{t \neq i,k} p_t(s_t) - \sum_{s_k = s_{kJ_k}} u_i(s_{ij}, s_{-i}) \prod_{t \neq i,k} p_t(s_t) \\ &= \sum_{s_k = s_{kl}} u_i(s_{ij}, s_{-i}) \sum_{l'=1}^{J_k} \frac{\partial u_{ij}}{\partial u_i(s_{kl'}, s_{-k})} - \sum_{s_k = s_{kJ_k}} u_i(s_{ij}, s_{-i}) \sum_{l'=1}^{J_k} \frac{\partial u_{ij}}{\partial u_i(s_{kl'}, s_{-k})} \end{aligned} \tag{1.3.5}$$

[11] McKelvey and Palfrey (1995) proved this theorem for the logit equilibrium by applying Sard's Theorem.

This implies that $\frac{\partial\sigma(\lambda\bar{u}(p))}{\partial p_{ij}}$ is a linear combination of columns of $\frac{\partial\sigma(\lambda\bar{u}(p))}{\partial u}$. Hence, rank $DF(p,\lambda,u) = \text{rank }(-I, \frac{\partial\sigma(\lambda\bar{u}(p))}{\partial\lambda}, \frac{\partial\sigma(\lambda\bar{u}(p))}{\partial u}) = \sum_{i=1}^{n} J_i - n$.

Applying the Transversality Theorem, for almost all u, 0 is a regular value of $F_u(p,\lambda)$. $F_u^{-1}(0)$ is then a C^1 one-dimensional manifold, which is diffeomorphic to a segment or a circle (Milnor, 1965, Lemma 4; Mas-Colell, 1974, Theorem 2). From Theorem 1.2 and Theorem 1.5, it is a segment that starts from the centroid as $\lambda = 0$ and converges to a Nash equilibrium as $\lambda \to +\infty$. □

Theorem 1.6 implies that for almost all normal form games, we can define a unique selection from the set of Nash equilibria by "tracing" the graph of the QRE correspondence beginning at the centroid of the strategy simplex (from Theorem 1.3, it is the unique solution when $\lambda = 0$) and continuing for larger and larger values of λ. [12] For given σ, we call the selected Nash equilibrium the *limiting QRE* of the game.

1.4 Equilibrium selection in symmetric games

This section studies the limiting QRE in n-person symmetric games. Theorem 1.7 points out that the limiting QRE of a symmetric game must be a symmetric Nash equilibrium. A normal form game is called symmetric if the players have identical strategy sets and payoff functions. That is, $S_i = S_j$ for all $i,j \in N$ and $u_i(s_1,...,s_n) = u_{\psi(i)}(s_{\psi(1)},...,s_{\psi(n)})$ for any permutation ψ and $s \in S$ (Dasgupta and Maskin, 1986). Denote an n-person symmetric game by $(N,\hat{S},\hat{u})$. For each player $i \in N$, $\hat{S} = \{\hat{s}_1,...,\hat{s}_J\}$ is the strategy set and $\hat{u} : \hat{S} \times \Box^J \to \mathbb{R}$ is the payoff function. Elements of $\Box^J$ are of the form $\hat{q} : \hat{S}^{n-1} \to \mathbb{N}_0^J$, where $\sum_{k=1}^{J} \hat{q}_k(s^{n-1}) = n-1$. Intuitively, $\hat{q}_k$ calculates the number of pure strategy $\hat{s}_k$ in the strategy profiles s^{n-1}. Therefore, payoff to a player using pure strategy $\hat{s}_i$ when the others adopt s^{n-1} is $\hat{u}(\hat{s}_i, \hat{q}(s^{n-1}))$.

Following the notations in section 1.2, $u_{ij}(p) = \sum_{s_i=\hat{s}_j} \prod_{k\neq i} p_k(s_k)\hat{u}(\hat{s}_j,\hat{q}(s_{-i}))$, where $s_{-i} \in \hat{S}^{n-1}$. QRE at noise level λ are the solutions of

$$p_{ij} = \sigma_{ij}(\lambda\bar{u}_i(p)) \tag{1.4.1}$$

Suppose that players have the identical quantal response function, i.e., $\sigma_{ij}(\lambda_i\bar{u}_i) = \sigma_{kj}(\lambda_k\bar{u}_k)$ if $\lambda_i\bar{u}_i = \lambda_k\bar{u}_k$ for all $i,k \in N$ and $j \in \{1,...,J\}$. Denote it by $\hat{\sigma} : \mathbb{R}^J \times [0,+\infty) \to \Delta^J$, where Δ^J is the set of probability distributions on $\hat{S}$. Eq.(1.4.1) is then written as

$$p_{ij} = \hat{\sigma}_j(\lambda\bar{u}_i(p)) \tag{1.4.2}$$

[12] As pointed out by Turocy (2005), the branch may have turning points, leading to intervals on which λ is decreasing while following the branch in the direction from the centroid at $\lambda = 0$ to the limiting Nash equilibrium. For the logit equilibrium, there are at most a finite number of turning points (Turocy, 2005). However, following the idea in section 1.2, it is possible to construct a quantal response function such that the branch has infinite turning points.

A QRE is called *symmetric* if $p_i = p_j$ for all $i, j \in N$. We next show that all QRE on T_π (which is the component of π that connects the centroid and a Nash equilibrium) are symmetric if T_π is diffeomorphic to a segment. Consider the equation

$$\hat{p}_i = \hat{\sigma}_i(\lambda \hat{u}(\hat{p})) \tag{1.4.3}$$

where $\hat{p} \in \Delta^J$, $\hat{p}_i = \hat{p}(\hat{s}_i)$ and $\hat{u}_j(\hat{p}) = \sum_{s^{n-1} \in \hat{S}^{n-1}} \prod_{k=1}^{n-1} \hat{p}(s_k^{n-1}) \hat{u}(\hat{s}_j, \hat{q}(s^{n-1}))$. For any given λ, it is easy to see that $\hat{p}$ is a solution of Eq.(1.4.3) if and only if $p_i = \hat{p}$ for all $i \in N$ is a solution of Eq.(1.4.2). Denote the set of symmetric QRE at noise level λ by $\hat{\pi}_\lambda$. Existence of a symmetric QRE follows from Brouwer's fixed point theorem. Define the graph of symmetric QRE correspondence by $\hat{\pi} = \{(\lambda, p) | \lambda \geq 0, p \in \hat{\pi}_\lambda\}$. Similarly as Theorem 1.5, it is easy to prove that $\hat{\pi}$ includes a component $T_{\hat{\pi}}$ that connects the centroid and a symmetric Nash equilibrium by applying Browder's Theorem. Since $\hat{\pi} \subseteq \pi$, we have $T_{\hat{\pi}} \subseteq T_\pi$. Therefore, in the case that T_π is a segment, $T_{\hat{\pi}}$ must also be a segment and $T_{\hat{\pi}} = T_\pi$. All QRE on T_π are then symmetric. As a consequence, we have Theorem 1.7.

Theorem 1.7

The limiting QRE of a symmetric game must be a symmetric Nash equilibrium.

Theorem 1.7 implies that if a symmetric game has a unique symmetric Nash equilibrium, the limiting QRE must be that equilibrium. However, the uniqueness of a symmetric Nash equilibrium does not imply the existence of the limiting QRE (see Example 1.4). Since the set of payoff functions for n-person symmetric games has Lebesgue measure zero in $R^{\sum J}$, the existence of the limiting QRE for almost all symmetric games is problematic. In particular, Theorem 2.1 in Chapter 2 indicates that the limiting QRE does not exist for any 2×2 symmetric game with a unique interior symmetric Nash equilibrium. Thus, the equilibrium selection is not well defined for one fourth of all 2×2 symmetric games.

Example 1.4

Consider a 2×2 symmetric game $\begin{pmatrix} 0,0 & 1,1 \\ 1,1 & 0,0 \end{pmatrix}$ with the logistic response function. QRE are the solutions of

$$\begin{aligned} p_{11} &= \frac{1}{1 + e^{\lambda(2p_{21}-1)}} \\ p_{21} &= \frac{1}{1 + e^{\lambda(2p_{11}-1)}} \end{aligned} \tag{1.4.4}$$

The game has a unique symmetric Nash equilibrium $p_{11} = p_{21} = \frac{1}{2}$ and two asymmetric Nash equilibria $p_{11} = 1, p_{21} = 0$ and $p_{11} = 0, p_{21} = 1$. For any given λ, this is a unique symmetric QRE, $p_{11} = p_{21} = \frac{1}{2}$. However, the graph of QRE correspondence has a bifurcation point and tracing the branch beginning at the centroid could reach all three Nash

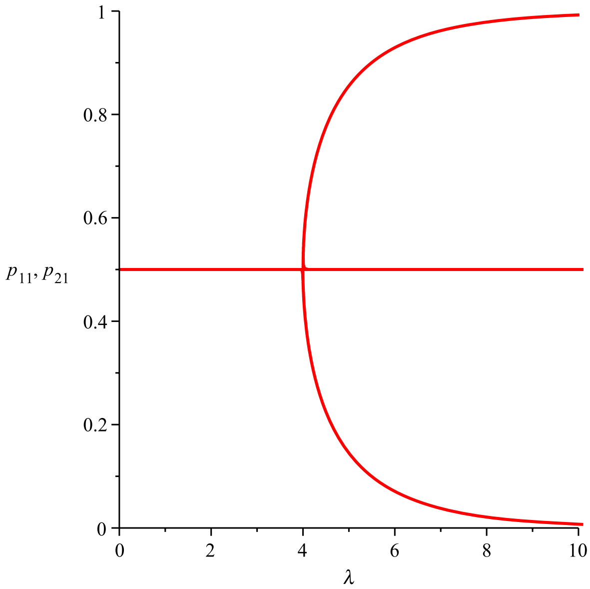


Figure 1.4.1: The graph of the QRE correspondence for Example 1.4.

equilibria. (See Figure 1.4.1) Hence, the equilibrium selection is not well defined.

Since asymmetric Nash equilibria cannot be the limiting QRE in symmetric games, we restrict the "tracing process" to the graph of symmetric QRE correspondence $\hat{\pi}$. Redefine the limiting QRE in symmetric games as the unique limiting point of $T_{\hat{\pi}}$ instead of T_{π}. Similarly as Theorem 1.6, we have the following theorem.

Theorem 1.8

If $\hat{\sigma}$ is C^2, for almost all symmetric games, $\hat{\pi}$ contains a unique branch that starts at the centroid as $\lambda = 0$ and converges to a symmetric Nash equilibrium as $\lambda \to +\infty$.

Under the new definition, the limiting QRE exists for almost all symmetric games. Moreover, if a symmetric game has a unique symmetric Nash equilibrium, it is the limiting QRE no matter how many asymmetric Nash equilibria the game has. As a consequence, following two corollaries can be obtained directly.

Corollary 1.1

If a two-person symmetric game has an interior ESS, it is the limiting QRE. [13]

---

[13]This follows from the fact that if a symmetric game has an interior ESS, then it is the unique symmetric Nash equilibrium (Hofbauer and Sigmund, 1998).

Corollary 1.2

The limiting QRE of an anti-coordination game is the interior Nash equilibrium. [14]

1.5 Two-person symmetric games

Consider a $J \times J$ two-person symmetric game. Let a_{ij} denotes the payoff to a player using strategy $\hat{s}_i$ when he meets strategy $\hat{s}_j$. Theorem 1.8 guarantees the existence of the limiting QRE for almost all two-person symmetric games. In order to decide the limiting QRE, we introduce a new concept.

Definition 1.1

A pure strategy $\hat{s}_i$ pairwise payoff dominates (PPD) another pure strategy $\hat{s}_j$ if $\sum_{k \in J_{ij}} a_{ik} > \sum_{k \in J_{ij}} a_{jk}$ for any $\{i, j\} \subseteq J_{ij} \subseteq \{1, ..., J\}$. If $\hat{s}_i$ PPD $\hat{s}_j$ for any $\hat{s}_j \in \hat{S}$ and J_{ij}, $\hat{s}_i$ is globally pairwise payoff dominant (GPPD).

Theorem 1.9

Suppose that σ is continuous, monotonic and cumulative. In two-person symmetric games, the limiting QRE is the GPPD Nash equilibrium if it exists.

Proof

Suppose that $\hat{s}_i$ is the GPPD Nash equilibrium. From Definition 1.1, $\hat{u}_i(\hat{p}) > \hat{u}_j(\hat{p})$ in region $\{\hat{p} \in \Delta^J | \hat{p}_i = \hat{p}_j \geq \hat{p}_k, \forall k \neq i, j\}$ for any $j \neq i$. Since $\hat{u}$ and σ are continuous, it follows for sufficiently small λ, $\hat{u}_i(\hat{p}) > \hat{u}_j(\hat{p})$ for any $j \neq i$ and $\hat{p} \in \hat{\pi}_\lambda$. If σ has monotonicity and cumulativity, QRE correspondence starting at the centroid will enter region $\Delta^J_{i\max} = \{\hat{p} \in \Delta^J | \hat{p}_i = \max\{\hat{p}_1, ..., \hat{p}_J\}\}$ and can not escape from it. Therefore, the QRE correspondence must converge to a Nash equilibrium in $\Delta^J_{i\max}$.

To complete the proof, it is enough to show that $\hat{p}_i = 1$ is the only Nash equilibrium in this region. For any $\hat{p} \in \Delta^J_{i\max}$, where $\hat{p}_i < 1$, suppose that $\max\{\hat{p}_1, ..., \hat{p}_{i-1}, \hat{p}_{i+1}, ..., \hat{p}_J\} = \hat{p}_j > 0$. From Definition 1.1, $\hat{u}_i(\hat{p}) > \hat{u}_j(\hat{p})$, which implies that it can not be a Nash equilibrium. As a consequence, $\hat{p}_i = 1$ is the limiting QRE. □

We next compare the limiting QRE to the solution concept of $\frac{1}{2}$ *dominant equilibrium* (Morris et al., 1995). A pure strategy profile $(\hat{s}_i, \hat{s}_i)$ is called $\frac{1}{2}$ dominant if for any $\hat{p} \in \Delta^J$ with $\hat{p}_i \geq \frac{1}{2}$, $\text{br}(\hat{p}) = \{\hat{s}_i\}$. $\frac{1}{2}$ dominant equilibrium is the sufficient condition for many equilibrium selection methods (see a literature review by Honda, 2012),

[14] A two-person symmetric game is said to have the anti-coordination property if any worst response to a mixed strategy is in the support of that mixed strategy. One famous example is the Hawk-Dove game. Kojima and Takahashi (2007) showed that every anti-coordination game has a unique symmetric Nash equilibrium. The equilibrium is in the interior of Δ.

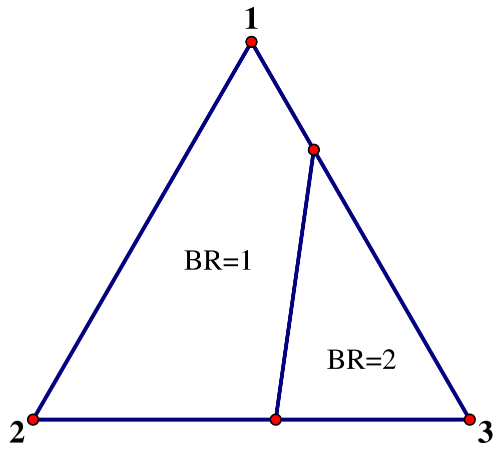


Figure 1.5.1: Best response regions for Example 1.5

e.g., the evolutionary methods (e.g., Kandori et al., 1993; Young, 1993), the potential game method (Monderer and Shapley, 1996), the global game method (Carlsson and van Damme, 1993), the incomplete information method (Kajii and Morris, 1997), the perfect foresight dynamics method (Matsui and Matsuyama, 1995), spatially dominance method (Hofbauer et al., 1997; Hofbauer, 1999).

From the definitions, a game has at most one GPPD Nash equilibrium and has at most one $\frac{1}{2}$ dominant equilibrium. If both exist, then two equilibrium must be the same. To show this, suppose that $\hat{s}_i$ is GPPD and $\hat{s}_j$ is $\frac{1}{2}$ dominant, where $i \neq j$. GPPD implies $a_{ii} + a_{ij} > a_{ji} + a_{jj}$ but $\frac{1}{2}$ dominance implies $\frac{a_{ii}+a_{ij}}{2} < \frac{a_{ji}+a_{jj}}{2}$. Here is a contradiction.

At a first glance, GPPD is more or less stronger than $\frac{1}{2}$ dominant since the GPPD Nash equilibrium strategy is the unique best response to the centroid of the strategy simplex. However, this intuition is wrong (see Example 1.5). Theorem 1.10 claims that the GPPD strategy is $\frac{1}{2}$ dominant if the symmetric game is a coordination game.

Example 1.5

Consider the following 3×3 symmetric game $\begin{pmatrix} 2 & 4 & 0 \\ 0 & 0 & 5 \\ 0 & 0 & 0 \end{pmatrix}$. Best response regions are shown in Figure 1.5.1. The first strategy is GPPD dominant but is not $\frac{1}{2}$ dominant. For instance, the second strategy is the best response to the strategy profile $(\frac{1}{2}, 0, \frac{1}{2})$.

Theorem 1.10

Suppose that the two-person symmetric game is a coordination game, i.e., $a_{ii} > a_{ji}$ for $j \neq i$. $\hat{s}_i$ is the GPPD Nash equilibrium if and only if for any $\hat{p} \in \Delta^J_{i\max}$, $\text{br}(\hat{p}) = \{\hat{s}_i\}$.

Proof

Without loss of generality, suppose that s_1 is the GPPD Nash equilibrium and $\hat{p}_1 \geq \hat{p}_2 \geq ... \geq \hat{p}_J$. For any $i \in \{2, ..., J\}$,

$$\hat{u}_1(\hat{p}) - \hat{u}_i(\hat{p}) = \sum_{j=1}^{J-1}(\hat{p}_j - \hat{p}_{j+1}) \sum_{k=1}^{j}(a_{1k} - a_{ik}) + \hat{p}_J \sum_{k=1}^{J}(a_{1k} - a_{ik}) \qquad (1.5.1)$$

From Definition 1.1, $\sum_{k=1}^{j} a_{1k} > \sum_{k=1}^{j} a_{ik}$ for $i \leq j$ and $\sum_{k=1}^{j} a_{1k} + a_{1i} > \sum_{k=1}^{j} a_{ik} + a_{ii}$ for $i > j$. Notice that $a_{ii} > a_{1i}$, we have $\sum_{k=1}^{j} a_{1k} > \sum_{k=1}^{j} a_{ik}$ for any $i, j \in \{2, ..., J\}$. This implies $\hat{u}_1(\hat{p}) - \hat{u}_i(\hat{p}) > 0$, i.e., $\text{br}(\hat{p}) = \{\hat{s}_i\}$.

On the other hand, suppose that for any $\hat{p} \in \Delta^J_{i\max}$, $\text{br}(\hat{p}) = \{\hat{s}_i\}$. In this case, for any given $j \neq i$ and $\{i, j\} \subseteq J_{ij} \subseteq \{1, ..., J\}$, we take $\hat{p}$ with $\hat{p}_k = \hat{p}_i$ for all $k \in J_{ij}$ and $\hat{p}_k = 0$ for others. Since $\hat{p} \in \Delta^J_{i\max}$, $\sum_{k \in J_{ij}} a_{ik} = \frac{\hat{u}_i(\hat{p})}{\hat{p}_i} > \frac{\hat{u}_j(\hat{p})}{\hat{p}_i} = \sum_{k \in J_{ij}} a_{jk}$. This implies that $\hat{s}_i$ is GPPD. □

From Theorem 1.10, GPPD Nash equilibrium strategy is the unique best response if it is the most frequent strategy in opponent's strategy profile, therefore must be $\frac{1}{2}$ dominant. In particular, the two conditions are equivalent in 2×2 coordination games. In 3×3 coordination games, GPPD dominant is stronger. A strategy is GPPD if and only if it is $\frac{1}{2}$ dominant and it is the unique best response to the centroid of the strategy simplex. Example 1.6 gives a 3×3 symmetric game, where the $\frac{1}{2}$ dominant equilibrium and the limiting QRE are different.

Example 1.6

Consider the following 3×3 symmetric coordination game $\begin{pmatrix} 6 & 0 & 2 \\ 0 & 5 & 3 \\ 2 & 3 & 4 \end{pmatrix}$. The first strategy is $\frac{1}{2}$ dominant and the third strategy is the best response to the centroid. Best response regions are shown in Figure 1.5.2. However, Gambit (McKelvey et al., 2010) suggests that the limiting logit equilibrium is the second strategy. (See Figure 1.5.2)

From Definition 1.1, a GPPD Nash equilibrium is also GPPD in the reduced form of the original game obtained by eliminating strictly dominated strategies. A natural question is to ask whether the limiting QRE depends only on the reduced form game. However, the answer is no. Goeree and Holt (2001) provided an example that the limiting logit equilibrium is subject to framing effects, i.e., duplicating a non Nash equilibrium strategy

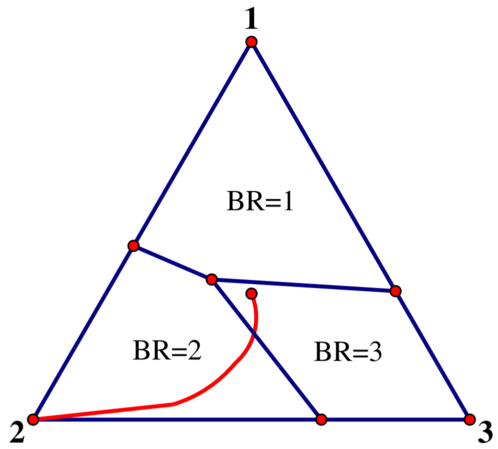


Figure 1.5.2: Best response regions and the QRE correspondence (red curve) for Example 1.6. Strategy 1 is $\frac{1}{2}$ dominant and strategy 3 is the best response to the centroid. However, strategy 2 is the limiting QRE.

which is never selected may affect the equilibrium selection (see also Goeree and Holt, 2004). More generally, Hilbe (2011) indicated that any differentiable quantal response function exhibits the framing effects. Here, we prove a much stronger result.

Theorem 1.11

Suppose that σ is continuous, monotonic and cumulative. In two-person symmetric games, any strict (symmetric) Nash equilibrium can be selected as the limiting QRE by appropriately adding a single strictly dominated strategy.

Proof

Without loss of generality, suppose that $a_{ij} > 0$ for all $i, j \in \{1, ..., J\}$ and the pure strategy profile $(\hat{s}_J, \hat{s}_J)$ is a strict Nash equilibrium. We add a new strategy $\hat{s}_{J+1}$, where $a_{iJ+1} = a_{J+1i} = 0$ for $i \in \{1, ..., J-1\}$, $a_{J+1J} = \max\{a_{1J}, ..., a_{J-1J}\}$ and $a_{JJ+1} > a_{J+1J+1} > \sum_{i=1}^{J-1}\sum_{j=1}^{J-1} a_{ij}$. Clearly, s_{J+1} is strictly dominated by s_J. (See Example 1.7)

For any $\hat{p} \in \{\hat{p} \in \Delta^{J+1} | \hat{p}_{J+1} \geq \max\{\hat{p}_1, ..., \hat{p}_{J-1}\}\}$ and any $k \in \{1, ..., J-1\}$

$$\hat{u}_J(\hat{p}) > \hat{u}_{J+1}(\hat{p}) > \hat{p}_J a_{J+1J} + \hat{p}_{J+1} \sum_{i=1}^{J-1}\sum_{j=1}^{J-1} a_{ij} \geq \hat{p}_J a_{kJ} + \sum_{i=1}^{J-1} \hat{p}_i a_{ki} = \hat{u}_k(\hat{p}) \qquad (1.5.2)$$

If σ is continuous, monotonic and cumulative, the QRE correspondence starting at the

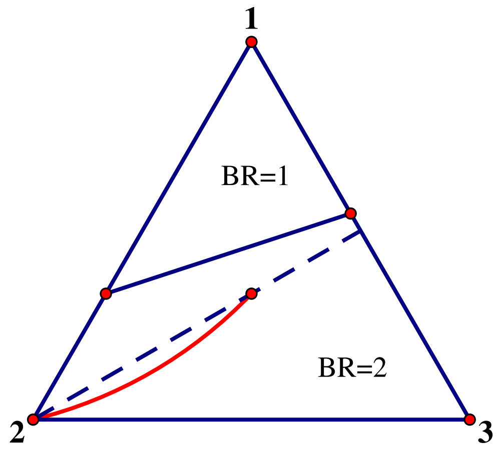


Figure 1.5.3: Best response regions and the QRE correspondence (red curve) for Example 1.7. The dash line denotes $\hat{p}_1 = \hat{p}_3$. The QRE correspondence starting at the centroid can not escape from $\{\hat{p} \in \Delta^3 | \hat{p}_3 \geq \hat{p}_1\}$.

centroid will enter and can not escape from $\{\hat{p} \in \Delta^{J+1} | \hat{p}_{J+1} \geq \max\{\hat{p}_1, ..., \hat{p}_{J-1}\}\}$. Therefore, the QRE correspondence must converge to the only Nash equilibrium, $\hat{p}_J = 1$, in the region. $\square$

Theorem 1.11 says that the limiting QRE of a game with strictly dominated strategies may be very different from that of the reduced form game. In fact, eliminating a strategy which is strictly dominated by the limiting QRE strategy may change the outcome of the equilibrium selection (see Example 1.7). The limiting QRE is highly sensitive to the addition and elimination of strictly dominated strategies.

Example 1.7

Consider the following 2×2 symmetric coordination game $\begin{pmatrix} 5 & 2 \\ 1 & 4 \end{pmatrix}$. The first strategy is the limiting QRE since it is $\frac{1}{2}$ dominant. We now add a strictly dominated strategy such that the second strategy is selected. From the proof of Theorem 1.11, an appropriate payoff matrix is $\begin{pmatrix} 5 & 2 & 0 \\ 1 & 4 & 7 \\ 0 & 2 & 6 \end{pmatrix}$, where the third strategy is strictly dominated by the second strategy. Best response regions are shown in Figure 1.5.3. In the new game, Gambit suggests that the limiting logit equilibrium is the second strategy. (See Figure 1.5.3)

1.6 Role games

A role game is a two-person symmetric game based on a bimatrix game (Selten, 1980; Gaunersdorfer et al., 1991; Weibull, 1995; Hofbauer and Sigmund, 1998; Berger, 2002). With the first move, nature randomly decides which player is player 1 (role 1) and which player is player 2 (role 2) in the later bimatrix game. After that, the two players play the bimatrix game, called the base game, according to the roles they have been assigned.

In this section, we consider a variant of the standard version, where players play both roles. [15] In detail, each player has to participate in two bimatrix games, where in one game acts as role 1 and in the other game acts as role 2. This modification does not change Nash equilibria of the role game. In fact, we could multiply the payoff matrix of the standard role game by the factor 2 to get the variant payoff values. We will see that each logit equilibrium of the base game corresponds to a logit equilibrium of the (variant) role game.

Let the base game be a $J_1 \times J_2$ bimatrix game. Denote the payoff to player i using strategy s_{ik} when he meets player j using strategy s_{jl} by a_{ikl}. Therefore, given a mixed strategy profile $p \in \Delta$, $u_{ik}(p_j) = \sum_{l=1}^{J_j} p_{jl} a_{ikl}$.

The base game leads to a $J_1 J_2 \times J_1 J_2$ symmetric role game. Denote a pure strategy in the role game by $\hat{s}_{ij}$, which means using strategy s_{1i} in role 1 and using strategy s_{2j} in role 2. The payoff to a player using strategy $\hat{s}_{ij}$ when he meets strategy $\hat{s}_{kl}$ is then given by $a_{1il} + a_{2jk}$. From the construction of the role game, each mixed strategy profile of the role game, $\hat{p} \in \Delta^{J_1 J_2}$, naturally corresponds to a mixed strategy profile of the bimatrix game, $p \in \Delta$, where $p_{1i} = \sum_{l=1}^{J_2} \hat{p}_{il}$ and $p_{2j} = \sum_{k=1}^{J_1} \hat{p}_{kj}$. For given $\hat{p} \in \Delta^{J_1 J_2}$, the average payoff of $\hat{s}_{ij}$ is then written as

$$\hat{u}_{ij}(\hat{p}) = \sum_{k=1}^{J_1} \sum_{l=1}^{J_2} \hat{p}_{kl}(a_{1il} + a_{2jk}) = u_{1i}(p_2) + u_{2j}(p_1) \tag{1.6.1}$$

Logit equilibria of the $J_1 \times J_2$ bimatrix game are the solutions of

$$\begin{aligned} p_{1i} &= \frac{e^{\lambda u_{1i}(p_2)}}{\sum_{k=1}^{J_1} e^{\lambda u_{1k}(p_2)}} \\ p_{2j} &= \frac{e^{\lambda u_{2j}(p_1)}}{\sum_{l=1}^{J_2} e^{\lambda u_{2l}(p_1)}} \end{aligned} \tag{1.6.2}$$

[15] Sigmund et al. adopted this variant in their paper "Reward and punishment" (2001), where the base game is the mini ultimatum game and the role game is the mini public goods game with punishment.

and in the corresponding $J_1J_2 \times J_1J_2$ role game, logit equilibria are the solutions of

$$\hat{p}_{ij} = \frac{e^{\lambda \hat{u}_{ij}(\hat{p})}}{\sum_{k=1}^{J_1} \sum_{l=1}^{J_2} e^{\lambda \hat{u}_{kl}(\hat{p})}} \tag{1.6.3}$$

Notice that $e^{\lambda \hat{u}_{ij}(\hat{p})} = e^{\lambda u_{1i}(p_2)} e^{\lambda u_{2j}(p_1)}$, there is a one-to-one mapping between the logit equilibria of the bimatrix game and the corresponding role game, where $\hat{p}_{ij} = p_{1i}p_{2j}$, $\sum_{l=1}^{J_2} \hat{p}_{il} = p_{1i}$ and $\sum_{k=1}^{J_1} \hat{p}_{kj} = p_{2j}$. Therefore, a Nash equilibrium p is the LLE in a bimatrix game if and only if the Nash equilibrium $(\hat{p}, \hat{p})$, where $\hat{p}_{ij} = p_{1i}p_{2j}$, is the LLE in the corresponding role game.

Furthermore, this result could be extended to any normal form game and its corresponding role game. If the base game is an n-person normal form game (N, S, u), the role game is an n-person $\prod_{i \in N} J_i$ -strategy symmetric game, where each player plays $n!$ different normal form games. [16] $p \in \pi_\lambda$ in the base game if and only if $(\hat{p}, ..., \hat{p}) \in \hat{\pi}_\lambda$ in the role game, where $\hat{p}_{i_1 \dots i_n} = \prod_{j=1}^{n} p_{ji_j}$, $i_j \in \{1, ..., J\}$. Therefore, the limiting logit equilibrium in any normal form game can be predicted by the corresponding symmetric role game.

However, this relation does not hold for more general quantal response functions. For instance, consider the following simple regular quantal response function

$$\sigma_{ij} = \frac{1 + \lambda u_{ij}}{J_i + \lambda \sum_{k=1}^{J_i} u_{ik}} \tag{1.6.4}$$

when payoffs are positive. It is easy to check that

$$\begin{aligned} \sum_{j=1}^{J_2} \hat{\sigma}_{ij} &= \sum_{j=1}^{J_2} \frac{1 + \lambda(u_{1i} + u_{2j})}{J_1J_2 + \lambda \sum_{k=1}^{J_1} \sum_{l=1}^{J_2} (u_{1k} + u_{2l})} \\ &= \frac{1 + \lambda(u_{1i} + c)}{J_1 + \lambda \sum_{k=1}^{J_1} (u_{1k} + c)} \neq \frac{1 + \lambda u_{1i}}{J_1 + \lambda \sum_{k=1}^{J_1} u_{1k}} = \sigma_{1i} \end{aligned} \tag{1.6.5}$$

if we take $u_{2j} = c > 0$ for all $j \in \{1, ..., J_2\}$.

1.7 Conclusion

McKelvey and Palfrey (1995) pointed out that the graph of logit equilibrium correspondence generically includes a unique branch connecting the centroid at $\lambda = 0$ to a unique Nash equilibrium as λ goes to infinity. They then suggested an equilibrium selection by tracing this branch. In this chapter, we extend this idea to quantal response functions

[16] Since the base game has n roles, there are $n!$ (the number of n-combinations) different normal form games.

that satisfy three axioms: continuity, monotonicity and cumulativity. From Brouwer's fixed point theorem, C^0 continuity is enough to guarantee the existence of a QRE for any given $\lambda = 0$. Monotonicity requires that players are "better responders" that play strategies with higher expected payoffs more often. As a result of monotonicity and continuity, the only QRE at $\lambda = 0$ is the centroid. Cumulativity ensures that players choose best responses as λ goes to infinity. Together with continuity, the QRE correspondence converges to Nash equilibria as λ goes to infinity. If a quantal response function satisfies the three axioms, the graph of QRE correspondence includes a path that connects the centroid at $\lambda = 0$ to at least one Nash equilibrium. However, C^0 continuity is too weak that the path is not necessarily nicely behaved. In exceptional cases, it may not be differentiable and bifurcation may arise. Such exceptional cases can be generically excluded by making differentiability assumption. If the quantal response function is C^2 continuous, except for a nowhere dense set of games, the path is diffeomorphic to a C^1 segment. This implies that for almost all normal form games, there is a unique selection from the set of Nash equilibria by "tracing" the graph of the QRE correspondence beginning at the centroid.

In the quantal response model, each player's payoffs are computed based on beliefs about other players' strategies, and in a QRE, the beliefs match the choice probabilities. In symmetric games with identical quantal response functions, the graph of QRE correspondence always contains a symmetric path (players have identical beliefs on this path) that connects the centroid to a symmetric Nash equilibrium. Therefore, the limiting QRE of a symmetric game must be a symmetric Nash equilibrium. If we further restrict the tracing process to the graph of symmetric QRE correspondence, the limiting QRE exists for almost all symmetric games. One implication directly from this result is that if a symmetric game has a unique symmetric Nash equilibrium, then it must be the limiting QRE.

In two-person symmetric games, a sufficient condition for the limiting QRE is GPPD. In coordination games, GPPD strategy is the unique best response when it is the most frequent strategy in opponent's strategy profile. This definition is closely related to the p-dominant equilibrium introduced by Morris et al. (1995), which says that each strategy of the strategy pair is a best response if the other player taking his strategy with probability at least p. In $J \times J$ coordination games, GPPD is stronger than $\frac{1}{2}$ dominant but weaker than $\frac{1}{J}$ dominant.

It is well known that the QRE is subject to framing effects: duplicating a strategy affects the equilibrium selection (Goeree and Holt, 2001, 2004). Such framing effects are inevitable if quantal response functions are differentiable (Hilbe, 2011). We offer a much stronger proposition: By appropriately adding a single strictly dominated strategy, any strict (symmetric) Nash equilibrium can be selected. Therefore, the limiting QRE is highly sensitive to the addition and elimination of strictly dominated strategies. [17] However, this does not imply that quantal response methods are without empirical content. Conversely, it is actually consistent with many experimental results. For instance, Cooper et al. (1990) provided evidence that strictly dominated strategies may influence equilibrium selection even though they are never selected as an outcome.

In this chapter, payoff functions are assumed to be linearly dependent on the noise factor λ. Previous results could be extended to more general cases if $\sigma_{ij}(\bar{u}_i, \lambda) = \frac{1}{J_i}$ for all $i \in N$ and $j \in \{1, ..., J_i\}$ when $\lambda = 0$. [18] For instance, consider the following regular quantal response function introduced by Luce (1959, see also Eq.(6.1) in Goeree et al., 2005) when payoffs are positive

$$\sigma_{ij}(\bar{u}_i, \lambda) = \frac{(u_{ij})^\lambda}{\sum_{k=1}^{J_i} (u_{ik})^\lambda} \tag{1.7.1}$$

It is easy to verify that continuity, monotonicity and cumulativity are satisfied, and the only QRE at $\lambda = 0$ is the centroid. Therefore, Eq.(1.7.1) yields a unique equilibrium selection for almost all games.

Finally, we show that there is a one-to-one mapping between the logit equilibria of a normal form game and the corresponding (variant) role game. Intuitively, QRE in the base game are the projection of QRE in the role game. Therefore, the limiting logit equilibrium in any normal form game can be predicted by the symmetric role game. In particular, if the base game is a 2×2 zero-sum game with a unique mixed Nash equilibrium p, the set of Nash equilibria in the role game is a continuum, where $\hat{p}_{11} + \hat{p}_{12} = p_{11}$ and $\hat{p}_{11} + \hat{p}_{21} = p_{21}$. Since the two games have the identical limiting QRE, logit equilibrium correspondence in the role game converges to the Wright equilibrium. [19] As pointed out by Berger (2002), this equilibrium is global attractive under the best response dynamics. Thus, the best response dynamics and the quantal response method select the same Nash equilibrium.

[17] Kim and Wong (2010) showed that for any symmetric normal form game, any strict (symmetric) Nash equilibrium can be selected as the unique long-run equilibrium (Kandori et al., 1993) by appropriately adding a single strategy which is strictly dominated by *all* original strategies.

[18] In this case, cumulativity is redefined as: $u_{ij} > u_{ik} \Rightarrow \lim\limits_{\lambda\to\infty} \frac{\sigma_{ik}(\bar{u}_i,\lambda)}{\sigma_{ij}(\bar{u}_i,\lambda)} = 0$.

[19] In the well known two-locus, two-alleles equation from populations genetics, the surface $\hat{p}_{11}\hat{p}_{22} = \hat{p}_{12}\hat{p}_{21}$ is called the Wright manifold. Berger (2002) then called the Nash equilibrium in this manifold the Wright equilibrium.

Chapter 2

Quantal response methods for equilibrium selection: 2×2 bimatrix games

Abstract

In this chapter, quantal response methods for equilibrium selection are analyzed in detail for 2×2 bimatrix coordination games. We show that in general not the risk-dominant equilibrium is selected. In the logarithmic game, the limiting QRE is the Nash equilibrium with the larger sum of square root payoff. Finally, we apply the quantal response methods to the public goods game with punishment. A cooperative equilibrium can be selected if punishment is strong enough.

Key words

Quantal response equilibrium; logit equilibrium; logrithmic game; equilibrium selection; punishment

2.1 Introduction

Quantal response equilibrium (QRE) introduced by McKelvey and Palfrey (1995) has been widely used to explain experimental data. In some applications, a limiting logit equilibrium, the limit of logit equilibrium as the noise approaches zero, is compared with limiting behavior in experiments (Anderson et al., 2001; Yi, 2002, 2005; Breitmoser et al., 2010). In contrast, few papers investigated the limiting QRE analytically. In particular, from our knowledge, the only theoretical work on equilibrium selection in normal form games is done by Turocy (2005). However, his claim that the QRE methods always select the risk dominant equilibrium, is wrong. [1]

In this chapter, we analyze the limiting QRE in 2×2 bimatrix games in detail. Section 2.2 defines logit equilibrium and reviews some basic results. Section 2.3 derives a formula for the quantal response equilibrium selection in 2×2 coordination games. Section 2.4 tests this formula with six different types of quantal response functions. For the logarithmic game (Harsanyi, 1973), we get a simple square root rule for the equilibrium selection. Finally, section 2.5 applies these results to the public goods game (PGG) with punishment. We compare the theoretical prediction with the past empirical studies qualitatively and hope to explain how punishment works in the real world.

2.2 Logit equilibrium

Let us start from a 2×2 bimatrix game, where A_1 and A_2 are two pure strategies of player A, and B_1 and B_2 are two pure strategies of player B. Suppose that a_{ij} denotes the payoff to player A using strategy A_i when it meets strategy B_j and b_{ij} the payoff to player B using strategy B_i when it meets strategy A_j. The bimatrix game is called a coordination game if pure strategy pairs (A_1, B_1) and (A_2, B_2) are both strict Nash equilibria. In this case, the payoff matrix can be normalized [2] as

$$\begin{array}{cc} & \begin{array}{cc} B_1 & B_2 \end{array} \\ \begin{array}{c} A_1 \\ A_2 \end{array} & \begin{pmatrix} 1-q, c(1-p) & 0,0 \\ 0,0 & q, cp \end{pmatrix} \end{array} \tag{2.2.1}$$

where $p = \frac{b_{22}-b_{12}}{b_{11}+b_{22}-b_{21}-b_{12}}$, $q = \frac{a_{22}-a_{12}}{a_{11}+a_{22}-a_{21}-a_{12}}$, $c = \frac{b_{11}+b_{22}-b_{21}-b_{12}}{a_{11}+a_{22}-a_{21}-a_{12}}$, and these parameters satisfy $0 < p, q < 1$, and $c > 0$. Besides of two strict pure Nash equilibria, the coordination game also has a mixed equilibrium (p, q).

[1] In 2×2 symmetric coordination games, the limiting QRE is the risk dominant equilibrium since the risk dominant equilibrium is $\frac{1}{2}$ dominant (See Theorems 1.10 and 1.11 in Chapter 1). But this statement is not true for 2×2 bimatrix coordination games.

[2] As pointed out by Goeree et al.(2005), structural quantal response functions involve only payoff differences.

In the framework of best response, players choose the strategy with the highest payoff. Denote the probability of player A using strategy A_1 by x and the probability of player B using strategy B_1 by y. One can easily show that A_1 is the best response strategy of player A if and only if $y > q$ and B_1 is the best response strategy of player B if and only if $x > p$.

Following McKelvey and Palfrey (1995), in a QRE, players are assumed boundedly rational and obverse noisy evaluations of the strategies values. In the 2×2 bimatrix game, player i will choose the first strategy if and only if

$$u_{i1} + \varepsilon_{i1} > u_{i2} + \varepsilon_{i2} \tag{2.2.2}$$

where u_{ij} denotes the payoff of player i using strategy j and ε_{ij} is the payoff disturbance. Best response function becomes probabilistic rather than deterministic. Suppose that player i's noise vector, $\varepsilon_i = (\varepsilon_{i1}, \varepsilon_{i2})$, is distributed according to a joint distribution with density function $p_i(\varepsilon_i)$, player i then adopts the first strategy with probability

$$\sigma_{i1}(u_{i1}, u_{i2}) = \int_{-\infty}^{\infty} \int_{-\infty}^{u_{i1}-u_{i2}+\varepsilon_{i1}} p_i(\varepsilon_i) d\varepsilon_{i2} d\varepsilon_{i1} \tag{2.2.3}$$

where $\sigma_{ij}(u_{i1}, u_{i2})$ is called the quantal response function.

The most common specification of QRE is the logit equilibrium. For any given $\lambda \geq 0$, the logistic quantal response function is given by

$$\sigma_{i1}(u_{i1}, u_{i2}) = \frac{e^{\lambda u_{i1}}}{e^{\lambda u_{i1}} + e^{\lambda u_{i2}}} = \frac{1}{1 + e^{\lambda(u_{i2} - u_{i1})}}. \tag{2.2.4}$$

This arises from (2.2.3) if all the noises follow the extreme value distribution with cumulative distribution function $\exp(-\exp(-\lambda\varepsilon - \gamma))$, where γ is Euler's constant. Therefore, if each player uses a logistic quantal response function, the corresponding logit equilibria are the solutions of

$$\begin{aligned} x &= \frac{1}{1 + e^{\lambda(q-y)}} \\ y &= \frac{1}{1 + e^{\lambda c(p-x)}} \end{aligned} \tag{2.2.5}$$

Consider the logit equilibria as a function of λ. $\lambda = 0$ means full noise and $\lambda = +\infty$ means no noise. For the set of logit equilibria, it is obvious that when $\lambda = 0$, Eq.(2.2.5) has a unique solution $(\frac{1}{2}, \frac{1}{2})$. On the other hand, when $\lambda \to +\infty$, the set of logit equilibria approaches the set of all three Nash equilibria of the game.

As shown by McKelvey and Palfrey (1995), for almost all normal form games, the graph of the logit equilibria correspondence contains a unique branch which starts for $\lambda = 0$ at the centroid of the strategy simplex and converges to a unique Nash equilibrium as λ goes to infinity. This defines a unique selection from the set of Nash equilibria by "tracing" the graph of the logit equilibria correspondence. The selected Nash equilibrium is called the *limiting logit equilibrium* of the game.

2.3 Equilibrium selection

In this section, we study the equilibrium selection by the structural QRE (Goeree et al., 2005), where the quantal response function is defined in Eq.(2.2.3). Notice that $\sigma_{ij}(u_{i1}, u_{i2})$ depends only on the payoff difference $u_i = u_{i1} - u_{i2}$, we write $\sigma_i(u_i)$ the probability of player i choosing the first strategy with payoff difference u_i. The density function p_i is called admissible if it is continuous, unbiased and independent across players (McKelvey and Palfrey, 1995). From Proposition 2 in Goeree et al., 2005, if p_i is admissible, $\sigma_i(u_i)$ satisfies:

(a) continuous,

(b) monotonically increasing in u_i,

(c) $\sigma_i(u_i) = 1 - \sigma_i(-u_i)$.

That is, $\sigma_i(u_i)$ is a cumulative density function of a symmetric distribution. For convenience, suppose that the disturbance on each player is identically distributed and follows a unimodal distribution. In this case, players have the same quantal response function, $\sigma = \sigma_1 = \sigma_2$, and $\sigma(u)$ is the cumulative distribution function of a unimodal symmetric distribution, where $\sigma'(0) > 0$, $\sigma''(u) \geq 0$ for $u < 0$ and $\sigma''(u) \leq 0$ for $u > 0$. The explicit formula of σ of course depends on the density function and we will investigate different types of noises in the next section.

Following the logit equilibrium, let us introduce the level of noise λ ($\lambda \geq 0$) into the structural QRE and write the quantal response function at level λ by $\sigma(\lambda u)$. $\lambda = 0$ means full noise and $\lambda = +\infty$ means no noise. For the 2×2 bimatrix game (2.2.1), QRE at level λ are the solutions of

$$\begin{aligned} x &= \sigma(\lambda(y - q)) \\ y &= \sigma(\lambda c(x - p)) \end{aligned} \tag{2.3.1}$$

As pointed out by Goeree et al. (2005), there exists a structure QRE of the game (2.2.1) for any admissible p_i. Let us now regard the solution of Eq.(2.3.1) as a 3-dimensional vector (x, y, λ), where (x, y) is the QRE at noise level λ. When $\lambda = 0$, Eq.(2.3.1) has a unique solution $(\frac{1}{2}, \frac{1}{2}, 0)$. When $\lambda = +\infty$, Eq.(2.3.1) has three solutions, $(0, 0, +\infty)$, $(1, 1, +\infty)$ and $(p, q, +\infty)$, which correspond to the three Nash equilibria of the coordination game. Similarly as the logit equilibrium, Eq.(2.3.1) induces a continuous path (x, y, λ) starting from the center point and to one of the Nash equilibria for almost all games (Theorem 1.6 in Chapter 1). This then defines a unique equilibrium selection. For given σ, we call the selected Nash equilibrium the *limiting QRE* of the game (2.2.1).

Theorem 2.1

(a) For $c = 1$, the limiting QRE is $(1, 1)$ if $p + q < 1$, and is $(0, 0)$ if $p + q > 1$. (b) For $c \neq 0$, the limiting QRE is $(1, 1)$ if $p + q < 1$ and $cp + q < \frac{1}{2} + \frac{c}{2}$, and is $(0, 0)$ if

$p+q>1$ and $cp+q>\frac{1}{2}+\frac{c}{2}$. [3]

Proof:

The intersecting points of the QRE correspondence and the plane $x+y=1$ satisfy

$$\sigma(\lambda(y-q))+\sigma(\lambda c(x-p))=1 \tag{2.3.2}$$

Since σ is the cumulative distribution function of a symmetric distribution, $\lambda(y-q+c(x-p))=0$.

(a) For $c=1$, $(\frac{1}{2},\frac{1}{2},0)$ is the only intersection if $p+q\neq 1$. This implies that the limiting QRE is $(1,1)$ if $x+y>1$ for any $\lambda>0$ and is $(0,0)$ if $x+y<1$ for any $\lambda>0$.

Let $\lambda\to 0^+$, Eq.(2.3.1) could be approximated as

$$\begin{aligned} x &\approx \frac{1}{2}+\lambda(y-q)\sigma'(0) \\ y &\approx \frac{1}{2}+\lambda(x-p)\sigma'(0) \end{aligned} \tag{2.3.3}$$

and

$$x+y\approx 1+\lambda(1-p-q)\sigma'(0) \tag{2.3.4}$$

Hence, the limiting point is $(1,1)$ if $p+q<1$ and is $(0,0)$ if $p+q>1$.

(b) For $c\neq 1$, suppose that $c<1$. Clearly, $(\frac{1}{2},\frac{1}{2},0)$ is an intersection, and another intersection is $(\frac{1-cp-q}{1-c},\frac{cp+q-c}{1-c},\lambda(p,q))$, where $\lambda(p,q)$ is the solution of

$$\frac{1-cp-q}{1-c}=\sigma(\lambda(p,q)c\frac{p+q-1}{1-c}) \tag{2.3.5}$$

If $(\frac{1-cp-q}{1-c},\frac{cp+q-c}{1-c})$ is not in region $[0,1]\times[0,1]\times[0,\infty)$, i.e., $cp+q<c$ or $1<cp+q$ or $(p+q-1)(cp+q-\frac{1}{2}-\frac{c}{2})<0$, $(\frac{1}{2},\frac{1}{2},0)$ is the only intersecting point. In this case, the limiting QRE is $(1,1)$ if $x+y>1$ for any $\lambda>0$ and is $(0,0)$ if $x+y<1$ for any $\lambda>0$.

Similarly as (a), for $\lambda\to 0^+$, Eq.(2.3.1) could be approximated as

$$\begin{aligned} x &\approx \frac{1}{2}+\lambda(y-q)\sigma'(0) \\ y &\approx \frac{1}{2}+\lambda c(x-p)\sigma'(0) \end{aligned} \tag{2.3.6}$$

and

$$x+y\approx 1+\lambda(\frac{1}{2}+\frac{c}{2}-q-cp)\sigma'(0). \tag{2.3.7}$$

[3] The bimatrix game Eq.(2.2.1) is symmetric if and only if $c=1$ and $p+q=1$. Thus, Theorem 2.1 implies that the limiting QRE does not exist for all 2×2 symmetric games with a unique symmetric mixed Nash equilibrium.

This implies that $x + y < 1$ if and only if $cp + q > \frac{1}{2} + \frac{c}{2}$. Hence, $(0, 0)$ is selected if $p + q > 1$ and $cp + q > \frac{1}{2} + \frac{c}{2}$, and $(1, 1)$ is selected if $p + q < 1$ and $cp + q < \frac{1}{2} + \frac{c}{2}$. □

For 2×2 bimatrix games, $(1, 1)$ is risk dominant if and only if the Nash products satisfy the inequality $c(1-q)(1-p) > cpq$, i.e., $p+q < 1$ (Harsanyi and Selten, 1988). Theorem 2.1 suggests that the quantal response equilibrium selection conforms to the Nash product rule if and only if $c = 1$, although the QRE method is very close to the "tracing procedure" of Harsanyi and Selten (1988; Harsanyi, 1975). On the other hand, $(1, 1)$ is the limiting QRE if $p < \frac{1}{2}$ and $q < \frac{1}{2}$. This implies that if a strategy is "risk dominant for both players", it will be selected by the QRE methods independent of c. Turocy (2005) made a mistake in his proof of Theorem 7 that applied this stronger condition instead of the Nash product rule.

As an extension of Theorem 2.1(b), we look at two limit cases $c = 0$ and $c = \infty$. If $c = 0$, Eq.(2.3.5) has a solution if and only if $q = \frac{1}{2}$. Hence, $(\frac{1}{2}, \frac{1}{2}, 0)$ is the unique intersection if $q \neq \frac{1}{2}$. From Eq.(2.3.7), $(1, 1)$ $((0, 0))$ is the limiting QRE if and only if $q < \frac{1}{2}$ $(q > \frac{1}{2})$, which is independent of p. This implies that the influence of player B on the equilibrium selection is negligible if its payoff is much less than player A. Similarly, if $c \to \infty$, $(1, 1)$ $((0, 0))$ is the limiting QRE if and only if $p < \frac{1}{2}$ $(p > \frac{1}{2})$, which is independent of q. This shows clearly that our equilibrium selection method depends crucially on c, in contrast to Nash products.

One may notice that limit set of QRE as $\lambda \to +\infty$ has three Nash equilibria, but Theorem 2.1 only mentioned the two pure Nash equilibria. In fact, Theorem 2.2 will show that there is no path connecting the mixed equilibrium (p, q) and the centroid $(\frac{1}{2}, \frac{1}{2})$ for almost all coordination games.

In order to derive a precise result of equilibrium selection, we introduce a new concept. For given c, define the *separatrix* as the curve in the $p-q$ plane separating the two regions where the limiting QRE are $(0, 0)$ and $(1, 1)$, respectively. Theorem 2.1 says that the separatrix lies between lines $p + q = 1$ and $cp + q = \frac{1}{2} + \frac{c}{2}$. Furthermore, $(1, 1)$ is selected for (p, q) below the separatrix and $(0, 0)$ is selected for (p, q) above the separatrix. (See Figure 2.3.1) In the rest of this section, we will derive an expression of the separatrix.

Define $f = \sigma^{-1}$. Clearly, f is an increasing function in $[0, 1]$, where $f(0) = -\infty$, $f(\frac{1}{2}) = 0$, $f(1) = +\infty$, $f''(x) \geq 0$ if $x < \frac{1}{2}$ and $f''(x) \leq 0$ if $x > \frac{1}{2}$. The projection of Eq.(2.3.1) on $x - y$ plane is then written as

$$(y - q)f(y) = c(x - p)f(x) \tag{2.3.8}$$

since

$$\begin{aligned} f(x) &= \lambda(y - q) \\ f(y) &= \lambda c(x - p) \end{aligned} \tag{2.3.9}$$

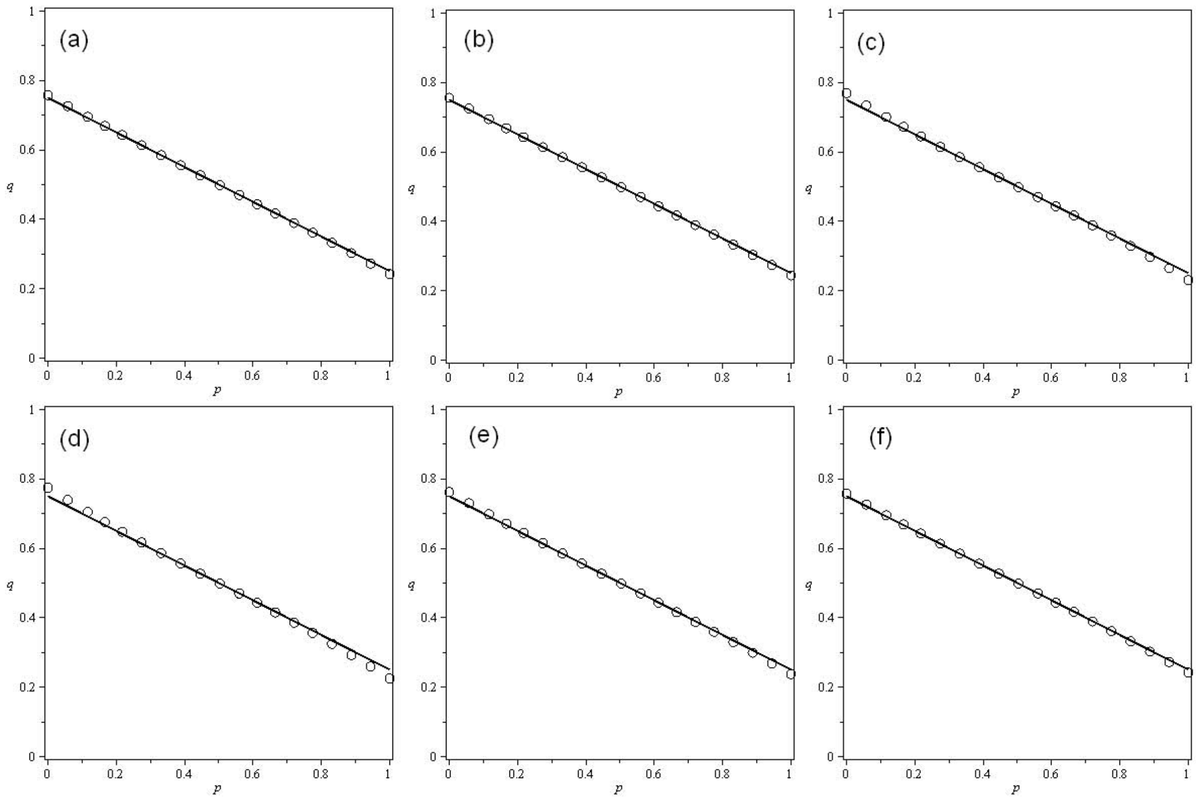


Figure 2.3.1: Separatrix on the $p-q$ plane, and its tangent line at $p = q = \frac{1}{2}$, for $c = 0.25$. Figures (a)-(f) are respectively the logit equilibrium, the probit equilibrium, the Cauchy noise, the exponential noise, the uniform noise and the logarithmic game. Circles denote numerically computed points on the separatrices. The slope of the tangent line is $-\frac{1}{2}$ (independently of σ). The Nash equilibrium $(1, 1)$ is selected for (p, q) in the region below the separatrix and $(0, 0)$ is selected for (p, q) in the region above the separatrix. The linear approximation is works well for these six types of quantal response functions.

From Eq.(2.3.9), each solution of Eq.(2.3.8) corresponds to a unique QRE. Therefore, we turn to investigate the graph of Eq.(2.3.8) instead of Eq.(2.3.1).

Theorem 2.2

For almost all parameters, the graph of Eq.(2.3.8) consists of two (disjoint) branches, where one passes through the mixed equilibrium (p, q) *and the other passes through the centroid* $(\frac{1}{2}, \frac{1}{2})$. *For the critical case, two branches intersect at a singular point.*

Proof:

One can easily calculate that Eq.(2.3.8) has four interior solutions, $(\frac{1}{2}, \frac{1}{2})$, $(p, \frac{1}{2})$, $(\frac{1}{2}, q)$ and (p, q), and four boundary solutions $(0, 0)$, $(1, 0)$, $(0, 1)$ and $(1, 1)$. For convenience, we divide the $x - y$ plane to nine regions by four lines $x = \frac{1}{2}$, $x = p$, $y = \frac{1}{2}$ and $y = q$, and study the graph of Eq.(2.3.8) in each region. Clearly, no solution in regions (2), (4), (6) and (8). Furthermore, $\lambda > 0$ in region (1) and (9), and $\lambda < 0$ in region (3) and (7). (See Figure 2.3.2 (a))

Without loss of generality, assume $p > \frac{1}{2}$ and $q < \frac{1}{2}$. Define

$$S(x, y) = (y - q)f(y) - c(x - p)f(x) \tag{2.3.10}$$

and the derivatives of $S(x, y)$ satisfy

$$\begin{aligned} S_x &= -c(x - p)f'(x) - cf(x) \\ S_{xx} &= -c(2f'(x) + (x - p)f''(x)) \\ S_y &= (y - q)f''(y) + f(y) \\ S_{yy} &= 2f'(y) + (y - q)f''(y) \end{aligned} \tag{2.3.11}$$

From the properties of f, we have $S_x < 0$ if $x > p$, $S_x > 0$ if $x < \frac{1}{2}$, $S_y > 0$ if $y > \frac{1}{2}$, $S_y < 0$ if $y < q$, $S_{xx} \leq 0$ if $\frac{1}{2} < x < p$ and $S_{yy} \geq 0$ if $q < y < \frac{1}{2}$.

Hence, in region (1), $S(x, y) = 0$ is a increasing curve from $(0, 0)$ to $(\frac{1}{2}, q)$ since $S_x > 0$ and $S_y < 0$; in region (3), it is a decreasing curve from (p, q) to $(1, 0)$ since $S_x < 0$ and $S_y < 0$; in region (5), it is a decreasing curve from $(0, 1)$ to $(\frac{1}{2}, \frac{1}{2})$ since $S_x > 0$ and $S_y > 0$; and in region (9), it is a increasing curve from $(p, \frac{1}{2})$ to $(1, 1)$ since $S_x < 0$ and $S_y > 0$. (See Figure 2.3.2 (a))

On the other hand, in region (5), we have $S_{xx} \leq 0$ and $S_{yy} \geq 0$, i.e., S is a convex function of y and a concave function of x. This implies that for given $\hat{y}$, $S(x, y) = 0$ has (a) two solutions $(\hat{x_1}, \hat{y})$ and $(\hat{x_2}, \hat{y})$ if $S(x^*, \hat{y}) > 0$, (b) one solution $(x^*, \hat{y})$ if and $S(x^*, \hat{y}) = 0$ (c) no solution if $S(x^*, \hat{y}) < 0$, where $S_x(x^*) = 0$ and $\hat{x_1} < x^* < \hat{x_2}$. Similarly, for given $\hat{x}$, $S = 0$ has (d) two solutions $(\hat{x}, \hat{y_1})$ and $(\hat{x}, \hat{y_2})$ if $S(\hat{x}, y^*) < 0$, (e) one solution $(\hat{x}, y^*)$ if $S(\hat{x}, y^*) = 0$ and (f) no solution if $S(\hat{x}, y^*) > 0$, where $S_y(y^*) = 0$ and $\hat{y_1} < y^* < \hat{y_2}$. Thus, the graph of $S(x, y) = 0$ in region (5) consists of two disjoint curves that separated by line

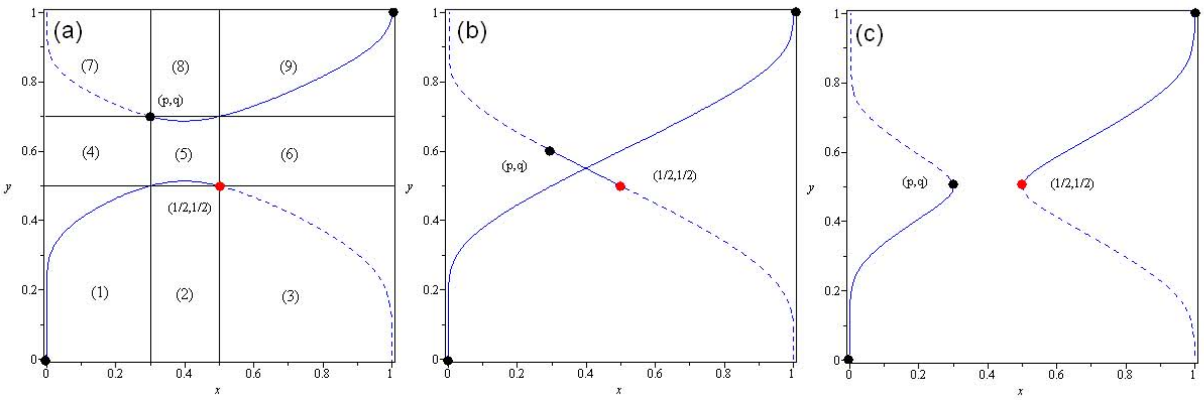


Figure 2.3.2: Logit equilibrium correspondence for the coordination game. $\lambda \geq 0$ on solid curves but $\lambda < 0$ on dashed curves. Black points are NE and red points are $(\frac{1}{2}, \frac{1}{2})$. Parameters are taken as $p = 0.3$, $c = 0.25$, $q = 0.7$ in (a), $q = 0.6005$ in (b), and $q = 0.5$ in (c). For almost all games, the graph of Eq.(2.3.8) consists of two branches, where one passes through the mixed equilibrium (p, q) and the other passes through the centroid $(\frac{1}{2}, \frac{1}{2})$. In the critical case (b), two branches intersect at a singular point.

$x = x^*$ if $S(x^*, y^*) > 0$ (from (a) and (f)) and separated by line $y = y^*$ if $S(x^*, y^*) < 0$ (from (c) and (d)).

In sum, the graph of $S(x, y) = 0$ consists of two branches, where the Nash equilibrium (p, q) and the centroid $(\frac{1}{2}, \frac{1}{2})$ are always on different branches. Furthermore, $(\frac{1}{2}, \frac{1}{2})$ is on the curve passing through $(0, 0)$ if and only if $S(x^*, y^*) > 0$ and $x^* > \frac{1}{2}$ or $S(x^*, y^*) < 0$ and $y^* > \frac{1}{2}$. For the critical case $S(x^*, y^*) = 0$, two branches intersect at a singular point (x^*, y^*). (See Figure 2.3.2 (b)) □

Theorem 2.2 implies that the limiting QRE is $(0, 0)$ (or $(1, 1)$) if and only if $(0, 0)$ (or $(1, 1)$) and $(\frac{1}{2}, \frac{1}{2})$ are on the same branch. On the other hand, (p, q) can not be selected for almost all games since there is no path from $(\frac{1}{2}, \frac{1}{2})$ to it. However, in the critical case, two branches intersect at a singular point (x^*, y^*) (we simply note it by (x, y) in later discussions) and tracing the branch of QRE correspondence beginning at the centroid could reach all three Nash equilibria. Thus, (p, q) is on the separatrix if and only if Eq.(2.3.8) has a singular point.

From the proof of Theorem 2.2, the singular point satisfies

$$\begin{aligned} S(x, y) &= (y - p)f(y) - c(x - p)f(x) = 0 \\ S_x(x, y) &= f(x) + (x - p)f'(x) = 0 \\ S_y(x, y) &= f(y) + (y - q)f'(y) = 0 \end{aligned} \tag{2.3.12}$$

and this yields

$$p = x + \frac{f(x)}{f'(x)},$$
$$q = y + \frac{f(y)}{f'(y)}. \tag{2.3.13}$$

Define $F(x) = x + \frac{f(x)}{f'(x)}$. It is easy to see that $F(\frac{1}{2}) = \frac{1}{2}$ since $f(\frac{1}{2}) = 0$. Notice that

$$F'(x) = 2 - \frac{f(x)f''(x)}{f'(x)^2} > 0 \tag{2.3.14}$$

, $F(x)$ is an increasing function. Hence, Eq.(2.3.8) could be written as

$$(F^{-1}(q) - q)f(F^{-1}(q)) = c(F^{-1}(p) - p)f(F^{-1}(p)). \tag{2.3.15}$$

For convenience, define $H(p) = (F^{-1}(p) - p)^{1/2} f(F^{-1}(p))^{1/2}$. The expression of the separatrix is then simplified as

$$H(q) = -c^{1/2} H(p) \tag{2.3.16}$$

where the minus is decided by Theorem 2.1.

Since $H(\frac{1}{2}) = 0$, the separatrix can be reduced to

$$q = \frac{1}{2} - c^{1/2}(p - \frac{1}{2}) \tag{2.3.17}$$

after ignoring the high order terms of $H(p)$. Eq.(2.3.17) shows that the first order term is independent of the quantal response function.

2.4 Different types of quantal response functions

Eq.(2.3.17) provides a linear approximation of the separatrix but one may doubt that whether this oversimplification is appropriate. In this section, we are going to derive the high order terms of the separatrix for six different types of quantal response functions. In general, the separatrix does not have an explicit formula (except in section 2.4.6). We provide both the power series and numerical simulations and show that the limiting QRE is affected little by the noise structure but mainly decided by the payoff matrix. This implies that Eq.(2.3.17) is a good approximation for the equilibrium selection by QRE methods. Based on the logarithmic game (Harsanyi, 1973), we find a simple rule to decide the limiting QRE: the strategy with larger sum of square root payoff is selected.

2.4.1 Logit equilibrium

The first example is the logit equilibrium discussed in section 2.2. If the noises follow the extreme value distribution with cumulative distribution function $\exp(-\exp(-\lambda\varepsilon-\gamma))$, the quantal response function is

$$\sigma(u)=\frac{1}{1+e^{-u}}. \tag{2.4.1}$$

By using the Taylor expansion, high order terms of $H(p)$ are

$$H(p)=\Delta p+\frac{1}{6}\Delta p^3+\frac{17}{120}\Delta p^5+o(\Delta p^7) \tag{2.4.2}$$

and the separatrix could be written as

$$\begin{aligned}\Delta q= \quad & -c^{1/2}\Delta p+(-\frac{1}{6}c^{1/2}+\frac{1}{6}c^{3/2})\Delta p^3\\ & +(\frac{1}{12}c^{3/2}-\frac{17}{120}c^{1/2}+\frac{7}{120}c^{5/2})\Delta p^5+o(\Delta p^7)\end{aligned} \tag{2.4.3}$$

where $\Delta p=p-\frac{1}{2}$ and $\Delta q=q-\frac{1}{2}$. (See Figure 2.3.1 (a))

2.4.2 Probit equilibrium

If the random perturbations follow the normal distribution (Palfrey and Prisbrey, 1997; Staudigl, 2011), $\sigma(u)$ is the cumulative distribution function of the normal distribution, i.e.,

$$\sigma(u)=\Phi(u)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{u}e^{-t^2/2}dt. \tag{2.4.4}$$

By using the Taylor expansion, high order terms of $H(p)$ are

$$H(p)=\Delta p+\frac{\pi}{24}\Delta p^3+\frac{19\pi^2}{1920}\Delta p^5+o(\Delta p^7) \tag{2.4.5}$$

and the separatrix could be written as

$$\begin{aligned}\Delta q= \quad & -c^{1/2}\Delta p+(-\frac{1}{24}c^{1/2}+\frac{1}{24}c^{3/2})\pi\Delta p^3\\ & +(\frac{1}{640}c^{3/2}-\frac{19}{1920}c^{1/2}+\frac{1}{640}c^{5/2})\pi^2\Delta p^5+o(\Delta p^7)\end{aligned} \tag{2.4.6}$$

(See Figure 2.3.1 (b))

2.4.3 Cauchy noise

Another famous noise structure is the Cauchy noise, which is the quotient distribution of two normal variables. An interpretation is that the noise comes from evaluating the frequencies of strategies. In this case, $\sigma(u)$ is the cumulative distribution function of the Cauchy distribution, i.e.,

$$\sigma(u) = \frac{1}{\pi}\arctan(u) + \frac{1}{2}. \tag{2.4.7}$$

By using the Taylor expansion, high order terms of $H(p)$ are

$$H(p) = \Delta p + \frac{\pi^2}{24}\Delta p^3 + \frac{13\pi^4}{1920}\Delta p^5 + o(\Delta p^7) \tag{2.4.8}$$

and the separatrix could be written as

$$\begin{aligned}\Delta q = & -c^{1/2}\Delta p + (-\frac{1}{24}c^{1/2} + \frac{1}{24}c^{3/2})\pi^2\Delta p^3 \\ & +(\frac{1}{192}c^{3/2} - \frac{13}{1920}c^{1/2} + \frac{1}{640}c^{5/2})\pi^4\Delta p^5 + o(\Delta p^7)\end{aligned} \tag{2.4.9}$$

(See Figure 2.3.1 (c))

2.4.4 Exponential noise

If the random perturbations follow the exponential distribution, $\sigma(u)$ is the cumulative distribution function of the Laplace distribution, i.e.,

$$\sigma(u) = \begin{cases} \frac{e^u}{2} & u < 0 \\ 1 - \frac{e^{-u}}{2} & 0 \le u \end{cases} \tag{2.4.10}$$

In this case, σ is not smooth at 0. Nevertheless, method in section 2.3 is well defined since $\sigma''(u) > 0$ for $u < 0$ and $\sigma''(u) < 0$ for $u > 0$.

For $p < 1/2$, high order terms of $H(p)$ are

$$H(p) = \Delta p + \frac{1}{4}\Delta p^2 + \frac{1}{6}\Delta p^3 + \frac{29}{192}\Delta p^4 + o(\Delta p^5) \tag{2.4.11}$$

and the separatrix could be written as

$$\begin{aligned}\Delta q = & -c^{1/2}\Delta p + (\frac{1}{4}c^{1/2} - \frac{1}{4}c)\Delta p^2 + (\frac{1}{8}c - \frac{1}{6}c^{1/2} - \frac{1}{6}c^{3/2})\Delta p^3 \\ & +(-\frac{19}{192}c + \frac{29}{192}c^{1/2} + \frac{1}{8}c^{3/2} - \frac{29}{192}c^2)\Delta p^4 + o(\Delta p^5)\end{aligned} \tag{2.4.12}$$

From the symmetry of $\sigma(u)$, for $p > 1/2$, high order terms of $H(p)$ are

$$H(p) = \Delta p - \frac{1}{4}\Delta p^2 + \frac{1}{6}\Delta p^3 - \frac{29}{192}\Delta p^4 + o(\Delta p^5) \tag{2.4.13}$$

and the separatrix could be written as

$$\begin{aligned}\Delta q = & -c^{1/2}\Delta p - (\frac{1}{4}c^{1/2} - \frac{1}{4}c)\Delta p^2 + (\frac{1}{8}c - \frac{1}{6}c^{1/2} - \frac{1}{6}c^{3/2})\Delta p^3 \\ & -(-\frac{19}{192}c + \frac{29}{192}c^{1/2} + \frac{1}{8}c^{3/2} - \frac{29}{192}c^2)\Delta p^4 + o(\Delta p^5)\end{aligned} \quad (2.4.14)$$

(See Figure 2.3.1 (d))

2.4.5 Uniform noise

If the perturbations follow the uniform distribution (Gale et al., 1995), $\sigma(u)$ is the cumulative distribution function of the triangular distribution

$$\sigma(u) = \begin{cases} 0 & u < -1 \\ \frac{(1+u)^2}{2} & -1 \le u < 0 \\ 1 - \frac{(1-u)^2}{2} & 0 \le u < 1 \\ 1 & 1 \le u \end{cases} \quad (2.4.15)$$

Similarly as the exponential distribution, σ is not smooth at 0 but $\sigma''(u) = 1 > 0$ for $u < 0$ and $\sigma''(u) = -1 < 0$ for $u > 0$.

For $p < 1/2$, high order terms of $H(p)$ are

$$H(p) = \Delta p + \frac{1}{8}\Delta p^2 + \frac{1}{16}\Delta p^3 + \frac{23}{512}\Delta p^4 + o(\Delta p^5) \quad (2.4.16)$$

and the separatrix could be written as

$$\begin{aligned}\Delta q = & -c^{1/2}\Delta p - (\frac{1}{8}c^{1/2} - \frac{1}{8}c)\Delta p^2 + (-\frac{1}{32}c - \frac{1}{16}c^{1/2} + \frac{1}{32}c^{3/2})\Delta p^3 \\ & +(-\frac{9}{512}c - \frac{23}{512}c^{1/2} + \frac{3}{256}c^{3/2} - \frac{1}{64}c^2)\Delta p^4 + o(\Delta p^5)\end{aligned} \quad (2.4.17)$$

From the symmetry of $\sigma(u)$, for $p > 1/2$, high order terms of $H(p)$ are

$$H(p) = \Delta p - \frac{1}{8}\Delta p^2 + \frac{1}{16}\Delta p^3 - \frac{23}{512}\Delta p^4 + o(\Delta p^5) \quad (2.4.18)$$

and the separatrix could be written as

$$\begin{aligned}\Delta q = & -c^{1/2}\Delta p + (\frac{1}{8}c^{1/2} - \frac{1}{8}c)\Delta p^2 + (-\frac{1}{32}c - \frac{1}{16}c^{1/2} + \frac{1}{32}c^{3/2})\Delta p^3 \\ & +(\frac{9}{512}c + \frac{23}{512}c^{1/2} - \frac{3}{256}c^{3/2} + \frac{1}{64}c^2)\Delta p^4 + o(\Delta p^5)\end{aligned} \quad (2.4.19)$$

(See Figure 2.3.1 (e))

2.4.6 Logarithmic game

The last example is the logarithmic game introduced by Harsanyi (1973). In this game, a player's utility has the form $\frac{\lambda}{\lambda+1}U + \frac{1}{\lambda+1}L$ with $\lambda \geq 0$, where U depends on payoff matrix and L depends on the player's own strategy. For 2×2 bimatrix games, payoff functions are defined as

$$\hat{u}_1(x) = \frac{\lambda}{\lambda+1}(xy(1-q) + (1-x)(1-y)q) + \frac{1}{\lambda+1}\log(x(1-x))$$
$$\hat{u}_2(y) = \frac{\lambda}{\lambda+1}(yxc(1-p) + (1-y)(1-x)cp) + \frac{1}{\lambda+1}\log(y(1-y)) \quad (2.4.20)$$

where $\hat{u}_i(x)$ is the payoff for player i using strategy x.

Suppose that individuals choose strategies maximizing their payoffs. The corresponding QRE is then the solution of

$$\frac{\partial \hat{u}_i(x)}{\partial x} = \frac{\lambda}{\lambda+1}(y-q) + \frac{1}{\lambda+1}\frac{1-2x}{x(1-x)} = 0$$
$$\frac{\partial \hat{u}_2(y)}{\partial y} = \frac{\lambda}{\lambda+1}(x-p) + \frac{1}{\lambda+1}\frac{1-2y}{y(1-y)} = 0 \quad (2.4.21)$$

where the quantal response function is

$$\sigma(u) = \frac{1}{2} + \frac{u}{4(1+\sqrt{u^2/4+1})} \quad (2.4.22)$$

From Eq.(2.3.12) to Eq.(2.3.16), we obtain

$$H(p) = \sqrt{\frac{1}{2} - \sqrt{\frac{1}{4} - (p-\frac{1}{2})^2}} \quad (2.4.23)$$

and the separatrix has an explicit formula

$$q = \frac{1}{2} - c^{1/2}\sqrt{\frac{1}{2} - \sqrt{\frac{1}{4} - (p-\frac{1}{2})^2}}\sqrt{1 - c(\frac{1}{2} - \sqrt{\frac{1}{4} - (p-\frac{1}{2})^2})} \quad (2.4.24)$$

Interestingly, Eq.(2.4.24) is equivalent to the following simpler expression

$$\sqrt{1-q} + \sqrt{c(1-p)} = \sqrt{q} + \sqrt{cp} \quad (2.4.25)$$

which means that the strategy with larger sum of square root payoff will be selected.

Since the equilibrium selection is affected little by the quantal response function, Eq.(2.4.25) then provides a simple way to estimate the limiting QRE directly from the non-normalized payoff matrix. Nash equilibrium (A_1, B_1) is selected by the QRE methods if it has larger sum of square root payoff, i.e.,

$$\sqrt{a_{11} - a_{21}} + \sqrt{b_{11} - b_{21}} > \sqrt{a_{22} - a_{12}} + \sqrt{b_{22} - b_{12}} \quad (2.4.26)$$

(See Figure 2.3.1 (f))

2.5 Application

In this section, we study the public goods game (PGG) with punishment (Sigmund et al., 2000, 2001). The only strict Nash equilibrium in this game is that do not contribute to the public pool and do not punish free riders. However, empirical researches indicated that punishment can curb free-riding. By applying the results in section 2.3, we find that a cooperative equilibrium is selected by the QRE method if the punishment is strong enough. For intermediate punishment, cooperation could also dominate the population when λ is not so large even if the limiting logit equilibrium is defection. By comparing the QRE model to past experiments, we hope to explain how punishment works in the real world.

2.5.1 Public goods game with punishment

Following Sigmund et al. (2001), we consider a two players PGG, where both can send a benefit b to their coplayer at a cost of a. In the second stage, they are offered the opportunity to punish their coplayer by imposing a fine. The fine amounts to a loss β to the punished player, and it entails a cost α to the punisher. Let us label with C (cooperator) those players who cooperate by sending a benefit and with D (defector) those who do not. Let P denotes those who punish defectors and N those who do not. The payoff matrix is then written as

$$\begin{array}{cc} & \begin{array}{cc} P & N \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} -a, b & -a, b \\ -\beta, -\alpha & 0, 0 \end{pmatrix} \end{array} \tag{2.5.1}$$

This mini PGG with punishment is obviously equivalent to the mini ultimatum game or the Prisoner's Dilemma game with punishment (Gale et al., 1995; Nowak et al., 2000; Dreber et al., 2008). The game has infinite number of Nash equilibria, one pure Nash equilibrium $(0, 0)$ and non-isolated Nash equilibria $(1, \hat{y})$, where $\frac{a}{\beta} \le \hat{y} \le 1$. However, all the cooperative equilibria are weakly dominated by the second-order free-riding, $(1, 0)$. Therefore, defecting and refusing to punish is the only strict Nash equilibrium. (See Figure 2.5.1 (a))

2.5.2 Quantal response method

We first normalize the payoff matrix as Eq.(2.2.1), where $p = 1$, $q = \frac{a}{\beta}$, $c = \frac{\alpha}{\beta}$. This is the limit case of the coordination game. From Eq.(2.3.1), the limit QRE set consist of three equilibria only, one asocial equilibrium $(0, 0)$ and two cooperative equilibria $(1, \frac{1}{2})$ and $(1, q)$. Similarly as Theorem 2.2, we can easily prove that the graph of Eq.(2.3.8) consists of two branches for almost all parameters, where the Nash equilibrium $(1, q)$ and

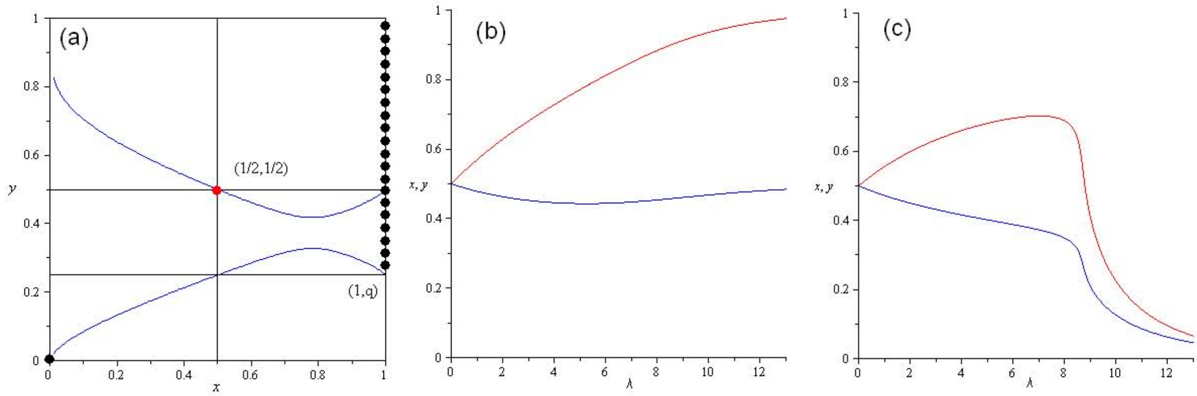


Figure 2.5.1: Logit equilibrium correspondence for the public goods game with punishment. $a = \alpha = 1$, $b = 2$, $\beta = 5$ in Figures (a) and (b), and $\beta = 4$ in Figure (c). $\lambda \geq 0$ on solid curves but $\lambda < 0$ on dashed curves. Black points are NE and red point is $(\frac{1}{2}, \frac{1}{2})$. Red curves and blue curves denote the frequencies of cooperation and punishment, respectively. (a) The graph of Eq.(2.3.8) consists of two branches, where one passes through the Nash equilibrium $(1, q)$ and the other passes through the centroid $(\frac{1}{2}, \frac{1}{2})$. (b) The limiting logit equilibrium is the cooperative equilibrium $(1, \frac{1}{2})$. (c) The limiting logit equilibrium is the asocial equilibrium $(0, 0)$ but cooperation can dominate the population for small value of λ.

the centroid $(\frac{1}{2}, \frac{1}{2})$ are always on different branches. (See Figure 2.5.1 (a)) By using the linear approximation Eq.(2.3.17), the limiting logit is $(1, \frac{1}{2})$, cooperate in the first stage and punish defectors with probability one half in the second stage, if

$$2\beta > 4a + \alpha + \sqrt{8a\alpha + \alpha^2} \tag{2.5.2}$$

Hence, quantal response methods choose the cooperative equilibrium if the punishment is strong enough. For instance, if the cost of cooperation is equal to the cost of punishment, i.e., $a = \alpha$, Eq.(2.5.2) can be simplified as $\beta > 4\alpha$. The limiting logit equilibrium is the cooperative equilibrium if the punishment/cost ratio (also called the effectiveness of punishment) is greater than four. (See Figure 2.5.1) On the other hand, if the cost of punishment is greatly larger than the cost of cooperation, i.e., $\alpha >> a$, Eq.(2.5.2) reduces to $\beta > \alpha$. In this case, punishment can be selected for lower effectiveness.

Intuitively, we say that cooperation is "dominant" in the population if more than half of all players contribute in the first stage. Denote the frequencies of C and P by x and y, respectively. From Eq.(2.3.10) and Eq.(2.3.11), the maximum value of x on the branch of QRE correspondence starting at the centroid is larger than one half if and only if $q < \frac{1}{2}$. This implies that if $\beta > 2a$, cooperation could dominate the population for some values of λ even if the QRE correspondence eventually converges to the asocial equilibrium. For instance, if $a = \alpha = 1$ and $c = 4$, the limiting logit equilibrium is $(0, 0)$ but the proportion of cooperators can reach 70 percent when $6 < \lambda < 8$. (Figure 2.5.1 (c))

2.5.3 How punishment works

Punishment of defectors is a key point for the explanation of cooperation. A large amount of empirical studies indicate that the effectiveness of punishment plays an important role in rasing contributions in the PGG. For instance, Sefton et al. (2007) used a punishment effectiveness of one and there is no differences between control groups and treatment. However, Fehr and Gachter (2002) considered a punishment effectiveness of three and the opportunity to punish increases cooperation significantly.

Recently, Nikiforakis and Normann (2008) compared PGGs with punishment/cost ratios from 0 to 4. They find that contributions to the public pool increase monotonically in effectiveness. With a punishment effectiveness of two or less, contributions remain constant at best or decline over time. Only a punishment effectiveness of three or four leads to high contribution rates and 4:1 punishment technology works better than other ratios in promoting cooperation. Our model confirms those empirical results, although they cannot be compared directly since the reduction to two players may affect an essential aspect of the PGG game.

Since punishment is often costly, this gives rise to an issue of second-order free-riding. Why people punish defectors? Explanations include reputation (Sigmund et al., 2001;

Ohtsuki et al., 2009; Hilbe and Sigmund, 2010), social norm (Henrich et al., 2006; Herrmann et al., 2008), individual preference (Fehr and Schmidt, 1999) or neurology (De Quervain et al., 2004). Different from above mechanisms, QRE method assumes that players are boundedly rational and adopt the probabilistic best response function. If the fine to defectors is sufficiently large, the frequency of cooperation increases rapidly. A minority of irrational punishers can force a majority of individuals learn to cooperate and population will converge to homogenous cooperation before all punishers turn to second-order free-riders. However, if the punishment is not strong enough, punishers go extinct at first and defectors will finally dominate the population.

2.6 Conclusion

McKelvey and Palfrey (1995) defined the equilibrium selection by tracing the logit equilibrium correspondence. In section 2.3, we extend their idea to the more general structural QRE with admissible noises (Goeree et al., 2005). Admissibility assumes that the errors on the payoff functions are continuous, unbiased and independent across players, which involves a large family of probability distributions, such as logistic distribution and normal distribution. For almost all 2×2 coordination games, the QRE correspondence consists of two (disjoint) branches, where one connects the mixed Nash equilibrium to a pure Nash equilibrium, and the second connects the centroid to another pure Nash equilibrium. The pure Nash equilibrium on the second branch is the limiting QRE. In the critical case where parameters are on the separatrix, two branches intersect at a singular point and tracing the QRE correspondence could reach all Nash equilibria.

For the expression of the separatrix, the first order term depends on the payoff matrix only, where the slope is $-\sqrt{c}$. In section 2.4, we derive the high order terms for six different quantal response functions. Both the power series and the numerical simulations show that the equilibrium selection is affected little by the noise distribution but mainly decided by the payoff matrix. This result is different from the noisy best response approach suggested by Staudigl (2011). He calculated the evolutionary path by the optimal control method, but linear terms of the separatrices for the probit noise and the logit noise are not the same. In the logarithmic game, we find a simple square root rule to decide the limiting QRE: A Nash equilibrium is selected if and only if it has larger sum of square root payoff. In particular, this formula is also a good approximation for the other equilibrium selections discussed in section 2.4 since they have similar separatrices. (See Figure 2.3.1) The square root rule is distinct from the Nash product rule of the risk dominant equilibrium, which is independent of c. Two rules are equivalent if and only if $c = 1$, which means that two types of players have the same learning rate. Unfortunately, Turocy (2005) made a mistake in his paper (Theorem 7).

In this chapter, we consider the structural QRE where noises are independently and identically distributed. However, our results are also true for non-identical noises since the proofs of Theorem 2.1 and 2.2 only need continuity and independence. In calculating the separatrix, we assume that noises follow a unimodal distribution. However, this assumption does not affect the linear approximation Eq.(2.3.17) since the first order term is independent of the noise structure. In addition, our technique also works for quantal response functions without explicit noise structures, such as the logarithmic game of Harsanyi (1973).

Finally, we apply the quantal response method to the public goods game with punishment. Equilibrium analysis indicates that the only strict Nash equilibrium is that players do not contribute to the public pool and do not punish free riders. This contradicts empirical evidence. However, punishment can promote cooperation if players are boundedly rational that make mistakes in evaluating the payoff functions. When punishment is strong enough, a cooperative equilibrium can be selected as the limiting QRE. In this case, a minority of irrational punishers force the population to evolve to homogenous cooperation before all punishers turn to second-order free riders.

Chapter 3

Social learning in the ultimatum game

Abstract

In the ultimatum game, two players divide a sum of money. The proposer suggests how to split and the responder can accept or reject. If the suggestion is rejected, both players get nothing. The rational solution is that the responder accepts even the smallest offer but human prefer the fair share. In this chapter, we study the ultimatum game by a learning-mutation process based on the quantal response equilibrium. Social learning is never stabilized at the fair outcome or the rational outcome, but leads to oscillations from offering 40 percent to 50 percent. More precisely, there is a clear tendency to increase the mean offer if it is lower than 40 percent, but will decrease when it reaches the fair offer.

Key words

Ultimatum game; quantal response equilibrium; learning-mutation process

3.1 Introduction

Ultimatum game introduced by Guth et al. (1982) is one of the most influential games in experimental economics that people in the real world do not behave rational. The setting of the game is quiet simple. Two players divide a sum of money. The proposer makes an offer how to split and the responder decides whether to accept. If the offer is rejected, both players get nothing. A rational responder ought to accept any non-zero offer. Therefore, a selfish proposer who thinks that the responder is rational should offer the minimal. Game theory predicts the rational outcome, however, empirical studies in human society, including both laboratory games and field games, prefer fair outcome. In hundreds of ultimatum games conducted in different countries in last 30 years, proposers on average offer 40 to 50 percent of the total sum to the responder. Responders usually accept offers higher than 40 percent and about half of all responders reject offers below 30 percent (Roth et al., 1995; Kagel and Roth, 1995; Camerer, 2003; Osterbeek and Kuilen, 2004; Cooper and Dutcher, 2010; Henrich et al., 2001, 2006, 2010).

How to understand people rejecting positive offers? One well known explanation is that irrational individuals have preference on fairness (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000). In these models, utility functions of players depend not only on their own payoff but also the payoff of the others. Responders reject low offers because the disutility of receiving a payoff less than the proposer is greater than the utility of getting small monetary benefits. On the other hand, the rejection of a unfair offer can be seen as a kind of punishment that inhibits selfish behaviors in later rounds. In iterated ultimatum game experiments, average offers are much more close to the fair share (Roth et al., 1995; Bolton and Ockenfels, 2000; Brenner and Vriend, 2006; Fischbacher et al., 2009). However, this contradicts the equilibrium analysis since the only subgame perfection is not to reject.

In this chapter, we study the iterated ultimatum game by social learning. To analyze the game, define individual strategy as $S(x, y, p)$, meaning giving x of the total sum to the responder when acting as proposer and rejecting any offer less than y with probability p (and accepting offers equal or higher than y with probability 1) when acting as responder, where $0 \leq x \leq 1$ and $0 < p \leq 1$. Following this definition, the rational strategy is written as $S(e, e, 1)$, where e is the minimum offer greater than 0, and the fair strategy is $S(\frac{1}{2}, \frac{1}{2}, 1)$. In our model, individuals strategies will be updated by a learning-mutation process.

3.2 The model

3.2.1 Mini ultimatum game

Before studying the ultimatum game with continuous strategies, we first consider the iterated mini ultimatum game with only two possible offers h and l, with $0 \leq l < h \leq 1$ (Gale et al., 1995; Nowak et al., 2000; Sigmund et al., 2001; Abbink, 2001; Falk et al., 2003). In each round, the proposer has to choose between the high offer h (labeled by H) or the low offer l (labeled by L), and the responder has to decide to reject the low offer l (labeled by H) or accept (labeled by L). The payoff matrix is then written as

$$\begin{array}{cc} & \begin{array}{cc} H & L \end{array} \\ \begin{array}{c} H \\ L \end{array} & \begin{pmatrix} 1-h, h & 1-h, h \\ 0, 0 & 1-l, l \end{pmatrix} \end{array} \tag{3.2.1}$$

where the proposer plays rows and the responder plays columns. [1] The mini game has a strict Nash equilibrium, (L, L), and non-isolated Nash equilibria, $(H, sH + (1-s)L)$, where $\frac{h-l}{1-l} \leq s \leq 1$. Since each equilibrium $(H, sH + (1-s)L)$ is weakly dominated by (H, L), (L, L) is the only subgame perfection. Therefore, rational players will choose (L, L) according to backward induction.

3.2.2 Quantal response equilibrium

There are many ways to model social learning (e.g., Selten and Stoecker, 1986; Gale et al., 1995; Roth and Erev, 1995; Abbink, 2001; Kirman and Virend, 2001). In this chapter, we study the quantal response equilibrium (QRE, also called the perturbed best response) introduced by McKelvey and Palfrey (1995, 1998; Goeree et al., 2005; Yi, 2005; Sandholm, 2010). In a quantal response equilibrium, players are assumed to be boundedly rational. They observe random perturbations on the payoffs of strategies and choose optimally according to those noisy observations. The most common specification of QRE is the logit equilibrium, where noises follow the extreme value distribution (Blume, 1993, 1995; Turocy, 2005). Let u_{ij} denotes the expected payoff of player i using strategy j, where $j = 1, , J_i$. For any given $\lambda \geq 0$, the logistic response function is defined as

$$\sigma_{ij}(\bar{u}_i) = \frac{e^{\lambda u_{ij}}}{\sum_{k=1}^{J_i} e^{\lambda u_{ik}}} \tag{3.2.2}$$

, where σ_{ij} is the probability that player i adopts strategy j and $\bar{u}_i = (u_{i1}, ..., u_{iJ_i})$. If each player uses a logistic response function, QRE or logit equilibria are the solutions of

[1] Payoff matrix Eq.(3.2.1) can also be explained as the Prisoner's Dilemma game with punishment (see Eq.(2.5.1) in chapter 2), where H and L correspond to Cooperation and Defection, and P means paying l to punish defector $1-l$. Similarly as the mini ultimatum game, (L, N) is the only subgame perfection.

$\pi_{ij} = \sigma_{ij}$, where π_{ij} is the frequency of strategy j in player i. λ has been interpreted as the intensity of experience (McKelvey and Palfrey, 1995). At $\lambda = 0$, players have no information about the game and each strategy is chosen with equal probability. As λ approaches infinity, players achieve full information about the game and play the best response.

The quantal response method has been widely used to explain experimental data. In iterated games, estimates of λ usually increase as the game progresses (McKelvey and Palfrey, 1995, 1998). As players gain experience from repeated observations, they can be expected to make more precise estimates and finally reach a Nash equilibrium. To describe this process, consider QRE as a function of λ. When $\lambda = 0$, the QRE is at the centroid of the simplex and when $\lambda = +\infty$, the QRE set consists of Nash equilibria only. As pointed out by McKelvey and Palfrey (1995), for almost all norm form games, the graph of the logit equilibria correspondence contains a unique branch which starts for $\lambda = 0$ at the centroid and converges to a unique Nash equilibrium as λ goes to infinity. This then defines a unique selection from the set of Nash equilibrium by "tracing" the graph of the logit equilibrium correspondence starting at the centroid. The selected Nash equilibrium is called the *limiting logit equilibrium* (LLE) of the game.

Following subsection 2.5.2 in chapter 2, for almost all mini ultimatum games, the LLE is one of two Nash equilibria only, either (L, L), giving the low offer and accepting the low offer, or $(H, \frac{H}{2} + \frac{L}{2})$, giving the high offer and rejecting the low offer with probability one half. Approximately, $(H, \frac{H}{2} + \frac{L}{2})$ is the LLE if and only if

$$2h < l + 1 - \sqrt{l(1-l)} \tag{3.2.3}$$

In the calculation of Eq.(3.2.3), high order terms of the Taylor expansion are ignored, which relate to the quantal response function. Subsection 2.4.1 in chapter 2 shows that coefficients of high order terms are very small therefore affect little to the equilibrium selection.

If the high offer is the fair offer, i.e., $h = \frac{1}{2}$, Eq.(3.2.3) tells us that social learning chooses the low offer. In fact, any high offer equal of greater than $\frac{1}{2}$ is unfavored. On the other hand, if the low offer is the rational decision, i.e., $l = e$, any high offer smaller than $\frac{1-\sqrt{e}}{2}$ is selected. Therefore, social learning does not always choose the rational outcome. For convenience, we say that offer x_1 *dominates* offer x_2 if x_1 is the LLE of the mini ultimatum game with offers x_1 and x_2. Dominant regions of x_1 and x_2 are shown in Figure 3.2.1. Offers lower than $\frac{1}{2}$ are dominated by slightly higher offers. For $l < \frac{1}{2}$, the right side of Eq.(3.2.3) is a convex function, where at the minimum $l^* = \frac{2-\sqrt{2}}{4} \approx 0.15$ and $h^* = \frac{3-\sqrt{2}}{4} \approx 0.4$. This implies that if $h^* \leq x < \frac{1}{2}$, x is also dominated by some low offers. In particular, h^* dominates almost all lower offers (the only exception is l^*).

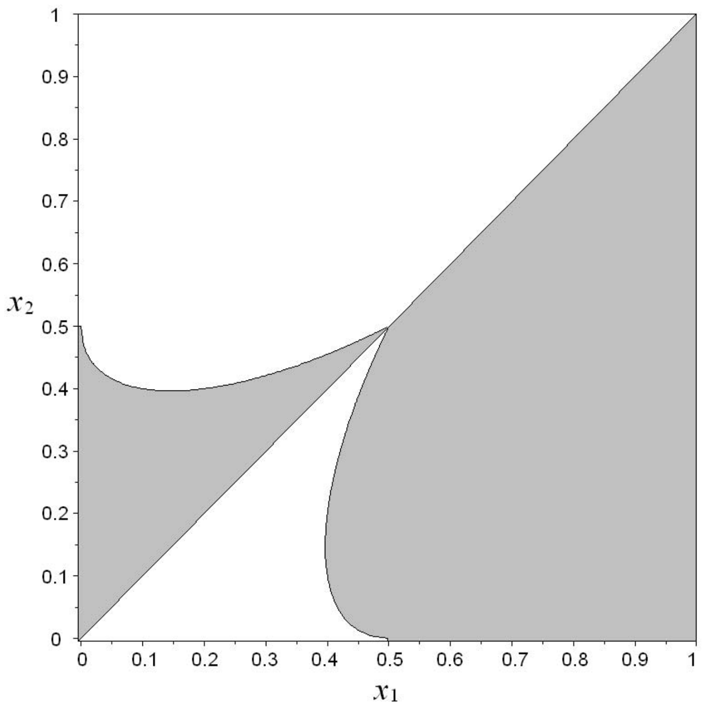


Figure 3.2.1: Pairwise invasibility plot. x_1 and x_2 are dominant in white and gray regions, respectively. Every offer x_1 lower than 0.5 is dominated by some higher offers and strategies equal or greater than 0.4 are also dominated by lower offers.

3.2.3 Learning-mutation process

Let us now introduce the learning-mutation process on the continuum of all strategies. Consider a population of n players. In each generation, players are randomly anonymously paired and play the iterated mini ultimatum game. In each group, roles of two members are decided randomly before the game starts and do not change in an interaction. Two players update their strategies by the quantal response learning and the interaction will stop if they reach a Nash equilibrium since in this situation both are unwilling to change. Mutations happen after all the groups reach Nash equilibria. With probability μ, players adopt a new strategy, plus or minus a small random value on their former strategies. (See Figure 3.2.2 for an example)

We first look at the learning process in one generation. Denote the mini ultimatum game where the proposer using strategy $S(x_1, y_1, p_1)$ and the responder using strategy $S(x_2, y_2, p_2)$ by $UG(S(x_1, y_1, p_1), S(x_2, y_2, p_2))$. In this game, the proposer offers x_1 and the responder rejects offers lower than y_2 with probability p_2. Without loss of generality, suppose that y_2 is the high offer h and x_1 is the low offer l. At the beginning, both players have the motivation to adjust their original strategies. That is, the proposer tends to increase his offer from x_1 to y_2 in order to avoid being refused, and meanwhile, the responder tends to decrease his acceptance level from y_2 to x_1. [2] Eq.(3.2.3) provides an approximated rule to decide the LLE. At the LLE, the responder either accepts the low offer or rejects the low offer with probability one half. Therefore, we always take $p = \frac{1}{2}$ and write $S(x, y) = S(x, y, \frac{1}{2})$ in later discussions. At the end of iterated game $UG(S(x_1, y_1), S(x_2, y_2))$, if x_1 dominates y_2, the proposer keeps his strategy unchanged but the responder adopts a new strategy $S(x_2, x_1)$. In contrast, if y_2 dominates x_1, the responder's strategy does not change but the proposer adopts a new strategy $S(y_2, y_1)$. We observe that learning always decreases the diversity of possible offers since in each mini game, one dominated offer is eliminated.

Next, consider a population that consists only of $S(x, y)$ players evolves under the learning (no mutation) process. At the end of iterated game $UG(S(x, y), S(x, y))$, strategies of two players are $S(x, y)$ and $S(x, x)$ if x dominates y, and are $S(y, y)$ and $S(x, y)$ if y dominates x. In the first case, the population will converge to a pure $S(x, x)$ population, while in the second case, it will converge to a pure $S(y, y)$ population. More generally, starting with any mixed population, the learning process will always lead to a homogenous population where the offer equals to the acceptance level since the diversity of possible offers is monotonically decreased.

We now analyze the case of weak mutation rate. As in the adaptive dynamics model, mutations occur rarely, so that a mutant will either vanish or has taken over the popu-

[2] This statement is also true if $x_1 > y_2$. In this case, the proposer wants to decrease his offer from x_1 to x_2 and the responder wants to increase his acceptance level from x_2 to x_1.

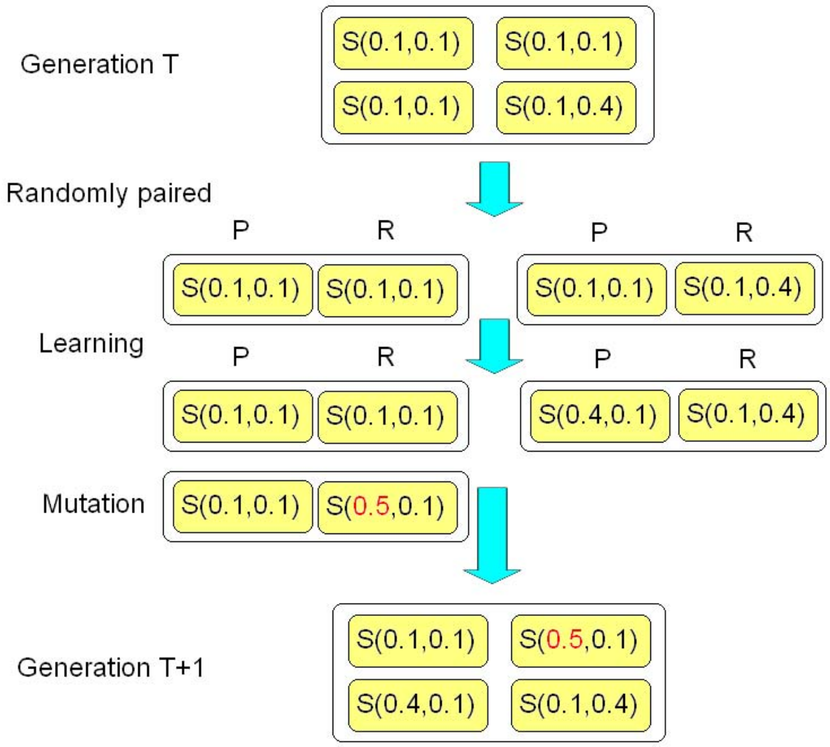


Figure 3.2.2: An example for the learning-mutation process in a population of four players from generation T to generation $T+1$. In generation T, three players adopt $S(0.1, 0.1)$ and one adopts a mutant strategy $S(0.1, 0.4)$. At the beginning, they are divided into two mini ultimatum games, $UG(S(0.1, 0.1), S(0.1, 0.1))$ and $UG(S(0.1, 0.1), S(0.1, 0.4))$, and update their strategies by the quantal response learning (P means proposers and R means responders). In the first group, players will not change their original strategies, while in the second group, the proposer will change the strategy to $S(0.4, 0.1)$ since 0.4 dominates 0.1. Mutations happen after all the pairs reach Nash equilibria. The responder in the first pair mutates to $S(0.5, 0.1)$ (the red number). As a result of learning and mutation, strategies in generation $T+1$ are $S(0.1, 0.1)$, $S(0.5, 0.1)$, $S(0.4, 0.1)$ and $S(0.1, 0.4)$.

lation before the next mutation arises (Hofbauer and Sigmund 1998; Geritz et al., 1998). For simplicity, we represent the strategy of the residents by $R(x) = S(x, x)$. Eq.(3.2.3) indicates that (a) if $R(x) < h^*$, the population could only be replaced by mutants using offer higher than $R(x)$, (b) if $h^* \leq R(x) < \frac{1}{2}$, both higher and lower offers may invade, (c) if $\frac{1}{2} \leq R(x)$, any lower offer could take over the population (See Figure 3.2.1). Generally speaking, the learning-mutation process leads to oscillations in the interval $[h^*, \frac{1}{2})$, where proposers offer 40 to 50 percent of the total sum to responders and responders reject offers below their expectation with probability one half. Once the resident strategy leaves the interval, learning and mutation will push it back. If we further assume that the mutational jumps are very small such that the resident strategy changes continuously, it is easy to verify that $\frac{dR(x)}{dt} > 0$ if $R(x) < \frac{1}{2}$ but $\frac{dR(x)}{dt} < 0$ if $R(x) \geq \frac{1}{2}$. $R(x) = \frac{1}{2}$ is a degenerate point of the adaptive dynamics, i.e., the resident strategy will decrease when near the fair offer.

Numerical simulations suggest that theoretical predictions of the weak mutation rate case could also be applied to describe the high mutation rate case, where the population has a high diversity of strategies. That is, the population mean offer and the mean acceptance level are nearly the same, and the mean offer increases if it is smaller than 0.4 but oscillates if it is between 0.4 to 0.5. Moreover, if the mutational jumps are very small, the mean offer converges to a interval very close to the fair offer. (See Figure 3.2.3)

3.3 Discussion

In the model, players have no information about the mini ultimatum game (i.e., $\lambda = 0$) at the beginning of each new interaction, no matter how many interactions they have played. The motivation is twofold. On the one hand, each player faces a new game in a new interaction since his opponent is anonymous and the payoff matrix of the mini ultimatum game depends on the strategies of both players. On the other hand, empirical evidences from repeated PD games (with punishment) [3] support this consideration. The frequency of cooperation in the initial round of each interaction is nearly the same and decreases over the rounds (Selten and Stoecker, 1986; Kagel and Roth, 1995; Dreber et al., 2008; Wu et al., 2009). We can then expect that players have zero information before each interaction and evolve entirely by social learning.

We consider that players are boundedly rational and choose the best response according to noisy observations. In each group, two players update their strategies simultaneously. At the beginning, the proposer is inclined to make the high offer due to the high

[3] The payoff matrix of the PD game is equivalent to that of the mini dictator game, which is a variation of the mini ultimatum game where responders are not allowed to reject. Selten and Stoecker (1986) and Kagel and Roth (1995) summarized the results on repeated PD games. Dreber et al. (2008) and Wu et al. (2009) are two recent studies on repeated PD games with punishment.

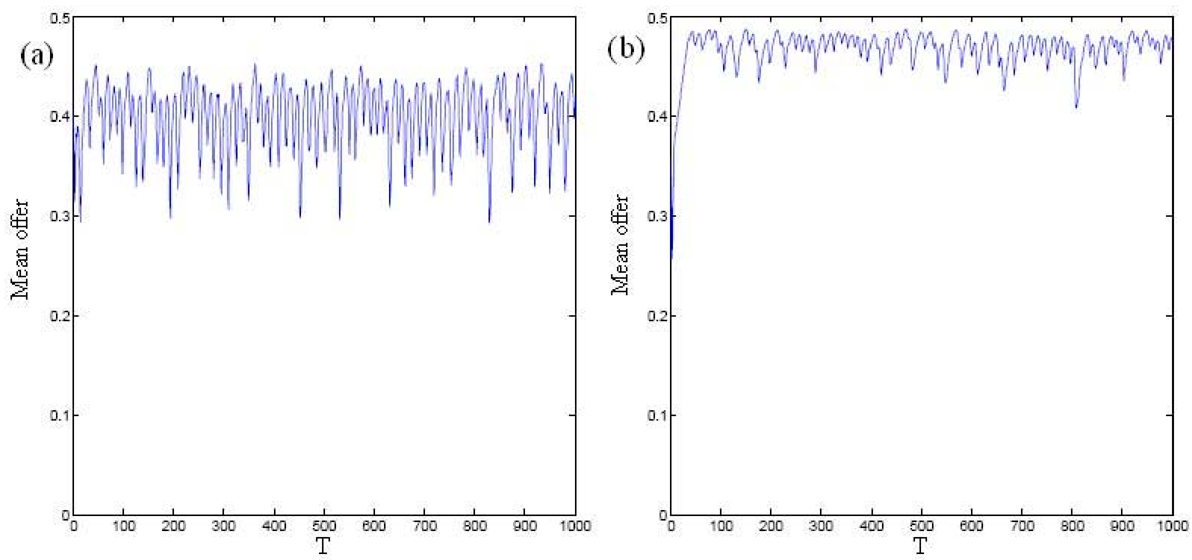


Figure 3.2.3: Time evolution of the population mean offer. The population size is 100 and evolves under the learning-mutation process. At the end of each generation, players adopt a new offer with probability 0.1. The mutational jumps follow the normal distribution, where variances in Figures (a) and (b) are 0.05 and 0.01, respectively. (a) The population mean offer increases if it is smaller than 0.4 but oscillates if it is between 0.4 to 0.5. (b) If the mutational jumps are very small, the mean offer converges to a interval very close to the fair offer.

rejection rate and the responder tends to accept the lower offer since rejecting is costly. The observation errors decrease as the game progresses and two players will finally reach a Nash equilibrium. Intuitively, their strategies converge to the high offer if the proposer learns faster than the responder, i.e., the proposer stops making the low offer before the responder stops rejecting. This happens when the low offer is small, which means the rejection of the low offer causes a greater loss to the proposer than to the responder. Thus, mistakes in evaluating the payoff functions lead to fairer solutions.

The emergence of equity is as complicated as the evolution of human society. Our model excluded many important issues, such as preference on fairness (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000) or on punishment (Charness and Rabin, 2002; Falk and Fischbacher, 2006), communication or information before an interaction (Levine, 1998; Nowak et al, 2000; Sigmund et al., 2001) and social networks (Page et al., 2000; Jong et al., 2008). Based on learning and mutation, we show that individuals entirely motivated by self interests can evolve toward fairness in the population.

Chapter 4

Equilibrium selections via replicator dynamics

Abstract

This chapter studies two equilibrium selection methods based on the replicator dynamics. A Nash equilibrium is called centroid dominant if the trajectory of the replicator dynamics starting at the centroid of the strategy simplex converges to it. On the other hand, an equilibrium is called basin dominant if it has the largest basin of attraction. Two concepts are compared with the risk dominant equilibrium in the context of 2×2 bimatrix coordination games. Main results include (a) if a Nash equilibrium is both risk dominant and centroid dominant, it must have the largest basin of attraction, (b) the basin dominant equilibrium must be either risk dominant or centroid dominant.

Key words

Equilibrium selection; replicator dynamics; risk dominance; basins of attraction

4.1 Introduction

One well known dynamic approach in evolutionary game theory is the replicator dynamics (Taylor and Jonker, 1978; Maynard Smith, 1982; Hofbauer and Sigmund, 1998). Replicator dynamics was first motivated biologically in the context of evolution (Maynard Smith, 1982; Nowak and Sigmund, 2004). Later, economists related this to learning and defined several equilibrium notions (Fudenberg and Harris, 1992; Samuelson and Zhang, 1992; Gale et al., 1995; Weibull, 1995; Binmore and Samuelson, 1997; Samuelson, 1997; Borgers and Sarin, 1997; Schlag, 1998; Cabrales, 2000; Imhof, 2005; Hilbe, 2010). These researches usually consider modified replicator dynamics that incorporate stochastic effects such as errors, mutations or finite populations. In this chapter, we study two equilibrium selections based on the canonical replicator dynamics.

The first is a homotopy approach by tracing the trajectory of the replicator dynamics starting at the centroid. For 2×2 coordination games (both symmetric games and bimatrix games), the trajectory approaches a unique Nash equilibrium as $t \to \infty$ (Hofbauer and Sigmund, 1998). This then defines a unique equilibrium selection from the set of Nash equilibria. We call this equilibrium the centroid dominant equilibrium of the game. A biological intuition is that natural selection leads to the centroid dominant equilibrium if each phenotype in the population has equal frequency. From the perspective of learning, if players choose their initial strategies randomly and imitate actions that perform better with a probability proportional to the expected payoff, the population will converge to the centroid dominant equilibrium.

The second method is to select a Nash equilibrium from the set of asymptotically stable equilibria of the replicator dynamics by comparing their basins of attraction. A Nash equilibrium is called basin dominant if it has the largest basin of attraction. This implies that a population with uncertain initial state is more likely to converge to the dominant equilibrium under the replicator dynamics.

For 2×2 symmetric coordination games, one can easily verify that the center point is attracted by the risk dominant equilibrium. Hence, the risk dominant equilibrium is both centroid dominant and basin dominant. In fact, in 2×2 symmetric games, most of the equilibrium notions we mentioned above choose the risk dominant equilibrium (as an exception, Binmore and Samuelson, 1997), but they usually select different equilibria in more general situations (Kim, 1996; Samuelson, 2002).

In this chapter, we focus on 2×2 bimatrix coordination games. Section 4.2 reviews the risk dominant equilibrium and the replicator dynamics. Section 4.3 studies the centroid dominance and derives an explicit formula for the centroid dominant equilibrium. Section 4.4 investigates the basin dominant equilibrium and shows some properties. Section 4.5 compares the three equilibrium notions and looks for their relations. Section 4.6 summarizes the main results and suggests some further developments.

4.2 Bimatrix games

Consider a 2×2 bimatrix game, where the two pure strategies of players in population A (call them player A) are A_1 and A_2, and two pure strategies of players in population B (call them player B) are B_1 and B_2. Let a_{ij} denotes the payoff to player A using strategy A_i when it meets player B using strategy B_j, and denote the payoff to player B in this interaction by b_{ji}. The payoff matrix is then written as

$$\begin{array}{cc} & \begin{array}{cc} B_1 & B_2 \end{array} \\ \begin{array}{c} A_1 \\ A_2 \end{array} & \begin{pmatrix} a_{11}, b_{11} & a_{12}, b_{21} \\ a_{21}, b_{12} & a_{22}, b_{22} \end{pmatrix} \end{array} \tag{4.2.1}$$

The bimatrix game is called a coordination game if pure strategy pairs (A_1, B_1) and (A_2, B_2) are both strict Nash equilibria. That is, $a_{11} - a_{21} > 0$ and $a_{22} - a_{12} > 0$ for player A, and $b_{11} - b_{21} > 0$ and $b_{22} - b_{12} > 0$ for player B. Besides of two strict Nash equilibria, the game also has a mixed equilibrium (p, q), where $p = \frac{b_{22}-b_{12}}{b_{11}+b_{22}-b_{21}-b_{12}}$ and $q = \frac{a_{22}-a_{12}}{a_{11}+a_{22}-a_{21}-a_{12}}$.

As introduced by Harsanyi and Selten (1988), for the 2×2 coordination game, (A_1, B_1) is said to risk dominate (A_2, B_2) if the Nash products satisfy

$$(a_{11} - a_{21})(b_{11} - b_{21}) > (a_{22} - a_{12})(b_{22} - b_{12}) \tag{4.2.2}$$

Define $L_{RD} : p + q = 1$. Hence, (A_1, B_1) is risk dominant (RD) if and only if (p, q) is below the line L_{RD} on $p - q$ plane.

Denote the frequency of strategy A_1 in A players' population and strategy B_1 in B players' population by x and y, respectively. The replicator dynamics for the bimatrix game Eq.(4.2.1) is

$$\begin{aligned} \frac{dx}{dt} &= x(1-x)((a_{12} - a_{22})(1-y) + (a_{11} - a_{21})y) \\ \frac{dy}{dt} &= y(1-y)((b_{12} - b_{22})(1-x) + (b_{11} - b_{21})x) \end{aligned} \tag{4.2.3}$$

Eq.(4.2.3) could be normalized as

$$\begin{aligned} \frac{dx}{dt} &= x(1-x)(y-q) \\ \frac{dy}{dt} &= cy(1-y)(x-p) \end{aligned} \tag{4.2.4}$$

, where (p, q) is the mixed equilibrium and $c = \frac{b_{11}+b_{22}-b_{21}-b_{12}}{a_{11}+a_{22}-a_{21}-a_{12}} > 0$.

It is easy to see that Eq.(4.2.4) has four boundary equilibria, $(0, 0)$, $(1, 0)$, $(0, 1)$ and $(1, 1)$, and one interior equilibrium, (p, q). For their stabilities, $(0, 0)$ and $(1, 1)$ are locally asymptotically stable that correspond to the two strict Nash equilibria (A_2, B_2) and (A_1, B_1), $(1, 0)$ and $(0, 1)$ are unstable, and the mixed equilibrium (p, q) is a saddle point (Hofbauer and Sigmund, 1998).

4.3 Centroid dominance

For 2×2 bimatrix coordination games, solutions of Eq.(4.2.4) includes a unique path starting for $t = 0$ at the centroid $(\frac{1}{2}, \frac{1}{2})$ and converging to an Nash equilibrium as $t \to \infty$ (Hofbauer and Sigmund, 1998). This implies that we can define a unique selection from the set of Nash equilibria by tracing the trajectory of the replicator dynamics. An equilibrium is called centroid dominant (CD) if and only if the solution of Eq.(4.2.4) with initial value at the centroid converges to it.

For convenience, denote the basins of attraction of two stable equilibria $(0, 0)$ and $(1, 1)$ by S_0 and S_1, respectively. This means that trajectories of Eq.(4.2.4) with initial points in region S_0 converge to $(0, 0)$ and with initial points in region S_1 converge to $(1, 1)$. However, if a initial point is on the curve separating S_0 and S_1 (we call this curve the *separatrix*), the trajectory goes to neither $(0, 0)$ nor $(1, 1)$ but to the mixed equilibrium (p, q). Hence, $(0, 0)$ is CD if and only if $(\frac{1}{2}, \frac{1}{2}) \in S_0$ and $(1, 1)$ is CD if and only $(\frac{1}{2}, \frac{1}{2}) \in S_1$. In another word, which equilibrium is selected is decided by the position of the separatrix, i.e., if it is above the centroid, $(0, 0)$ is CD, and if it is below the centroid, $(1, 1)$ is CD. (See Figure 4.3.1)

Therefore, we turn our attention to the separatrix in the rest of this section. For given (p, q, c), denote the separatrix by

$$L_{(p,q,c)} : y = l_{(p,q,c)}(x). \tag{4.3.1}$$

Intuitively, $L_{(p,q,c)}$ consists of two trajectories of Eq.(4.2.4), where one from $(1, 0)$ to (p, q) and another from $(0, 1)$ to (p, q), i.e., points on $L_{(p,q,c)}$ satisfy

$$\frac{dx}{dy} = \frac{x(1-x)(y-q)}{cy(1-y)(x-p)} \tag{4.3.2}$$

This implies that $l_{(p,q,c)}(x)$ is monotonically decreasing in x and $(x - p)(y - q) \leq 0$.

Let us now derive an expression of the separatrix. From Eq.(4.3.2),

$$(1-x)^{c(1-p)} x^{cp} = K(1-y)^{1-q} y^q \tag{4.3.3}$$

where K is a constant depending on the initial point. Notice that the separatrix passes through (p, q),

$$(\frac{1-x}{1-p})^{c(1-p)} (\frac{x}{p})^{cp} = (\frac{1-y}{1-q})^{1-q} (\frac{y}{q})^q \tag{4.3.4}$$

For convenience, define

$$F_{(p,q,c)}(x, y) = (\frac{1-x}{1-p})^{c(1-p)} (\frac{x}{p})^{cp} - (\frac{1-y}{1-q})^{1-q} (\frac{y}{q})^q \tag{4.3.5}$$

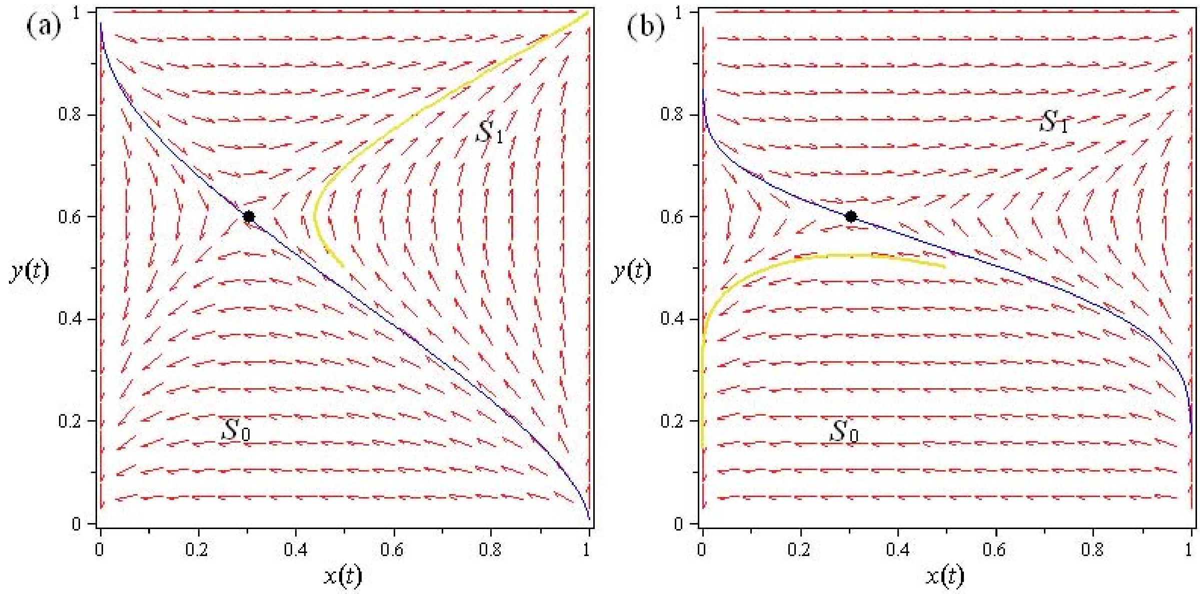


Figure 4.3.1: Phase portraits of the replicator dynamics. Parameters are taken as $p = 0.3$, $q = 0.6$, $c = 0.5$ in Figure (a) and $c = 0.1$ in Figure (b). Black points are the mixed equilibrium $(0.3, 0.6)$, blue curves are the separatrices, and yellow curves are the trajectories of the replicator dynamics with initial value $(\frac{1}{2}, \frac{1}{2})$. $(1, 1)$ is RD but CD equilibrium and BD equilibrium depend crucially on c. In Figure (a), $(1, 1)$ is CD and BD, and in Figure (b), $(0, 0)$ is CD and BD.

It is clearly that $F_{(p,q,c)} = 0$ for each point on $L_{(p,q,c)}$, but we need to be careful that not all solutions of $F_{(p,q,c)} = 0$ are on $L_{(p,q,c)}$, e.g., pure strategy equilibria $(0,0)$ and $(1,1)$. A point (x,y) is on the separatrix if and only if both conditions $F_{(p,q,c)}(x,y) = 0$ and $(x-p)(y-q) \leq 0$ hold.

For separatrices with different (p,q,c), we have the following lemma.

Lemma 4.1

(a) $l_{(p,q,c)}(x) > l_{(\hat{p},q,c)}(x)$ if and only if $\hat{p} < p$. (b) $l_{(p,q,c)}(x) > l_{(p,\hat{q},c)}(x)$ if and only if $\hat{q} < q$.

Proof:

(a) Notice that (p,q) and $(\hat{p},q)$ are on $L_{(p,q,c)}$ and $L_{(\hat{p},q,c)}$, respectively, we only need to show that $L_{(p,q,c)}$ and $L_{(\hat{p},q,c)}$ have no intersection in $(0,1) \times (0,1)$.

For any (x,y),

$$F_{(p,q,c)}(x,y) - F_{(\hat{p},q,c)}(x,y) = F_{(p,q,c)}(x,0) - F_{(\hat{p},q,c)}(x,0) \tag{4.3.6}$$

, where $F_{(p,q,c)}(0,0) = F_{(p,q,c)}(1,0) = F_{(\hat{p},q,c)}(0,0) = F_{(\hat{p},q,c)}(1,0) = 0$. Notice that

$$\frac{d(\ln F_{(p,q,c)}(x,0) - \ln F_{(\hat{p},q,c)}(x,0))}{dx} = c\frac{(p-\hat{p})(1-2x)}{x(1-x)} \tag{4.3.7}$$

, $F_{(\hat{p},q,c)}(x,0) - F_{(p,q,c)}(x,0) \neq 0$ for $0 < x < 1$. This implies that $F_{(p,q,c)}(x,y) \neq F_{(\hat{p},q,c)}(x,y)$ for $0 < x < 1$, i.e., $L_{(p,q,c)}$ and $L_{(\hat{p},q,c)}$ have no intersection.

The proof of (b) is similar. □

From Lemma 4.1, we obtain Theorem 4.1.

Theorem 4.1

For any given c, the mixed equilibrium (p,q) is CD if and only if (p,q) satisfies $F_{(p,q,c)}(\frac{1}{2},\frac{1}{2}) = 0$, where $(\frac{1}{2}-p)(\frac{1}{2}-q) \leq 0$. This defines a curve on $p-q$ plane, denote it by L_{CD}. $(0,0)$ is CD if and only if (p,q) is above L_{CD} and $(1,1)$ is CD if and only if (p,q) is below L_{CD}.

Proof:

The mixed equilibrium (p,q) is CD if and only if $(\frac{1}{2},\frac{1}{2})$ is on the separatrix, i.e., $F_{(p,q,c)}(\frac{1}{2},\frac{1}{2}) = 0$ and $(\frac{1}{2}-p)(\frac{1}{2}-q) \leq 0$.

Lemma 4.1 indicates that for given (p,q,c), if $(\frac{1}{2},\frac{1}{2})$ is attracted by $(0,0)$, $(\frac{1}{2},\frac{1}{2})$ is also attracted by $(0,0)$ for $(p,\hat{q},c)$, where $\hat{q} > q$. This implies the second part of the theorem. □

In addition to Theorem 4.1, an explicit formula for deciding the pure strategy CD equilibrium can be summarized as follows.

Corollary 4.1

Nash equilibrium $(0,0)$ *is CD if (a)* $p > \frac{1}{2}$ *and* $q > \frac{1}{2}$*, or (b)* $p < \frac{1}{2}$, $q > \frac{1}{2}$ *and* $F_{(p,q,c)}(\frac{1}{2},\frac{1}{2}) > 0$*, or (c)* $p > \frac{1}{2}$, $q < \frac{1}{2}$ *and* $F_{(p,q,c)}(\frac{1}{2},\frac{1}{2}) < 0$*. Nash equilibrium* $(1,1)$ *is CD if (d)* $p < \frac{1}{2}$ *and* $q < \frac{1}{2}$*, or (e)* $p < \frac{1}{2}$, $q > \frac{1}{2}$ *and* $F_{(p,q,c)}(\frac{1}{2},\frac{1}{2}) < 0$*, or (f)* $p > \frac{1}{2}$, $q < \frac{1}{2}$ *and* $F_{(p,q,c)}(\frac{1}{2},\frac{1}{2}) > 0$*.*

Theorem 4.1 claims that $p - q$ plane is divided into two regions by

$$L_{CD} : (1-p)^{c(1-p)}p^{cp} = 2^{1-c}(1-q)^{1-q}q^q \tag{4.3.8}$$

, where $(p-\frac{1}{2})(q-\frac{1}{2}) \leq 0$. By applying the implicit function theorem, several properties of L_{CD} can be verified easily.

Corollary 4.2

For any given $c > 0$*, (a)* L_{CD} *is monotonically decreasing in* $p - q$ *plane; (b)* $(p,q) = (\frac{1}{2},\frac{1}{2})$ *is on* L_{CD} *and the slope at this point is* $-\sqrt{c}$*. (c) If* $c = 0$*,* L_{CD} *is* $q = \frac{1}{2}$*; if* $c = 1$*,* L_{CD} *matches* $L_{RD} : p + q = 1$*; if* $c \to \infty$*,* L_{CD} *is* $p = \frac{1}{2}$*. (d) For* $c \neq 1$*,* $(\frac{1}{2},\frac{1}{2})$ *is the only intersection of* L_{CD} *and* L_{RD}*.*

Corollary 4.2 points out that if $c = 1$, the trajectory of the replicator dynamics starting at the centroid always converges to the RD equilibrium. Intuitively, $c = 1$ can be understood as that payoffs for two types of players are equally weighted. However, L_{RD} and L_{CD} are no longer identical for any $c \neq 1$. Figure 4.3.1 shows clearly that the CD equilibrium depends crucially on c in contrast to the RD equilibrium.

4.4 Basin dominance

In this section, an equilibrium is considered dominant if and only if it has the largest basin of attraction. To formulate this definition, let $s_0(p,q,c)$ and $s_1(p,q,c)$ denote the sizes of the basins of attraction of $(0,0)$ and $(1,1)$, respectively. Using the notations in previous sections, $s_0(p,q,c)$ can be calculated by the integral

$$s_0(p,q,c) = \int_0^1 l_{(p,q,c)}(x)dx. \tag{4.4.1}$$

Hence, $(0,0)$ is basin dominant (BD) if and only if $s_0(p,q,c) > \frac{1}{2}$ and $(1,1)$ is BD if and only if $s_0(p,q,c) < \frac{1}{2}$.

The main goal of this section is to find a function f^*, where $q = f^*(p, c)$, such that $s_0(p, f^*(p, c), c) = \frac{1}{2}$. Theorem 4.2 guarantees the existence and uniqueness of f^*.

Theorem 4.2

There exists a unique continuous function $q = f^*(p, c)$ such that $s_0(p, q, c) = \frac{1}{2}$, where it has following properties: (a) f^* is a decreasing function of p, (b) $s_0(p, q, c) > \frac{1}{2}$ if and only if $q > f^*(p, c)$, (c) f^* is central symmetric, i.e., $1 - q = f^*(1 - p, c)$.

Proof:

To show the existence and uniqueness of f^*, it is enough to prove that $s_0(p, q, c)$ is continuously increasing in q. From Eq.(4.4.1), $s_0(p, q, c)$ is continuously increasing in q if the separatrix $y = l_{(p,q,c)}(x)$ is continuously increasing in q for any given p, c and x. From Eq.(4.3.4), the continuity is obvious. To see the monotonicity, we calculate the derivative of $l_{(p,q,c)}(x)$

$$
\begin{aligned}
\frac{dy}{dq} &= -\frac{\partial F_{(p,q,c)}/\partial q}{\partial F_{(p,q,c)}/\partial y} \\
&= -\frac{(1-y)^{1-q} y^q}{(1-q)^{1-q} q^q} \ln(1 + \frac{y-q}{(1-y)q}) / \frac{(q-y)(1-y)^{-q} y^{q-1}}{(1-q)^{1-q} q^q} \\
&= y(1-y) \frac{\ln(1 + \frac{y-q}{(1-y)q})}{y-q} > 0. \qquad (4.4.2)
\end{aligned}
$$

This implies the existence and uniqueness of f^*.

Since $s_0(p, q, c)$ is continuously increasing in q, property (b) is obvious. Similarly as Eq.(4.4.2), it is easy to prove that $l_{(p,q,c)}(x)$ is continuously decreasing in p. This yields property (a). Finally, property (c) is directly from the symmetry of $L_{(p,q,c)}$. $\square$

Interestingly, Theorem 4.2 shows that CD equilibrium selection and BD equilibrium selection have the similar structures. For convenience, denote the curve $q = f^*(p, c)$ by L_{BD}. For given c, properties (a) and (c) say that L_{BD} is monotonically decreasing and divides the $p - q$ plane into two regions, where $(0, 0)$ is selected for (p, q) in the upper region and $(1, 1)$ is selected for (p, q) in the lower region.

In general, deriving an explicit expression of f^* is very difficult since Eq.(4.4.1) is implicit. Alternatively, we calculate L_{BD} for some special parameters. These results will provide an intuition for L_{BD}.

Firstly, we introduce a new notation. Denote the time derivatives for points on curve L by $D(L)$. For instance, on the separatrix $L_{(p,q,c)}$

$$
D(L_{(p,q,c)}) = \frac{d(F_{(p,q,c)}(x, y))}{dt}\Big|_{L_{(p,q,c)}} \qquad (4.4.3)
$$

$D(L) \neq 0$ means that the solutions of the replicator dynamics with initial values on L go away from it. Since the separatrix consists of two solutions, $D(L_{(p,q,c)}) = 0$. On the other hand, if a curve L satisfies $D(L) = 0$ and passes through three points $(0,1)$, $(1,0)$ and (p,q), it must be the separatrix.

Theorem 4.3

(a) If $c \to 0$, $s_0(p,q,c) > \frac{1}{2}$ if and only if $q > \frac{1}{2}$; (b) if $c = 1$, $s_0(p,q,c) > \frac{1}{2}$ if and only if $p + q > 1$; (c) if $c \to \infty$, $s_0(p,q,c) > \frac{1}{2}$ if and only if $p > \frac{1}{2}$.

Proof:

(a)If $c \to 0$, Eq.(4.2.4) becomes

$$\begin{aligned} \frac{dx}{dt} &= x(1-x)(y-q) \\ \frac{dy}{dt} &\to 0 \end{aligned} \tag{4.4.4}$$

Clearly, the separatrix of S_0 and S_1 is $L_{(p,q,c)} : y = q$. Thus, $s_0(p,q,c) > \frac{1}{2}$ if and only if $q > \frac{1}{2}$.

(b)If $c = 1$, Eq.(4.2.4) could be written as

$$\begin{aligned} \frac{dx}{dt} &= x(1-x)(y-q) \\ \frac{dy}{dt} &= y(1-y)(x-p) \end{aligned} \tag{4.4.5}$$

Consider the time derivatives of $L : x + y = 1$,

$$D(L) = x(1-x)(1-p-q) \tag{4.4.6}$$

If $p + q > 1$, points on L are attracted by $(0,0)$ since $D(L) < 0$. This implies $s_0(p,q,c) > \frac{1}{2}$. Conversely, if $p+q < 1$, L is attracted by $(0,0)$, which implies $s_0(p,q,c) < \frac{1}{2}$. For the critical case $p+q = 1$, we have $D(L) = 0$. Notice that L passes through $(0,1)$, $(1,0)$ and (p,q), it is the separatrix of S_0 and S_1. Thus, we have $s_0(p,q,c) = s_1(p,q,c) = \frac{1}{2}$.

(c) The proof is similar to (a). $\square$

Theorem 4.3 studies L_{BD} for extreme values of c. From another angle, we next derive L_{BD} for $p = q = \frac{1}{2}$. In this case, Eq.(4.3.4) becomes

$$2^{2c}((1-x)x)^c = 2^2(1-y)y \tag{4.4.7}$$

Notice that the separatrix goes through $(0,1)$ and $(1,0)$, $L_{(1/2,1/2,c)}$ can be written down explicitly

$$y = \begin{cases} \frac{1+(1-2^{2c}(x(1-x))^c)^{1/2}}{2} & 0 \leq x \leq \frac{1}{2} \\ \frac{1-(1-2^{2c}(x(1-x))^c)^{1/2}}{2} & \frac{1}{2} \leq x \leq 1 \end{cases} \tag{4.4.8}$$

Thus, the size of the basin of attraction of $(0,0)$ is

$$s_0(\frac{1}{2},\frac{1}{2},c) = \int_0^{1/2} \frac{1+(1-2^{2c}(x(1-x))^c)^{1/2}}{2}dx + \int_{1/2}^{1} \frac{1-(1-2^{2c}(x(1-x))^c)^{1/2}}{2}dx = \frac{1}{2} \quad (4.4.9)$$

since

$$\int_0^{1/2} (1-2^{2c}(x(1-x))^c)^{1/2}dx = \int_{1/2}^{1} (1-2^{2c}(x(1-x))^c)^{1/2}dx. \quad (4.4.10)$$

This implies $f^*(\frac{1}{2},c) = \frac{1}{2}$ for any c.

As a conclusion, properties of L_{BD} are summarized in Corollary 4.3.

Corollary 4.3

For any given $c > 0$, (a) L_{BD} is monotonically decreasing in $p-q$ plane; (b) $(p,q) = (\frac{1}{2},\frac{1}{2})$ is on L_{BD} and the slope at this point is $-\sqrt{c}$. (c) If $c = 0$, L_{BD} is $q = \frac{1}{2}$; if $c = 1$, L_{BD} matches $L_{RD} : p + q = 1$; if $c \to \infty$, L_{BD} is $p = \frac{1}{2}$. (d) For $c \neq 1$, $(\frac{1}{2},\frac{1}{2})$ is the only intersection of L_{BD} and L_{RD}.

Properties (a) and (c) are obtained directly from Theorem 4.2 and Theorem 4.3, and we leave the proofs of properties (b) and (d) in Appendix A.1.

Corollary 4.3 implies that RD equilibrium does not always have the largest basin of attraction in 2×2 bimatrix games. If compares Corollary 4.2 and Corollary 4.3, one can find that L_{CD} and L_{BD} have very similar properties. However, since the BD equilibrium does not have an explicit formula, relation between L_{CD} and L_{BD} is still unclear.

4.5 Relations among different notions

Previous sections discuss the relation between L_{RD} and L_{CD} and the relation between L_{RD} and L_{BD}. In this section, we are going to link the three equilibrium notions.

Theorem 4.4

For any given c, L_{BD} is between L_{RD} and L_{CD}. (See the proof in Appendix A.2)

Corollary 4.4

(a) A strategy that is both RD and CD must be BD. (b) The BD strategy must be either RD or CD.

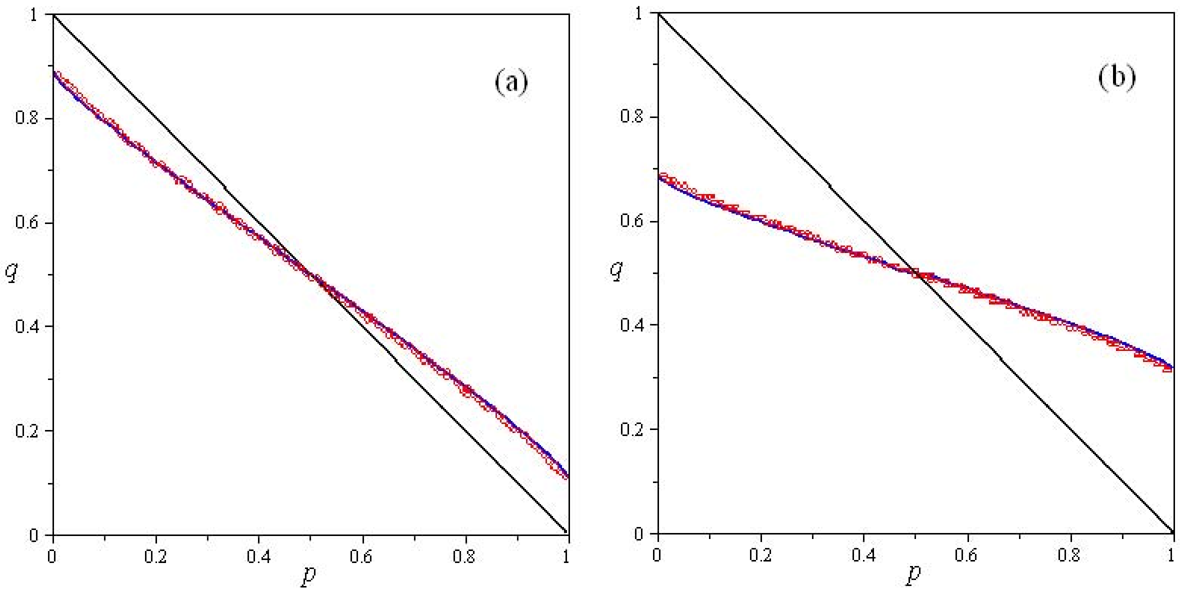


Figure 4.5.1: Relations between RD, CD and BD. Parameters are taken as $c = 0.5$ in Figure (a) and $c = 0.1$ in Figure (b). Black line, blue curve and red points are L_{RD}, L_{CD} and L_{BD}, respectively. L_{BD} is between L_{RD} and L_{CD} and is very close to L_{CD}.

As a complement of Corollary 4.3, Corollary 4.4 provides an alternative way for finding the BD strategy. When facing a 2×2 bimatrix game, we can first calculate the RD equilibrium and CD equilibrium. If two methods point to the same equilibrium, it must be also BD. However, the thought does not work if two methods choose different equilibria. Numerical simulations suggest that L_{BD} is very close to L_{CD}. (See Figure 4.5.1) Therefore, if RD strategy and CD strategy are different, the CD strategy is more likely to have the largest basin of attraction.

4.6 Discussion

In this chapter, we studies two equilibrium notions, centroid dominance and basin dominance, based on the canonical replicator dynamics. A Nash equilibrium is called centroid dominant if the centroid of the strategy simplex is in its basin of attraction. This predicts that in a population where individuals choose their initial strategies randomly, replicator dynamics converge to the centroid dominant equilibrium. On the other hand, a Nash equilibrium is called basin dominant if it has the largest basin of attraction. Following this concept, a population with uncertain initial state has larger probability to evolve to the basin dominant equilibrium.

We then compare them with the risk dominant equilibrium. For 2×2 bimatrix coordination games, three methods have similar structures. For given c, each of them yields a curve separating $p - q$ plane and the equilibrium equilibrium is decided by the relative position of (p, q) and the curve, i.e., $(0, 0)$ is dominant if (p, q) is above the curve and $(1, 1)$ is dominant if (p, q) is below the curve.

For these curves, L_{RD} is a line with slope -1 but the shapes of L_{CD} and L_{BD} depend crucially on c. If $c = 1$, three curves are identical, which implies that the three methods choose the same Nash equilibrium. More precisely, the trajectory of the replicator dynamics starting at the centroid converges to the risk dominant equilibrium, and this equilibrium also has the largest basin of attraction. If $c \neq 1$, three methods are no longer equivalent. In this case, the risk dominant equilibrium is not always preferred.

Centroid dominant equilibrium can be calculated by Corollary 4.1, but there is no explicit formula for the basin dominant equilibrium. We only know that the curves of two methods have very similar properties. For instance, L_{BD} and L_{CD} are all monotonically decreasing in $p - q$ plane and the slopes at $(\frac{1}{2}, \frac{1}{2})$ are both $-\sqrt{c}$.

Instead of deriving the explicit expression, Theorem 4.4 provides an alternative way for finding the basin dominant strategy by comparing the curves of different methods. It claims that L_{BD} is always between L_{RD} and L_{CD}. Hence, if a strategy is both risk dominant and centroid dominant, it must has the largest basin of attraction. However, if the risk dominant equilibrium and centroid dominant equilibrium are different, numerical simulation suggests that the centroid dominant strategy is more likely to have the largest basin of attraction.

As an extension, centroid dominance and basin dominance are also well defined for 3×3 symmetric coordination games. In these games, Zeeman (1980; see also Hofbauer and Sigmund, 1998) showed that trajectories of the replicator dynamics converge to Nash equilibria, and all mixed Nash equilibria are unstable. However, we can not expect a simple formula to decide which equilibrium is selected even for the centroid dominance because the solutions of 3-strategy replicator dynamics do not always have explicit expressions. For further studies, a starting point is coordination games with diagonal payoff matrix. Intuitively, the payoff dominant strategy must be both centroid dominant and basin dominant. Another development is to compare the center and basin dominance equilibria under different evolutionary/learning dynamics. It is well known that a strict Nash equilibrium usually has different basins of attraction under the replicator dynamics and the best response dynamics. In particular, Golman and Page (2010) constructed a class of 3×3 symmetric games for which the overlap in the two basins of attraction is arbitrarily small. This implies that best response dynamics lead to a different equilibrium than replicator dynamics almost always.

As an extension, centroid dominance and basin dominance are also well defined for 3×3 symmetric coordination games. In these games, Zeeman (1980; see also Hofbauer

and Sigmund, 1998) showed that trajectories of the replicator dynamics converge to Nash equilibria, and all mixed Nash equilibria are unstable. However, we can not expect a simple formula to decide which equilibrium is selected even for the centroid dominance since the solutions of 3-strategy replicator dynamics do not always have explicit expressions. For further studies, a starting point is coordination games with diagonal payoff matrix. Intuitively, the payoff dominant strategy must be both centroid dominant and basin dominant. Another development is to study the relations between basin dominance and other equilibrium notions. In 3×3 symmetric games, a strict Nash equilibrium usually has different basins of attraction under the replicator dynamics and the best response dynamics. In particular, Golman and Page (2010) constructed a class of 3×3 symmetric games for which the overlap in the two basins of attraction is arbitrarily small.

Appendix

A.1 Proof of Corollary 4.3

Proof of Corollary 4.3 (b)

For any given $c > 0$, $(p,q) = (\frac{1}{2}, \frac{1}{2})$ *is on* L_{BD} *and the slope at this point is* $-\sqrt{c}$.

Proof

Eq.(4.4.9) shows that $(p,q) = (\frac{1}{2}, \frac{1}{2})$ is on L_{BD} for any $c > 0$. Let us now calculate the slope at this point.

For (p,q) close to $(\frac{1}{2}, \frac{1}{2})$ and (x,y) on $L_{(1/2,1/2,c)}$, we have

$$\begin{aligned} & s_0(p,q,c) - \frac{1}{2} \\ = & \int_0^1 (l_{(p,q,c)}(x) - l_{(1/2,1/2,c)}(x))dx \\ = & \int_0^1 ((p-\frac{1}{2})\frac{\partial l_{(p,q,c)}(x)}{\partial p}|_{(p,q)=(\frac{1}{2},\frac{1}{2})} + (q-\frac{1}{2})\frac{\partial l_{(p,q,c)}(x)}{\partial q}|_{(p,q)=(\frac{1}{2},\frac{1}{2})})dx \\ = & -\int_0^1 ((p-\frac{1}{2})\frac{\partial F_{(p,q,c)}/\partial p}{\partial F_{(p,q,c)}/\partial y}|_{(p,q)=(\frac{1}{2},\frac{1}{2})} + (q-\frac{1}{2})\frac{\partial F_{(p,q,c)}/\partial q}{\partial F_{(p,q,c)}/\partial y}|_{(p,q)=(\frac{1}{2},\frac{1}{2})})dx \\ = & -\int_0^1 \frac{2^{2c-2}(x(1-x))^c}{y-1/2}(c(p-\frac{1}{2})\ln(\frac{x}{1-x}) - (q-\frac{1}{2})\ln(\frac{y}{1-y}))dx \qquad \text{(A.1.1)} \end{aligned}$$

If (p,q) is on the tangent of L_{BD}, i.e.,

$$q - \frac{1}{2} = \frac{\partial f^*(p,c)}{\partial p}|_{(p,q)=(\frac{1}{2},\frac{1}{2})}(p-\frac{1}{2}) \qquad \text{(A.1.2)}$$

, Eq.(A.1.1) should not include first-order term of p.

For convenience, define $\ln(\frac{x}{1-x}) = e(x)$. The Taylor expansion of $y = l_{(1/2,1/2,c)}(x)$ at $e = 0$ is

$$y = \frac{1}{2} + \frac{\sqrt{c}}{4}e + o(e)^3 \qquad \text{(A.1.3)}$$

Hence,

$$\ln(\frac{y}{1-y}) = 4(y-\frac{1}{2}) + o(y-\frac{1}{2})^3 = \sqrt{c}e + o(e)^3 \tag{A.1.4}$$

Eq.(A.1.1) is then simplified as

$$-\int_0^1 2^{2c}(x(1-x))^c(\sqrt{c}(p-\frac{1}{2}) + q - \frac{1}{2} + o(e^2))dx \tag{A.1.5}$$

In order to eliminate the first-order term of $(p-\frac{1}{2})$, we need $\sqrt{c}(p-\frac{1}{2}) + q - \frac{1}{2} = 0$. This implies that the slope of L_{BD} at $(\frac{1}{2}, \frac{1}{2})$ is $-\sqrt{c}$. □

Proof of Corollary 4.3 (d)

For given $c \neq 1$, $(\frac{1}{2}, \frac{1}{2})$ is the only intersection of L_{BD} and L_{RD}.

Proof

Without loss of generality, suppose that $0 < c < 1$ and $p < \frac{1}{2}$. For $p + q = 1$, consider the derivatives of $L_{(1/2,1/2,c)}$

$$D(L_{(1/2,1/2,c)}) = c(x(1-x))^c(p-q)(1-x-y) \tag{A.1.6}$$

Notice that points on $L_{(1/2,1/2,c)}$ satisfy

$$\begin{aligned} 4(y-x)(1-x-y) &= (1-2x)^2 - (2y-1)^2 \\ &= (2^2(x(1-x)))^c - 2^2(x(1-x)) > 0 \end{aligned} \tag{A.1.7}$$

, this implies that $1-x-y > 0$ if $x < y$ and $1-x-y < 0$ if $y < x$. Thus, trajectories of Eq.(4.2.4) with initial points on $L_{(1/2,1/2,c)}$ are always attracted by $(0,0)$ since $D(L_{(1/2,1/2,c)}) < 0$ in region $(0, \frac{1}{2}) \times (\frac{1}{2}, 1)$ and $D(L_{(1/2,1/2,c)}) > 0$ in region $(\frac{1}{2}, 0) \times (0, \frac{1}{2})$. Thus, $s_0(p,q,c) > \frac{1}{2}$. From Theorem 4.2, L_{BD} is below L_{RD} for $0 < c < 1$ and $p < \frac{1}{2}$.

Similarly, L_{BD} is above L_{RD} for $0 < c < 1$ and $\frac{1}{2} < p$. Therefore, $(\frac{1}{2}, \frac{1}{2})$ is the only intersection of L_{BD} and L_{RD}. □

A.2 Proof of Theorem 4.4

Proof of Theorem 4.4

For any given c, L_{BD} is between L_{RD} and L_{CD}.

Proof

Without loss of generality, suppose $0 < c < 1$ and $p < \frac{1}{2}$. In this case, Corollary 4.2 indicates that L_{CD} is below L_{RD}. Hence, we have to prove: (i) L_{BD} is below L_{RD} and (ii) L_{BD} is above L_{CD}.

(i) can be obtained directly from property (d) of Corollary 4.3 (see also Appendix A.1) but the proof of (ii) is more complicated. Our basic idea is to calculate $s_0(p, q, c)$ for (p, q, c) on L_{CD} (denote by $s_0(L_{CD})$ for simplicity). From Theorem 4.2, L_{BD} is above L_{CD} if $s_0(L_{CD}) < \frac{1}{2}$.

The proof consists of three parts: (a) L_{CD} is below $L_{(1/2,1/2,c)}(p, q)$ if $0 < c < 1$ and $p < \frac{1}{2}$, where $L_{(1/2,1/2,c)}(p, q)$ denotes the separatrix $L_{(1/2,1/2,c)}$ on $p-q$ plane (i.e., replace x and y by p and q, respectively), (b) $L_{L_{CD}}$ is below $L_{(1/2,1/2,c)}$ if $0 < x < \frac{1}{2}$, where $L_{L_{CD}}$ denotes the separatrix $L_{(p,q,c)}$ with (p, q, c) on L_{CD}, and (c) $s_0(L_{CD}) < \frac{1}{2}$.

(a) From Corollary 4.2, $p < 1 - q < \frac{1}{2}$ for points on both curves. For convenience, denote the expressions of $L_{(1/2,1/2,c)}(p, q)$ and L_{CD} by $q_1 = l_{(1/2,1/2,c)}(p)$ and $q_2 = l_{CD}(p)$, respectively. Clearly, $p = q = \frac{1}{2}$ is an intersection.

We now investigate the existence of intersection in region $p < 1-q < \frac{1}{2}$. In this region, slopes of two curves are

$$\begin{aligned} \frac{dq_1}{dp} &= c\frac{q_1(1-q_1)(1-2p)}{p(1-p)(1-2q_1)} < 0 \\ \frac{dq_2}{dp} &= c\frac{\ln p - \ln(1-p)}{\ln q_2 - \ln(1-q_2)} < 0 \end{aligned} \tag{A.2.1}$$

, where at $p = q = \frac{1}{2}$, both of them equal to $\frac{1}{2}$.

For $p < 1 - q < \frac{1}{2}$,

$$\frac{q(1-q)(1-2p)}{p(1-p)(1-2q)} < \frac{\ln p - \ln(1-p)}{\ln q - \ln(1-q)} \tag{A.2.2}$$

since

$$\begin{aligned} &\frac{d(\frac{p(1-p)}{1-2p}(\ln p - \ln(1-p)))}{dp} \\ &= \frac{(\ln p - \ln(1-p))(p^2 + (1-p)^2) + 1 - 2p}{(1-2p)^2} \end{aligned} \tag{A.2.3}$$

and

$$\begin{aligned} &\frac{d((\ln p - \ln(1-p))(p^2 + (1-p)^2) + 1 - 2p)}{dp} \\ &= \frac{(1-2p)^2}{p(1-p)} - 2(1-2p)(\ln p - \ln(1-p)) > 0 \end{aligned} \tag{A.2.4}$$

From Eq.(A.2.2), we have $q_1 > q_2 > \frac{1}{2}$ for $p \to \frac{1}{2}^-$. To see this, consider the linear approximations of q_1 and q_2 near $(\frac{1}{2}, \frac{1}{2})$,

$$q_1 = \frac{1}{2} + (p - \frac{1}{2})\frac{dq_1}{dp}$$
$$q_2 = \frac{1}{2} + (p - \frac{1}{2})\frac{dq_2}{dp} \quad \text{(A.2.5)}$$

If $q_2 > q_1 > \frac{1}{2}$,

$$\frac{dq_1}{dp} = \frac{q_1(1-q_1)(1-2p)}{p(1-p)(1-2q_1)} < \frac{\ln p - \ln(1-p)}{\ln q_1 - \ln(1-q_1)} < \frac{\ln p - \ln(1-p)}{\ln q_2 - \ln(1-q_2)} = \frac{dq_2}{dp} \quad \text{(A.2.6)}$$

which contradicts Eq.(A.2.5).

On the other hand, Eq.(A.2.2) also implies that $\frac{dq_1}{dp} < \frac{dq_2}{dp}$ at all possible intersections. Therefore, $L_{(1/2,1/2,c)}(p,q)$ and L_{CD} can not meet each other in region $p < 1 - q < \frac{1}{2}$.

Notice that $l_{CD}(0) < 1 = l_{(1/2,1/2,c)}(0)$, L_{CD} is below $L_{(1/2,1/2,c)}(p,q)$ if $0 < c < 1$ and $p < \frac{1}{2}$.

(b) From part (a), we only need to show that $L_{L_{CD}}$ and $L_{(1/2,1/2,c)}$ have no intersection in region $0 < x < \frac{1}{2}$.

Suppose that (x, y) is an intersection, the slopes of two curves at (x, y) are then given by $k_{L_{L_{CD}}}(x,y) = \frac{x(1-x)(y-q)}{cy(1-y)(x-p)}$ and $k_{L_{(1/2,1/2,c)}}(x,y) = \frac{x(1-x)(y-1/2)}{cy(1-y)(x-1/2)}$, respectively.

Since the separatrix is decreasing, (x, y) is either in region $0 < x < p$ or $\frac{1}{2} < y < q$. If $0 < x < p$, we have $k_{L_{L_{CD}}}(x,y) < k_{L_{(1/2,1/2,c)}}(x,y) < 0$ since (p,q,c) is below $L_{(1/2,1/2,c)}$. Notice that boundary point $(0,1)$ is on both curves, two curves have no intersection in region $0 < x < p$. Similarly, if $\frac{1}{2} < y < q$, we have $0 > k_{L_{CD}}(x,y) > k_{L_{(1/2,1/2,c)}}(x,y)$. Notice that $(\frac{1}{2}, \frac{1}{2})$ is on both curves, they also have no intersection in region $\frac{1}{2} < y < q$. Therefore, $L_{L_{CD}}$ is below $L_{(1/2,1/2,c)}$ if $0 < x < \frac{1}{2}$.

(c) From part (b), we have $F_{(1/2,1/2,c)}(x,y) = 2^{2c}((1-x)x)^c - 2^2(1-y)y < 0$ for (x,y) on $L_{L_{CD}}$. Notice that (x,y) satisfies $F_{L_{CD}}(x,y) = 2^c(1-x)^{c(1-p)}x^{cp} - 2(1-y)^{1-q}y^q = 0$, this yields $F_{L_{CD}}(1-x, 1-y) = 2^c x^{c(1-p)}(1-x)^{cp} - 2y^{1-q}(1-y)^q < 0$, which implies that $l_{L_{CD}}(x) + l_{L_{CD}}(1-x) < 1$ if $x < \frac{1}{2}$.

Thus,

$$\begin{aligned} s_0(L_{CD}) &= \int_0^{1/2} l_{L_{CD}}(x)dx + \int_{1/2}^1 l_{L_{CD}}(x)dx \\ &= \int_0^{1/2} l_{L_{CD}}(x)dx + \int_0^{1/2} l_{L_{CD}}(1-x)dx < \frac{1}{2} \end{aligned} \quad \text{(A.2.7)}$$

Finally, from Theorem 4.2, L_{BD} is above L_{CD}. □

Chapter 5

The evolution of sanctioning institutions: An experimental approach to the social contract

Abstract

A vast amount of empirical and theoretical research on public good games indicates that the threat of punishment can curb free-riding in human groups engaged in joint enterprises. Since punishment is often costly, however, this raises an issue of second-order free-riding: indeed, the sanctioning system itself is a public good which can be exploited. Most investigations, so far, considered peer punishment: players could impose fines on those who exploited them, at a cost to themselves. Only a minority considered so-called pool punishment. In this scenario, players contribute to a punishment pool before engaging in the public good game, and without knowing who the free-riders will be. Theoretical investigations have shown that peer punishment is more efficient, but pool punishment more stable. Social learning should lead to pool punishment if sanctions are also imposed on second-order free-riders, but to peer punishment if they are not. Here we describe an economic experiment which tests this prediction. We find that pool punishment only emerges if second-order free riders are punished, but that peer punishment is more stable than expected. Basically, our experiment shows that social learning can lead to a spontaneously emerging social contract, based on a sanctioning institution to overcome the free rider problem.

Key words

Public goods game; experiments; collective action; punishment; institution; social learning

5.1 Introduction

Coercion plays an essential role in overcoming social dilemmas. The corresponding line of reasoning goes back at least as far as Hobbes' 'Leviathan' from 1651, and the practical implementation can be traced throughout history. The selfish motivations endangering collective actions have to be suppressed by positive and negative incentives (Olson 1965; Boyd and Richerson, 1992; Andreoni et al., 2003; Rockenbach and Milinski, 2006). In particular, the threat of punishment curbs the temptation to free-ride, i.e., to exploit the contributions of others without offering an adequate return.

Institutions can be viewed as tools for providing incentives (Ostrom, 2005). It has been shown that even in small-scale societies far removed from 'Leviathan'-like states, grass-root institutions can deal, often efficiently, with the tasks of monitoring joint efforts and sanctioning defectors (Ostrom, 1990; Henrich, 2006; Baldassarri and Grossman, 2011).

The role of punishment in boosting cooperation is one of the best studied topics in experimental economy. However, most investigations deal with so-called peer-punishment (see, e.g., Fehr and G?chter, 2000, 2002; Fehr and Rockenbach, 2003; Casari, 2005; Fowler, 2005; G?chter et al., 2008; Hermann et al., 2008; Henrich et al., 2006; Sigmund, 2007; Dreber at al., 2008; Egas and Riedl, 2008; Chaudhuri, 2011). Typically, the players in a public good game are allowed to impose fines on exploiters, at a cost to themselves. The threat of punishment can lead to considerable increases in the level of cooperation in the collective action. Many players are willing (and frequently even eager) to shoulder the costs of imposing fines on cheaters.

In most aspects of everyday life, the task of punishing exploiters has eventually been taken over by institutions (Ostrom, 2005; Guillen et al., 2006). This is remarkable, given the wide-spread tendency for moralistic aggression. In developed societies, peer-punishment is not only obsolete, but even explicitly forbidden. Under conditions of anarchy, individuals have to take punishment into their own hands, but in all better-regulated communities, punishment is delegated to institutions. How can we envisage this important step in social development?

Evidently, this question can be approached from many different angles. Here, we use an economic experiment to test how individuals who want to coerce their group to cooperate decide between inflicting punishment directly or using the intermediary of an institution. The foremost problem, in such an experiment, is how to implement the sanctioning institution (Tyler and Degoey, 1995; Casari and Luini, 2009; Kosfeld et al., 2009; Andreoni and Gee, 2011). Which is the essential feature distinguishing institutional from peer-punishment? Some argue that it is the delegation of punishment. However, individuals who want to exert personal revenge can recur to 'hiring a gun', and this would still count as peer-punishment (Van Vugt et al., 2009). A more pronounced difference is that sanctioning institutions are established in advance, and thus entail running costs even in the case that no one commits a punishable offense. A county would have to pay its sheriff even if nobody commits a crime. We tried to model this as 'pool-punishment' (Yamagishi, 1986; Sigmund et al, 2010, 2011; Kamei et al., 2011; Markussen et al., 2011; Traulsen et al., 2012). Players who want to use such a sanctioning tool have to pay a fee, even before the public good interaction takes place, or at least before they are informed of its outcome, and thus before they know whether there will be any exploiters to punish. Pool punishers can be viewed as paying a tax towards a police. We note that instead of pool- or peer-punishment, some authors use the terms 'formal' and 'informal' sanctions (Kamei et al., 2011; Markussen et al., 2011).

In our experiment, we investigated small groups, or 'toy-communities', of 12 to 14 players. Each such group played 50 rounds of a public good game. Within each group, players could decide, before each round, whether to join a public good game (A) without punishment, (B) with peer-punishment, (C) with pool-punishment or (D) not to participate. These games were played separately, i.e., the outcome of one game did not affect the outcomes of the other games in the group. Players were anonymous, and prevented from communicating. Usually, both features do not hold under realistic conditions, but we imposed them in order to focus on the alternative choices in punishment mechanisms. All that players learned, after each round, was how many opted, in their group, for each game, and which payoff they obtained. They then could choose whether to opt for (A), (B), (C) or (D) in the next round. We thus observed, in each toy community, whether social learning led to institutional punishment or not.

It is clear that if punishment works, i.e., if it leads to all-out cooperation, then peer-punishment is more efficient than pool-punishment, since it entails no running costs. However, theoretical considerations (Sigmund et al. 2010, see relevant theory in Appendix B.1) imply that pool-punishment is more stable, provided that it is also directed at those participants in the game who do not contribute to the punishment pool. Indeed, if cooperation is achieved, i.e., if no one needs to be punished, then a peer-punisher cannot be distinguished from a non-punisher. This means that second-order free-riders (defined as those who contribute to the public good, but not to the sanctions) cannot be spotted, and thus cannot be punished. By contrast, those who do not contribute to the punishment pool are just as visible as those who do not contribute to the public good, and can be punished just as well. A system implementing this is highly immune against exploitation, but requires payment of a tax.

In our experiment, a clear majority chose peer punishment in the first round. Most players switched to pool punishment in later rounds, but (as predicted) only if punishment was also imposed on second-order free-riders. The experiment involved 238 first-year students from universities in Vienna. Interactions were anonymous. Players were randomly allocated to 18 groups of 12 to 14 players each, for the duration of 50 rounds. We implemented 2 treatments with 9 groups each: in the 'second-order treatment', players were offered a pool punishment which sanctioned second-order free riders, and in the 'first-order treatment' a pool punishment game which did not. The former treatment led to the emergence of pool punishment in six out of the nine groups, the latter in none. Peer punishment slowly declined over rounds in both treatments. Roughly speaking, it was not displaced by pool punishment, but eroded gradually. Contributions to the public good were vastly more frequent in the treatment with second-order pool punishment. In a nutshell, players were allowed to 'vote with their feet' (the expression seems to be due to Tiebout, 1956), and they decided in favor of a sanctioning institution, but only if this institution coerced participants to contribute not merely to the public good, but also to its own upkeep. Under this additional commitment, the institution was adopted by the group, in a kind of 'social contract' which was achieved without explicit communication or deliberation, and uniquely based on social learning from the own experience and that of others.

In section 5.2, we describe the experiment, in section 5.3, we display the results, and in section 5.4, we offer a discussion and conclusions. The theoretical background, the instructions for the players and the detailed results of every group are contained in Appendix B.

5.2 The experiment

The 18 groups of 12 to 14 players (our 'toy-communities') were the independent sample points of our experiment. Players in different groups did neither interact nor communicate with each other for the duration of the experiment. The players were not told that the number of rounds was fixed beforehand at 50, so as to prevent end-round effects. In each round, players were given 3 monetary units (MU) and asked to choose one of three variants of public good (PG) games: (A) PG without punishment; (B) PG with peer punishment; (C) PG with pool punishment. The players could also decide (D) not to participate in any of these games. Such non-participants received an additional 0.5 MU. The idea, here, was that when not participating in a joint enterprise, an individual can engage in some useful activity which does not depend on the decisions of others. Once they had chosen one of these games, they interacted (through contributions and punishment) with those group-members who had chosen the same game. Players who opted for one of the games (A), (B) or (C), but found no co-players to join them, were treated as non-participants (D), and received an additional 0.5 MU, independently of what the others did. Once the round was over, the players learned how many (in their group) had played (A), (B), (C) or (D), which strategy they had chosen in their particular game, and which payoff they had obtained. (See payoff functions of different games in Appendix B.1). They could use this information to decide for which game to opt in the next round. Players did not learn about who did what, so there was no possibility to establish a reputation. Players knew that they would be paid immediately after the game, at a rate of 10 cents (euro) per MU, without having to give away their identity (as players) to their co-players or to the experimenters. The guaranteed minimal payoff was 10 euro.

Players participating in a PG game of type (A) (no punishment) could decide whether or not to contribute 1 MU to the common pool, knowing that their contribution would be multiplied by 3 and divided equally among all other players in their game, irrespective of whether these co-players had contributed or not. Thus contributors did not benefit from their own contribution. This slightly deviates from the ordinary type of PG games, where the contributors receive a return from their own contribution, usually a fraction inversely proportional to the number of participants in their PG game. Our version has the same structure as the Mutual Aid Game (Sugden 1986) and was also considered in Wilson (1975), Yamagishi (1986) and Fletcher and Zwick (2004). We adopted this version (which makes the social dilemma harder to solve) because the number of participants in the PG game can fluctuate in our experiment [1], which introduces a complicating factor which we wished to avoid (Sigmund, 2010). It may be noted that if all players contribute, everybody gains the same in both cases (namely 2 MU), irrespective of whether one obtains a return from the own contribution or not.

Players choosing to participate in a PG game of type (B) (peer punishment) would first play a PG game as described above, and then, in a second stage of the same round, be shown the number of non-contributors (i.e., defectors) in their game. Contributors could then decide whether or not to punish these free-riders. The fine-to-fee ratio is fixed to 2:1 in (B) [2]. Each punisher would have to pay a fee of 0.5 MU per defector, and that each defector would have to pay a fine of 1 MU per punisher. Again, if all cooperate, everyone gains 2 MU.

Players participating in a PG game of type (C) (pool punishment) had to choose between three options: (i) not to contribute anything, (ii) to contribute to the common pool (i.e., to pay 1 MU so that 3 MU would be shared among all other members who had chosen (C)), or (iii) to contribute to both the common pool and the punishment pool. This last alternative requires the players to pay 1 MU to the common pool and an additional 0.5 MU into the punishment pool. Thus if all cooperate, everyone gains 1.5 MU. This PG game was played in two variants, denoted as 'first-order variant' resp. 'second-order variant'. In the first-order variant, players knew that everyone who had not contributed to the common pool would be fined 1MU per punisher. In the second-order variant, players knew that everyone who had not contributed to both pools would be fined 1MU per punisher. Hence, in the second-order variant of game (C), second-order free-

[1] In our experiment, a PG game has at least 2 and at most 12 to 14 players. Isaac and Walker (1988) showed that the effect of group size on individual decisions is very weak if the self return remains constant.

[2] The effectiveness of fee-to-fine ratio 1:2 has been studied by many researchers (e.g., Carpenter, 2007; Nikiforakis and Normann, 2008). This ratio is enough to maintain cooperation.

riders were punished, while in the first-order variant, they were not. The fine to fee ratio can greatly vary, depending on the number of defectors and pool punishers. In groups 1-9 (with altogether 120 subjects), the game of type (C) was offered in the first-order variant, and in groups 10-18 (with 118 subjects) in the second-order variant.

We note that this is a complex game, without obvious money-maximizing strategies for the individuals choosing (B) and (C), since payoff depends on how many decide for the different options. In order to provide the players with an appreciation of the issues involved, they were given, at the start of the session, 25 practice rounds (see Appendix B.2). They knew that these rounds would not count towards their score and that groups would be reshuffled before the experiment started. More precisely, players were first given, via computer screen, a brief introduction into game (A) (no punishment), then played five rounds of game (A). The same then happened with games (B) (peer punishment) and (C) (pool punishment). Finally, they all played 10 rounds with the option, in each round, to choose between the three games (A), (B) and (C), or (D) to abstain from participation (exactly as later in the actual experiment). Thus players could familiarize themselves with their options, in the practice rounds, but were precluded from sharing their experiences. Immediately after the practice rounds, the 'toy communities' were assembled randomly, and engaged in their 50 rounds of social learning.

After each round, players were shown the payoffs for all strategies used in their group, and had 15 seconds to decide which game (A), (B). (C) or (D) to join next. The tightness of the schedule and the complexity of the task provided a strong motivation to be guided by the size of the payoffs, i.e., to engage in social learning. We also did not shrink from using loaded language in the instructions, for instance by calling punishment 'punishment'. Since our main aim was to compare different sanctioning technologies, we felt justified in acknowledging the underlying, common intention to uphold norms of collaboration. In particular, asocial punishment or revenge were not offered as options to our players. In a similar minimalistic spirit, we avoided the issue of increasing group returns.

5.3 Results

In the actual experiment, we observed strong changes in behavior in most of the 18 groups, especially during the initial phase. 12 of the groups eventually settled down, in the sense that the majority opted for the same game for each of the last 10 rounds. Six of these groups settled down for pool punishment. All six belonged to the second-order treatment. In three groups playing the second-order treatment, and three groups playing the first-order treatment, players settled for peer punishment. The null hypothesis that pool punishment is equally likely in both treatments can be rejected with a significance of $p < 0.05$ ($n_1 = 9$, $n_2 = 9$, two-sided binomial sample test). Based on the theoretical model, we had indeed expected pool punishment to emerge in the second-order treatment only.

The average frequency of pool punishment increased during the first rounds, in the second-order treatment, and overtook the frequency of peer punishment. In fact, the initial frequencies of (A), (B), (C) and (D), in the first-order treatment, corresponded closely with the initial frequencies in the second-order treatment, but then the frequencies evolved very differently (see Figure 5.3.1). Frequencies of peer punishers decreased in both treatments, but only slowly. Frequencies of pool punishment decreased in the first-order treatment, but increased in the second-order treatment. (See Table 5.3.1 (a))

More precisely, in the first round of the second-order treatment, 55 per cent of players choose peer punishment and 36 per cent pool punishment. The initial frequencies in the first-order version were 56 per cent and 31 per cent, respectively. However, in the first-order treatment, both frequencies declined, to reach 48 per cent and 19 per cent, respectively, by round 50. By contrast, the evolution in the second-order treatment reversed frequencies, so that after 50 rounds, 63 per cent of players opted for the pool punishment game but only 33 per cent for the peer punishment game (Figure 5.3.1 (b)). This reversal took place in the first 20 rounds. The regression line is $y = 0.326 + 0.0146x$ (where y represents the frequency of pool-punishment and x the round), with correlation coefficient $R = 0.9167$ and P-value< 0.0001. Obviously, players approached both first- and second-order treatments with similar expectations, but then underwent a very different learning experience. (See Table 5.3.1 (c))

Table 5.3.1: Regression lines

Table 5.3.1 (a): Popularity of different games

	Regression line (50 rounds)	R	P-value
First-order peer game	$y = 0.6347 - 0.0031x$	0.4761	P-value< 0.001
First-order pool game	$y = 0.2749 - 0.0029x$	0.4362	P-value< 0.001
Second-order peer game	$y = 0.5220 - 0.0044x$	0.6671	P-value< 0.001
second-order pool game	$y = 0.4325 + 0.0042x$	0.5939	P-value< 0.001

Table 5.3.1 (b): Frequencies of C and D

	Regression line (50 rounds)	R	P-value
First-order C	$y = 0.6689 - 0.0065x$	0.8812	P-value< 0.001
First-order D	$y = 0.3292 - 0.0015x$	0.1761	P-value$= 0.024$
Second-order C	$y = 0.8783 + 0.001x$	0.1779	P-value$= 0.023$
Second-order D	$y = 0.1164 - 0.001x$	0.1707	P-value$= 0.029$

Table 5.3.1 (c): Voting for different games in the second-order treatment

	Regression line (First 20 rounds)	R	P-value
Second-order peer game	$y = 0.6191 - 0.0136x$	0.8983	P-value< 0.001
Second-order pool game	$y = 0.3262 + 0.0146x$	0.9167	P-value< 0.001

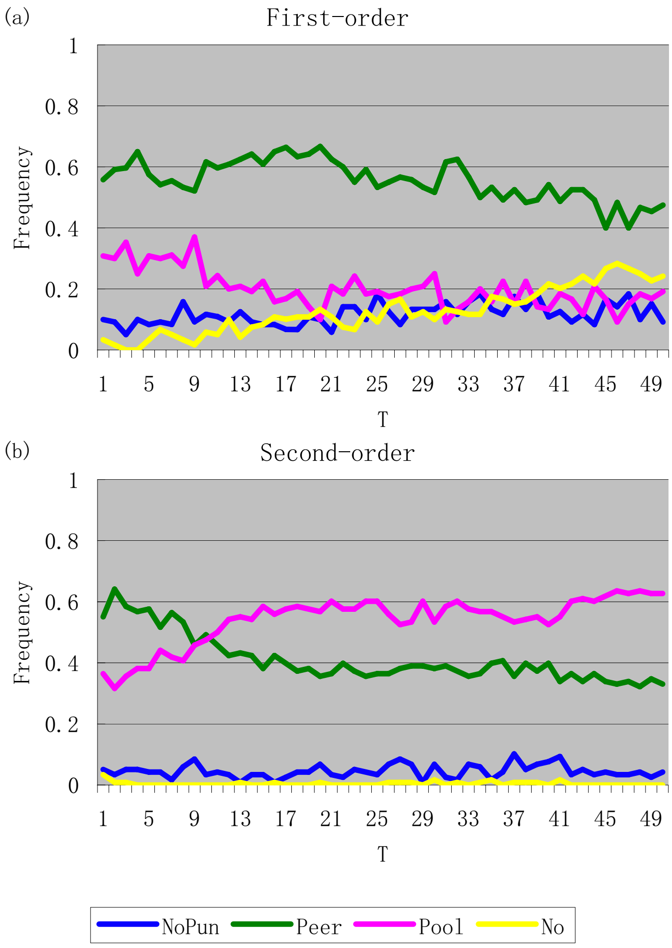


Figure 5.3.1: Time-evolution of the frequencies of players voting for the games (A), (B), (C) or (D). Here, (A) denotes the game without punishment (NoPun), (B) the game with peer-punishment (Peer), (C) with pool punishment (Pool) and (D) denotes non-participation (No).

Notes: y represents the frequency and x the round. R is the correlation coefficient.

If we average over all 50 rounds, we find a significant preference for peer punishment in the first-order treatment, and a less significant preference for pool punishment in the second- order treatment (Figure 5.3.3 (a)). The latter treatment leads to a very pronounced cooperative behaviour. Indeed, the frequency of contributions was significantly higher in the second-order treatment than in the first-order treatment (88.2 per cent vs.48.9 per cent, Mann-Whitney U-test, $n_1 = 9$, $n_2 = 9$, $p = 0.0373$), and it hardly changed over the 50 rounds (Figure 5.3.2 (b)). We can see (Figure 5.3.3 (c) and Appendix B.2) that average payoff values differ by little, but that peer punishment clearly yields the highest payoff in the first-order treatment, whereas it shares front rank with pool punishment, in the second order treatment.

In the first-order treatment, peer punishment was preferred by a wide margin: game (B) was chosen in 55.6 per cent of all decisions, game (C) in 20.2 per cent and game (A) in 11.7 per cent (Figure 5.3.3 (a)). A majority (62 per cent) of players opting for the peer punishment game contributed to the public good, but did not punish. Their payoff was higher than that of the punishers (4.636 vs. 4.1, Mann-Whitney U-test, $n_1 = 9$, $n_2 = 9$, $p = 0.077$). (It is obvious that within any round, this has to hold, if some players defect; we see here that it also holds on average). The non-contributors in the peer punishment game earned marginally more than the non-participants, namely 3.61 MU (the difference is not significant). All in all, 48.9 per cent of all decisions were in favour of contributing to the public good, rather than defecting (35.6 per cent) or abstaining from the game (15.5 per cent). But the time evolution over 50 rounds tells a more pessimistic story (Figure 5.3.2 (a)). Three-fourth of players cooperated in the first round but half of them gave up in later rounds. The regression line is $y = 0.669 - 0.0065x$ (where y represents the frequency of cooperation and x the round), with correlation coefficient $R = 0.8812$ and P-value < 0.0001. (Table 5.3.1 (b)) Moreover, in the first-order pool punishment games, cooperation did not take off. Only a tiny fraction of the decisions (54 out of 1149) favoured investing into the punishment pool.

In the second-order treatment, the preferences change drastically. Pool punishment, i.e., game (C), was chosen in 54.1 per cent of all decisions, and almost always (in 3155 of 3174 cases) was combined with a decision to actually contribute to the pool. Peer punishment (B) was chosen in 41 per cent of the decisions. Interestingly, players who chose the peer punishment game rarely decided to actually punish (only 9 per cent did), and the average payoff for those who engaged in peer punishment, 3.78 MU, was significantly less than that of second-order free-riders (4.77 MU, Mann-Whitney U-test, $n_1 = 9$, $n_2 = 9$, $p = 0.0106$). But this minority of punishers sufficed to keep free-riding down to 16 per cent. Few decisions (4.5 per cent) were in favour of the alternative (A), i.e., joining a PG game without punishment. The average payoff for the peer punishment game was

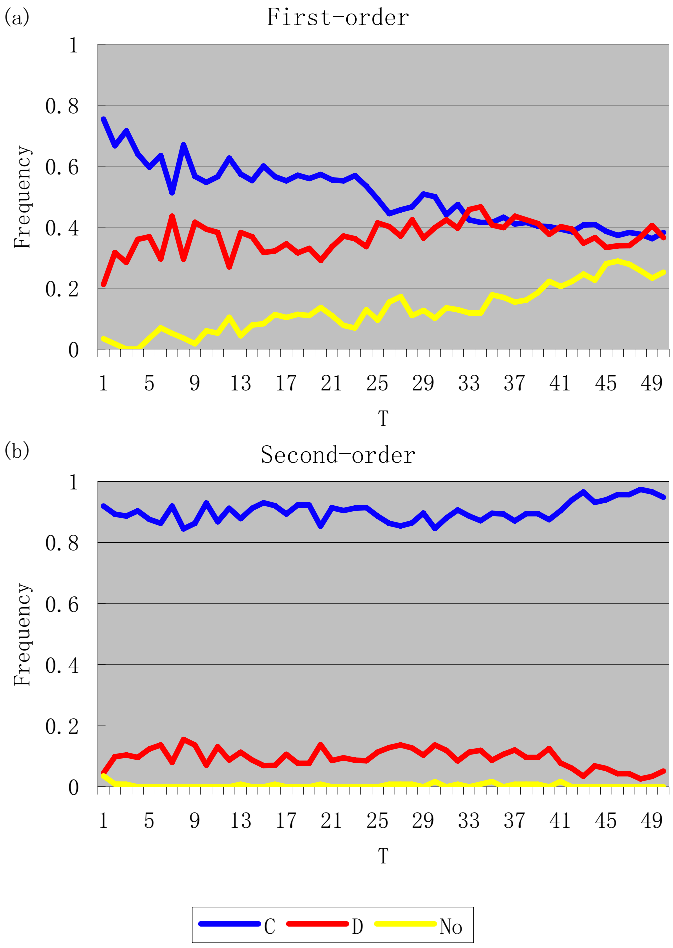


Figure 5.3.2: Time-evolution of the frequencies of cooperation (C, blue), defection (D, red) and non-participation (No, yellow) over 50 rounds in the first- and the second-order treatments. (a) In the first-order treatment, defection was chosen by about one-third of the players in each round. The number of contributions declined in favor of non-participation. (b) In the second-order treatment, almost all the players chose to contribute. This cooperative regime was stably sustained.

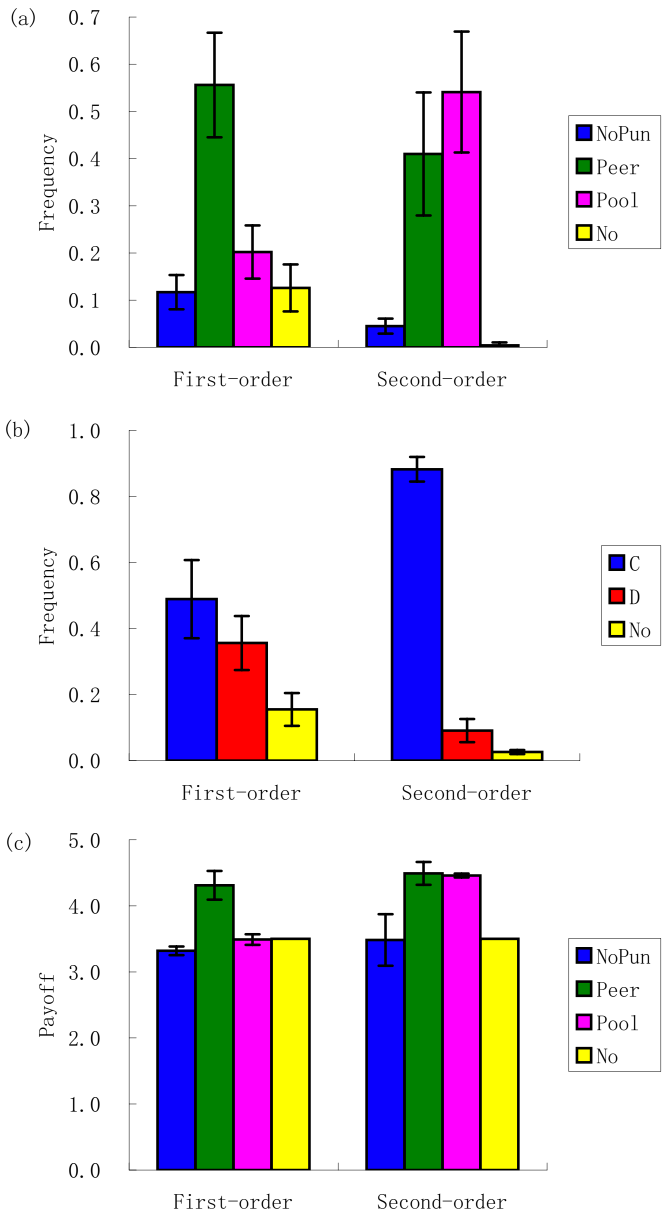


Figure 5.3.3: (a) Frequencies of the decisions in favor of the different games, over 50 rounds, for the first- and the second-order treatments. In the first-order treatment, peer punishment is favored. In the second-order treatment, pool punishment is more frequent, but error bars overlap. (b) Frequencies of the decisions to contribute to the public good, to defect (i.e., not to contribute) and to opt for non-participation, averaged over 50 rounds. Contribution is strongly promoted in the second-order treatment. (c) Payoffs obtained for the different games, averaged over fifty rounds, do not greatly differ. Nevertheless, in the first-order treatment, peer punishment games, and in the second-order treatment, both peer and pool punishment games provided the highest average payoff.

insignificantly larger than for the pool-punishment game (4.49 vs. 4.46), but those who actually peer-punished had a significantly lower payoff than the pool-punishers (3.78 vs. 4.49, Mann-Whitney U-test, $n_1 = 9$, $n_2 = 9$, $p = 0.004$).

The average payoff for those choosing a given game is almost the same for both treatments, with one exception: the payoff for choosing pool punishment has substantially increased in the second order treatment, because almost all players contributed to the common pool in the second-order treatment, but less than a third did so in the first-order version (Figure 5.3.3 (c)).

5.4 Discussion

In principle, a public good is non-excludable. In this sense, our PG game is misnamed, since players can decide not to participate. It may be better to call it a 'voluntary contribution game' or a 'collective-action game', for instance, or a 'mutual aid game', but we wanted to use the term most common in experimental games. There certainly exist enterprises or resources from which one cannot abstain: the global climate is the best example. Such compulsory interactions do not belong to the class considered here. Nevertheless, it could well be that the main 'efficiency vs. stability' result still holds for compulsory games. We decided to consider only the voluntary interactions in our experiment for two reasons: first, because the theoretical results guiding our predictions were derived for this class of games only, and second because, in the course of the experiment, we sometimes (but rarely) encounter players who do not make up their mind quickly enough, or who are the only individual choosing a given PG game of type (A), (B) or (C). In this case, it is practical to assign them option (D), namely 'non-participation'. This, incidentally, hardly affects the statistics.

The important role of second-order free-riding is well-known (Oliver, 1980), and our experiment confirms it. In the second-order treatment, pool punishment effectively prohibits this possibility, whereas in the first-order treatment, it does not. Apparently, pool-punishers notice that they are exploited, in the first-order treatment, and react against this breach in equity (Bolton and Ockenfels, 2000). Voting for the second-order treatment implies a higher commitment. In our experimental design, we did not allow for second-order peer punishment. The reason is twofold. On the one hand, theoretical models predict that it has no effect on the outcome (Sigmund et al., 2010). On the other hand, economic experiments have confirmed this in similar situations (Cinyabuguma et al., 2006; Kiyonari and Barclay, 2008; Traulsen et al., 2011).

We have reduced all individual decisions to choices between two, three or four alternatives. It would be interesting to investigate scenarios where players have a larger range of strategies, for instance by allowing them to choose between ten levels of contribution to the public good, or different degrees of punishment. Similarly, we have proposed only one, extremely rudimentary form of institution. It is easy to think of better designs, for instance by allowing part, at least, of the unused funds to return to the players who have contributed to the punishment pool. We refrained from doing this, because we did not want to make it too easy for institutional punishment to emerge. The fact that as many groups ended up with peer- as with pool-punishment suggests that we succeeded in this 'calibration'. Moreover, our experiment is already complex enough as it stands, and we feared to make it cognitively too demanding by adding more choices. As it was, the practice rounds needed to familiarize the players with their options took almost one hour (as long as the subsequent experiment).

Our main objective was to compare two different versions of pool punishment (rather than pool with peer). We note that there exist at least three experiments (independently conceived and as yet unpublished) comparing pool with peer punishment, or 'informal' with 'formal' sanctions (Kamei, Puttermann and Tyran, 2011; Markussen, Puttermann and Tyran, 2011; Traulsen, Rohl and Milinski, 2011). In Markussen et al. 2011, fixed groups of five players play for 24 rounds, and can vote, at specific instants, between two different regimes (corresponding, in our setup, to decisions between (A) and (B), (B) and (C), or (A) and (C)). In Kamei et al, their choice is between (B) and (C) with various parameters for the sanctions. Informal sanctioning does remarkably well. (The papers by Ertan, Page and Puttermann, 2009 and Boyd, Gintis and Bowles, 2010 confirm that peer punishment works well when players have an opportunity for coordinating.) Formal sanctions (which did not include second order punishment) fared poorly. The experiment by Traulsen, Rohl and Milinski 2012 presents players with the opportunity to use both mechanisms jointly, and finds that pool punishment prevails if it includes second-order punishment. In contrast to these papers, we describe how players 'vote with their feet' between competing games.

Our experiment is close in spirit and design to an experiment by Gurerk, Irlenbusch and Rockenbach, 2006. In this experiment, players were given the choice between a PG game with and one without peer punishment. The majority started with a clear preference for the treatment without punishment, but switched after a few rounds to the peer-punishment treatment, apparently guided by payoff considerations. Essentially, we kept the three-staged structure (choice of treatment, decision to contribute, decision to punish), but added pool punishment and non-participation as additional choices. (In contrast to the paper by Gurerk et al., 2006, we did not allow for rewarding; a related endogenous choice between peer punishing and rewarding has been investigated by Sutter, Haigner and Kocher, 2010.) The option of pool punishment adds an important element, as it essentially provides the opportunity for a tacit social contract establishing a sanctioning institution. To our knowledge, this is the first experiment demonstrating that such a social contract can emerge through social learning based on comparing the (frequency dependent) payoff values of diverse options.

The great attention that peer punishment has attracted in economic experiments is at least in part due to the fact that it does not presuppose the selection of an institution over another. Such a selection is necessarily culture-specific. Instead, peer-punishment scenarios mimic conditions of anarchy (i.e., the philosophers' 'state of nature'). It may be noted that nevertheless, institutions loom large in the background of such experiments: players are submitted to strict rules, and monitored by lab assistants who effectively act as authorities. Conditions of true anarchy, as would exist among the inmates of a prison or a kindergarten after the permanent removal of guards, can obviously not be implemented in economic experiments.

Since we wanted to favor conditions for social learning, we provided the players with information on the frequencies and average payoffs obtained by the various strategies in their group. However, we refrained from giving them opportunities to build up individual profiles, for instance reputations or significant differences in resources. Needless to say, this does not imply that reputations or differences in resource holding power are irrelevant for the evolution of institutions. Similarly, we did not consider other regarding preferences (Fehr and Schmidt, 1999) or contests between groups, although such struggles played doubtlessly an important role in human evolution (Choi and Bowles, 2007).

Our players were given the choice between one type of peer and one type of pool punishment. They could order them, as from a menu. Needless to say, such an approach cannot tell how such opportunities for sanctioning emerge, i.e., how the dishes were prepared. What are the roots of sanctioning institutions? Cooperation has frequently arisen through biological evolution (Maynard Smith and Szathmary, 1995), often via subtle mechanisms suppressing competition (Frank, 1995), and there exist many examples of animals punishing each other (Clutton-Brock and Parker, 1995). In particular, parents repress competition between their offspring, in many species, and it may be that this eventually led, in human populations, to institutionalized sanctioning. Offspring would simply have to remain with their parents (a costly option providing some safety) rather than leave and defend their interests single-handedly. It seems that institutions, once they have arisen, apply themselves to curb the vengeful and aggressive instincts fuelling peer-punishment. It would be interesting to explore this, both by modeling and by experiment. In our experimental setup, we have not allowed pool-punishers to sanction peer-punishers, or punished players to retaliate (Cinyabuguma et al., 2006; Nikiforakis, 2008). We also excluded communication and deliberation, although theoretical models, field observations and experiments have stressed the importance of communication in sanctioning exploiters (Walker et al., 2000; Bochet et al., 2006; Ertan et al., 2009). If individuals can look for allies, or deliberate with their peers, , stable systems of incentives can arise (Casari and Luini, 2009; Ertan et al., 2009; Boyd et al., 2010). We aimed for a minimalistic scenario based on social learning, and showed that it can lead to the emergence of a rudimentary type of institutionalized coercion helping to overcome individuals' selfish preferences.

Appendix B

B.1 Payoff values

In this section, we briefly sketch some of the relevant theory from Sigmund et al. (2010). First of all, let us consider the PG game of type (A) (no punishment). There are m players in the group. They can decide whether or not to contribute an amount $c > 0$, knowing that this will be multiplied by $r > 1$ and divided among all other players in the group. If m_C is the number of those players who contribute, and m_D the number of those who don't (with $m_C + m_D = m$), then the payoff for a contributor is

$$P^C = rc\frac{m_C - 1}{m - 1} - c \tag{B.1.1}$$

and that for a defector

$$P^D = rc\frac{m_C}{m - 1}. \tag{B.1.2}$$

Clearly, we always have $P^D > P^C$ (the difference is independent of m_C). If all players contribute, their payoff is $(r-1)c$, which is independent of group size m. The dominant strategy is to refuse to contribute. In our experiment, $c = 1$ MU, $r = 3$ MU and $m \geq 2$ is variable. Now let us consider the PG game of type (B) (peer punishment). Let us suppose that m_{Pe} the number of players who contribute and punish those who do not contribute, m_C the number of players who contribute, but do not punish, and m_D the number of those who neither contribute nor punish (with $m_{Pe} + m_C + m_D = m$). Let β be the size of the fine that each non-contributor receives from each punisher, and γ the fee each punisher has to pay for each non-contributor he or she punishes. Then we obtain as payoff values

$$\begin{aligned}
P^C &= rc\frac{m_C + m_{Pe} - 1}{m - 1} - c \\
P^{Pe} &= rc\frac{m_C + m_{Pe} - 1}{m - 1} - c - \gamma m_D \\
P^D &= rc\frac{m_C + m_{Pe}}{m - 1} - \beta m_{Pe}
\end{aligned} \tag{B.1.3}$$

There is no dominant strategy. The group optimum is obtained whenever $m_D = 0$. Clearly, we have $P^C \geq P^{Pe}$ (with equality if and only if $m_D = 0$). The state when no one contributes is a strict Nash equilibrium. Other (non-strict) equilibria exist for $m_D = 0$ and $m_{Pe} \geq \frac{c+\beta}{\beta}$. In our experiment, $\beta = 1$ MU and $\gamma = 0.5$MU so that states with two or more peer punishers, but no defector are also Nash equilibria.

Finally, let us consider games of type (C) (pool punishment). There are m_C players who contribute to the common pool, but not to the punishment pool, m_{Po} players who contribute to both pools, and m_D players who contribute to neither pool (with $m_{Po} + m_C + m_D = m$). Pool punishers have to contribute an amount c to the common pool and an amount F to the punishment pool. In the first-order variant, everyone who does not contribute to the common pool is fined by an amount Bm_{Po}, whereas in the second-order variant, everyone who does not contribute to both pools is fined by that amount. The payoff values are

$$P^{Po} = rc\frac{m_C + m_{Po} - 1}{m - 1} - c - F \tag{B.1.4}$$

and in the first-order variant (C1)

$$\begin{aligned} P^C &= rc\frac{m_C + m_{Po} - 1}{m - 1} - c \\ P^D &= rc\frac{m_C + m_{Po}}{m - 1} - Bm_{Po} \end{aligned} \tag{B.1.5}$$

resp. in the second-order variant (C2)

$$\begin{aligned} P^C &= rc\frac{m_C + m_{Po} - 1}{m - 1} - c - Bm_{Po} \\ P^D &= rc\frac{m_C + m_{Po}}{m - 1} - Bm_{Po} \end{aligned} \tag{B.1.6}$$

In our experiment, we used $B = 1$MU and $F = 0.5$MU. In the first-order variant, we have again $P^C > P^{Po}$ so that $m_D = m$ is the only equilibrium. In the second-order variant, $m_{Po} = m$ is another equilibrium (as long as $c + F \leq B(m - 1)$, which for our parameter values means that there are at least three punishers). We note that this equilibrium is not efficient, since $m_C = m$ provides a higher per capita payoff.

In Sigmund et al. (2010), it is shown that in the second-order version, pool punishment is more stable than peer punishment, although it is less efficient.

B.2 Experiments

The experiment took place in a computer lab of the Vienna University of Economics and Business (WU) on six days. On three days, the first-order treatment was played, and on the other three days the second-order treatment. The lab has 50 computers and for each of the six sessions, some 40 students (3 groups) played together. The interactions were anonymous, and via PCs. Cardboard dividers ensured that the students could not see each other. Players were not allowed to communicate, or to ask questions.

Table B.2.1: Group size in the first-order treatment and the second-order treatment

Group sizes in the first-order treatment									
group 1	group 2	group 3	group 4	group 5	group 6	group 7	group 8	group 9	Total
13	13	13	13	13	13	14	14	14	120
Group sizes in the second-order treatment									
group 1	group 2	group 3	group 4	group 5	group 6	group 7	group 8	group 9	Total
13	13	13	13	13	13	14	14	14	120

The practice rounds lasted about 45 min, almost for as long as the subsequent experiment (students knew that the sessions would at most for two hours, but were not told the number of rounds, so as to avoid end round effects). All players were given the same instructions (in German, see screen shots in Appendix B.4). The groups were then re-shuffled before the actual experiment started, and remained unchanged for its entire duration. The translation of the instructions for the practice rounds and the experiment can be found at the end of Appendix B.3. The average income was 19.6 euro (minimum 15.3, maximum 24.9). All steps were time-limited. Players knew that if they did not decide within 15 seconds, they would be allocated a random decision. Since the players had familiarized themselves with each game, this happened only 9 times in 11900 decisions, and is omitted from the statistics

In the groups 1-9, which offered the first-order treatment of pool punishment, peer punishment was preferred, as can be seen in Figure B.2.1 (a) and Table B.2.2 (a).

Table B.2.2 (a): Decisions in the first-order treatment

Groups 1-9: Popularity of the different games (including non-participation)

Decisions	Number of times	Percentage	Average payoff
(D) non-participation	754	0.126	3.500
(A) no-punishment game	701	0.117	3.342
(B) peer punishment game	3330	0.556	4.299
(C) pool punishment game	1208	0.202	3.492
Totals	5993	1	3.924

After including among non-participants those players who found no partners

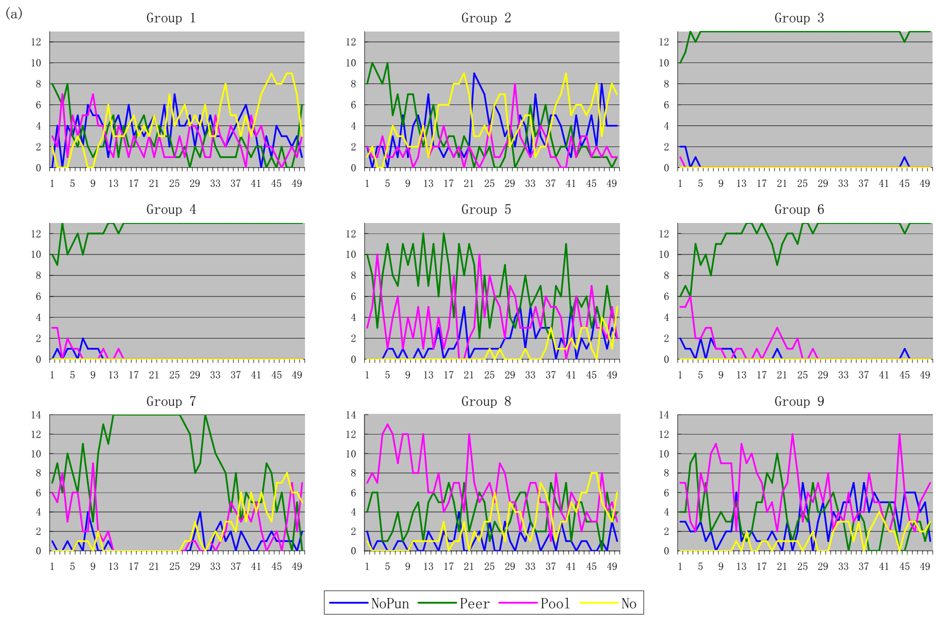
Decisions	Number of times	Percentage	Average payoff
(D) non-participation	926	0.155	3.500
(A) no-punishment game	618	0.103	3.32
(B) peer punishment game	3300	0.551	4.31
(C) pool punishment game	1149	0.192	3.49

Decisions within each game

Decisions	Number of times	Percentage	Average payoff
Contribution in (A)	99	0.017	2.601
Non-contribution in (A)	519	0.087	3.458
Contribution, but no punishing, in (B)	2049	0.342	4.636
Non-contribution in (B)	859	0.143	3.614
Peer-punishment and contribution in (B)	392	0.065	4.100
Contribution, but no punishing, in (C1)	338	0.056	3.486
Non-contribution in (C1)	757	0.126	3.566
Pool-punishment and contribution in (C1)	54	0.009	2.477

In the first-order pool punishment games, cooperation did not take off. Only a tiny fraction of the decisions in this group (54 out of 1149) favored investing into the punishment pool. The large majority seems to have sensed that the punishment threat would not be carried out, and defected. Defection was the most profitable decision in the pool punishment game, but the average payoff (3.566 MU) was only slightly higher than what non-participants obtained. (This difference was not significant). Peer punishment was clearly preferred. The average payoff obtained by opting for the peer-punishment game was 4.3 MU, higher than for opting for a pool-punishment game (3.49 MU, Mann-Whitney U-test, $n_1 = 9, n_2 = 9, p = 0.11$) or the game without punishment (3.34 MU, Mann-Whitney U-test, $n_1 = 9, n_2 = 9, p = 0.03$). Indeed, the average payoff values in the pool punishment or no-punishment games were lower than the non-participation payoff of 3.5 MU. A majority (62 percent) of players opting for the peer punishment game contributed to the public good, but did not punish. All in all, 48.9 percent of all decisions were in favor of contributing to the public good, rather than defecting (35.6 percent) or abstaining from the game (15.5 percent). But as mentioned in the main text, the time evolution over the fifty rounds shows a clear decline in contributions over time. We also note that free-riding was the most frequent and most successful behavior in the pool punishment game, but that the average payoff (3.566 MU) was only insignificantly higher than what non-participants obtained. Remarkably, the payoff for defecting in the games without punishment was almost the same (3.458 MU).

In the groups 10 to 18, pool-punishment was offered in the second-order treatment, i.e., it included punishing those who contributed to the common pool but not to the



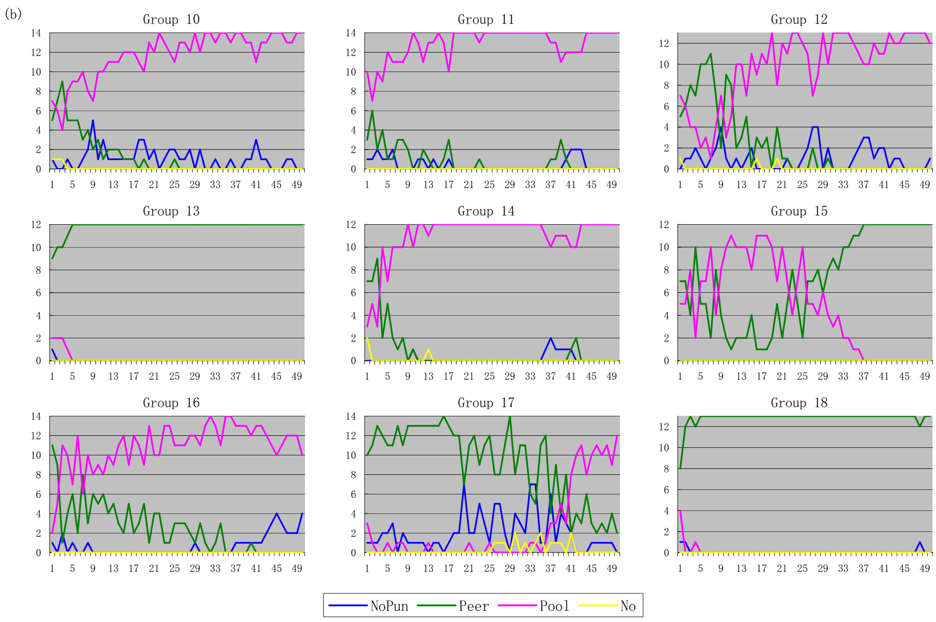


Figure B.2.1: The time-evolution, over fifty rounds, of the frequencies of players voting for the games (A), (B), (C) or (D). Game (A) is the game without punishment (NoP, blue), (B) the game with peer-punishment (Peer, green), (C) with pool punishment (Pool, pink) and (D) means non-participation (No, yellow). (a) The first-order treatment, groups 3, 4 and 6 settled on the peer punishment game, (in the sense that during each of the last 10 rounds, more than half of the players opted for it). The six other groups remained undecided. (b) The second-order treatment, groups 10, 11, 12, 14, 16, 17 settled on the pool punishment game, and groups 13, 15, 18 settled on the peer punishment game.

punishment pool. This time, pool punishment was preferred, as can be seen in Figure B.2.1 (b) and Table B.2.2 (b). (We note that in 49 out of 215 cases, declaring oneself to be peer punisher was cost-free, since there were no defectors to be punished.) Only 4.5 percent of all decisions were in favor of alternative (A). The free-riders, in that case, did about as poorly as in the peer punishment game (3.696 vs 3.689), since they found only few to exploit. Almost no decision was in favor of non-participation. In many more cases, non-participation was the unintended consequence of choosing a game that was not chosen by anyone else in the group. Second-order free-riding (i.e., opt for the peer punishment game, and contribute, but do not punish) achieved the highest payoff, 4.77 MU (see Figure 5.3.3 (c)).

The time-evolution in the different groups is interesting (see Figures B.2.1 and B.2.2). In seven of the nine groups where pool punishment was offered in the first-order treatment, the initial majority voted for peer punishment and in the other two groups, the initial majority voted for pool punishment. Three groups (3, 4 and 6) quickly reached consensus on peer punishment but all other groups went to chaos. During fifty rounds, players persisted in switching from one game to another. We note that in the three peer punishment groups, two-thirds of the players, in each round, decided not to opt for punishment. The threat of the remaining third sufficed to ensure co-operation, although that threat had rarely to be carried out.

Table B.2.2 (b): Decisions in the second-order treatment

Groups 10-18: Popularity of the different games (including non-participation)

Decisions	Number of times	Percentage	Average payoff
(D) non-participation	23	0.004	3.500
(A) no-punishment game	265	0.045	3.483
(B) peer punishment game	2421	0.410	4.490
(C) pool punishment game	3189	0.541	4.459
Totals	5898	1	4.424

After including among non-participants those players who found no partners

Decisions	Number of times	Percentage	Average payoff
(D) non-participation	154	0.026	3.500
(A) no-punishment game	181	0.031	3.475
(B) peer punishment game	2389	0.405	4.503
(C) pool punishment game	3174	0.538	4.464

Decisions within each game

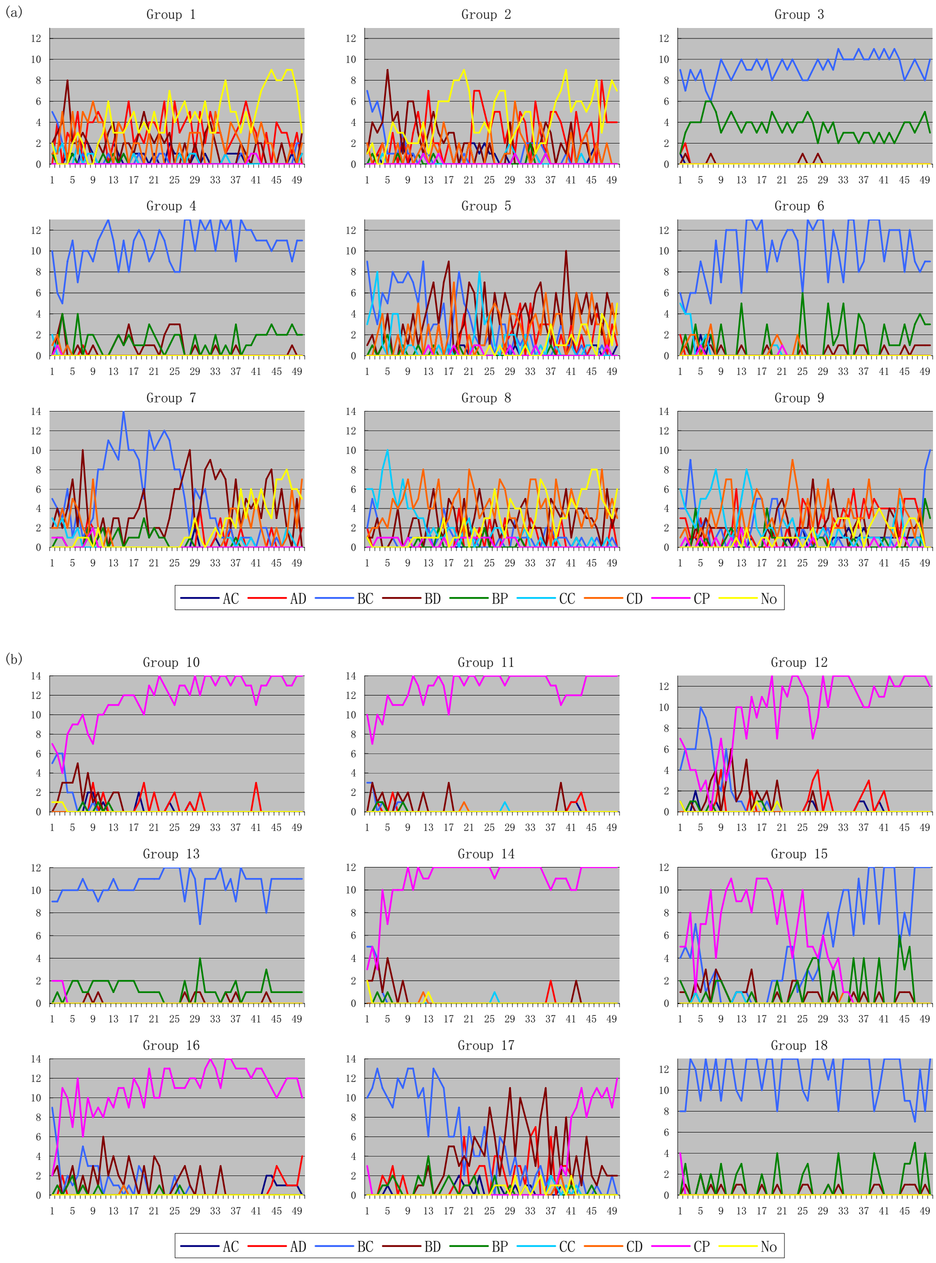


Figure B.2.2: The time-evolution, over fifty rounds, of the frequencies of the strategies. Here AC, AD, BC, BD, CC and CD denote contribution resp. defection in (A), (B) and (C), BP denotes peer-punishment, CP pool-punishment and No non-participation.

Decisions	Number of times	Percentage	Average payoff
Contribution in (A)	43	0.007	2.767
Non-contribution in (A)	138	0.023	3.696
Contribution, but no punishing, in (B)	1781	0.302	4.770
Non-contribution in (B)	393	0.067	3.689
Peer-punishment and contribution in (B)	215	0.036	3.776
Contribution, but no punishing, in (C2)	11	0.002	-0.955
Non-contribution in (C2)	8	0.001	0.313
Pool-punishment and contribution in (C2)	3155	0.535	4.493

There was not much switching in the groups where the second-order treatment of pool punishment was played. Despite the fact that in the first round, more players voted for peer than for pool punishment (65 vs 43), pool-punishment emerged in six of the nine groups as consensus solution. In three groups (13, 17 and 18), the initial majority for peer punishers was large enough to ensure the fixation of peer punishment within a few rounds. However, group 17 collapsed eventually, since the threat of peer punishment was not actually carried out. The players then turned to the pool treatment. A switch in the opposite direction occurred in group 15. After some initial oscillations, the pool-punishment game emerged as the majority choice, but it was never unanimous, and eventually became replaced by the peer-punishment treatment.

There are two related problems in establishing the statistics. One is that players opting for a game may end up with no partners, and thus become non-participants. Their decision was registered, and included in the statistics, but their payoff (3.5 MU) was not included in the average payoff for the game of their choice, since that game was cancelled. If we had added instead their 3.5 MU to the average, not much would have changed. The second problem is how to count the decisions in favor of peer punishment in those peer punishment games where no defection took place. If a player sees that there is no one to punish, and then chooses 'peer-punishment', this can indicate an earnest commitment to uphold the sanctioning system to guarantee cooperation (Masclet et al., 2003), but it could just as well be a mere cost-free gesture. If conversely a player chooses 'non-punishment', this can either indicate a decision for second-order free riding, or merely mean that the player was aware that there was no need for sanctions anyway. There were 108 such rounds (out of 900). In computing average payoffs and frequencies, we decided to take the players statements at face value. But we also computed a 'skeptical' version (not shown here), where players who actually did not punish were counted as non-punishers, no matter whether they declared themselves to be peer-punishers or not. Frequencies and the average payoffs are different, but the main conclusions remain unaffected.

The experiment was motivated by a theoretical analysis (Sigmund et al, 2010). This analysis predicts that the emergence of pool punishment is possible only if second-order free-riders are also punished. This is confirmed in our experiment. On the other hand, we expected that peer punishment would be replaced, in that case, by pool punishment. As it turned out, we did not observe this anticipated 'trading efficiency for stability'. Rather, we found examples for switches in both directions (groups 15 and 17, see Appendix B.2). A look at the time evolution in each group (see Appendix B.2, Figures B.2.1 and B.2.2) suggests that in both treatments, peer punishment offered a modicum of stability, but that when it failed, it gave way to asocial behavior (i.e., non-participation or defection) in the first-order treatment, and to pool punishment in the second-order treatment. As a consequence, contributions were stably sustained in the second-order treatment, at a very high level, whereas they declined, and were ultimately overtaken by defections, in the first-order treatment (see Figure 5.3.2). This good performance of peer punishment may be due to the fact that retaliatory punishment was not possible in our design (Cinyabuguma et al., 2006; Nikiforakis, 2008). Moreover, in contrast to the theoretical model (Sigmund et al., 2010), pool-punishers could not punish peer-punishers in our experiment. They belonged to different games. It is possible that 'cross-punishment' can change this outcome (Traulsen et al., 2011).

The initial phase of our experiment displayed a high rate of change in behavior in most groups. On average, more than one-fourth of the players switched to another decision between one round and the next, during the first twenty rounds. In the last ten rounds, the average switching rate was only 5.6 percent in the twelve groups that had settled on peer or pool punishment, but 50 percent in the others.

Another question that was not addressed here is whether the option to abstain from the game ('non-participation'), which is crucial for the theoretical analysis (Sigmund et al., 2010), is also necessary for the experiment. For the analysis, it was assumed that innovative behavior ('mutation') is much rarer than copying behavior. In that case, non-participation is necessary as an escape from the homogeneous state of defection. Since actual human populations display high degrees of polymorphism (Traulsen et al., 2010), non-participation may not be needed. On the other hand, voluntary participation is likely to increase the perceived legitimacy of the sanctioning institution, and hence its efficiency (Tyler and Degoey, 1995; Ertan et al., 2009).

B.3 Instructions

B.3.1 Instructions for the practice rounds (translated into English).

Welcome and thank you for showing up. Your minimal payoff will be 10 euros (guaranteed). We first start with some practice games. These do not count towards your score. You can experiment.

COMMUNITY GAME

In each round, you receive 3 MU and must decide whether or not to contribute 1 MU to your co-players' payoff.

I CONTRIBUTE means: you pay 1 MU and 3 MU will be distributed equally among all your co-players.

I DON'T CONTRIBUTE means: you keep 1 MU. This will not change your co-player's score.

You have 30 seconds for each round to decide and CONFIRM. If you do not decide in time, the computer will make a random decision. After each round, you will see the scores.

EXAMPLE

If all contribute, all end up with 5 MU. If no one contributes, all end up with 3 MU. In mixed groups, contributors always end up with less than the non-contributors.

DO YOU WANT TO CONTRIBUTE TO YOUR GROUP?

YES

NO

The round is played.

The scores are displayed.

This is repeated 5 times, with a reflection time of 30 seconds per round.

COMMUNITY GAME WITH OPTION TO PUNISH

This game consists of 2 stages. At the start of each round you receive 3 units. The first stage is the community game, as above. You can decide whether or not to contribute 1 unit. You will then see the scores in your group, and how many contributed. In the second stage, contributors can decide whether or not to punish all those who did not contribute. If you punish, you have to pay 0.5 MU per non-contributor. Each non-contributor is then fined 1 MU. You will then see the final score of the round.

EXAMPLE

If 4 players punish a non-contributor, this costs each punisher 0.5 MU, and the punished player 4 MU.

If 3 players punish 2 non-contributors, this costs each punisher 1 MU and each punished player 3 MU.

If 2 players punish 3 non-contributors, this costs each punisher 1.5 MU and each punished player 2 MU.

DO YOU WANT TO CONTRIBUTE TO YOUR GROUP?

YES

NO

x players out of y contributed.

DO YOU WANT TO PUNISH ALL NON-CONTRIBUTORS?

YES

NO

The round is played.

The scores are displayed.

This is repeated 5 times, with a reflection time of 30 seconds for each decision.

COMMUNITY GAME WITH PUNISHMENT DEVICE

At the start of each round you receive 3 MU. Again, you can decide to contribute 1 MU to the group or not. Contributors can additionally decide to pay for a punishment device. This costs the contributor 0.5 MU.

In the first-order treatment: Each punishment device will punish all non-contributors by 1 MU.

In the second-order treatment: Each punishment device will punish all non-punishers by 1 MU (irrespective of whether they contributed or not).

EXAMPLE FOR THE SECOND ORDER TREATMENT

If 3 players chose a punishment mechanism, each pays 0.5 MU and 3 MU will be removed from the account of each player who did not chose the punishment mechanism. Even if every player choses the punishment mechanism and no-one will be punished, the costs for the punishment mechanism will have to be paid.

DO YOU WANT TO CONTRIBUTE TO THE GROUP? DO YOU WANT A PUNISHMENT DEVICE?

JUST CONTRIBUTE TO THE GROUP

NEITHER, NOR

BOTH

The round is played.

The scores are displayed.

This is repeated 5 times, with 30 seconds per decision.

B.3.2 Instructions for the full game with option to choose a game (still in the practice rounds)

You will now have to decide, for each round, which game to play. You will receive 3 units for each round. You can choose to join

A: COMMUNITY GAME WITH NO PUNISHMENT
B: COMMUNITY GAME WITH OPTION TO PUNISH
C: COMMUNITY GAME WITH PUNISHMENT DEVICE

You can also decide not to play the game. In this case, you receive an additional 0.5 MU, but you cannot improve.

13 players participate in each round. But the sizes of the groups playing A, B or C are variable. If no co-player joins your group, you receive 0.5 MU and your game is cancelled. At the end of each round, you will see the scores.

OPT FOR YOUR GAME:
A: COMMUNITY GAME WITH NO PUNISHMENT
B: COMMUNITY GAME WITH OPTION TO PUNISH
C: COMMUNITY GAME WITH PUNISHMENT DEVICE
D: NO GAME

The round is played.

The scores are displayed.

This is repeated 10 times, with 30 seconds per decision.

B.3.3 Instructions for the experiment (after the practice rounds)

Now you will be paid according to your score (1 MU is 10 cents, so that 10 MU = 1 euro). The average payoff will be around 20 euros.

OPT FOR YOUR GAME:
A: COMMUNITY GAME WITH NO PUNISHMENT
B: COMMUNITY GAME WITH OPTION TO PUNISH
C: COMMUNITY GAME WITH PUNISHMENT DEVICE
D: NO GAME

The round is played.

The scores are displayed.

Repeat this for 50 rounds, with 15 seconds per decision

B.4 Screen shots

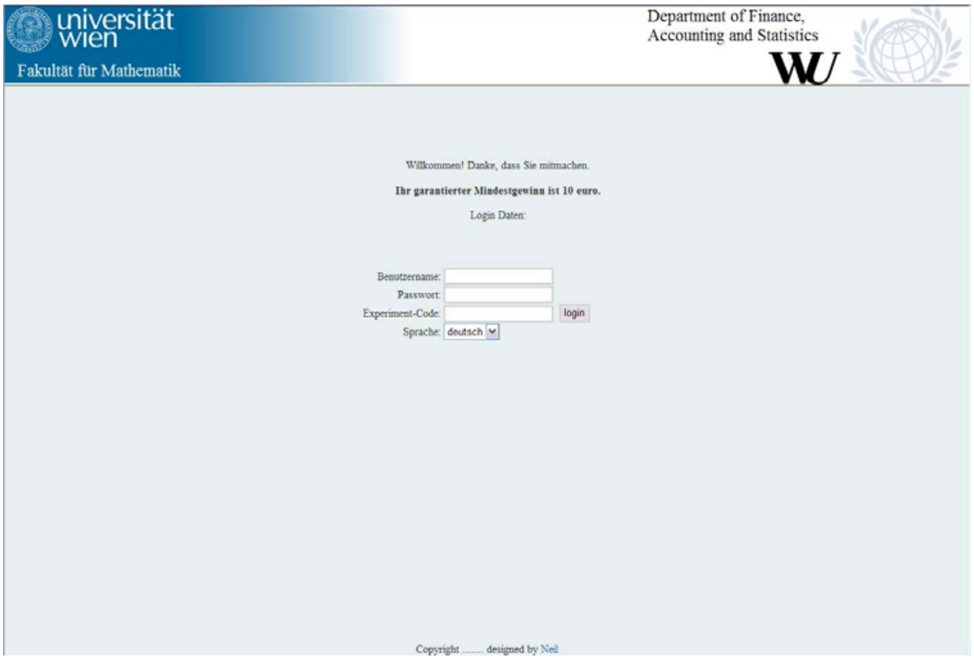


Figure B.4.1: Login page

Figure B.4.2: Practice rounds, instruction, game (A)

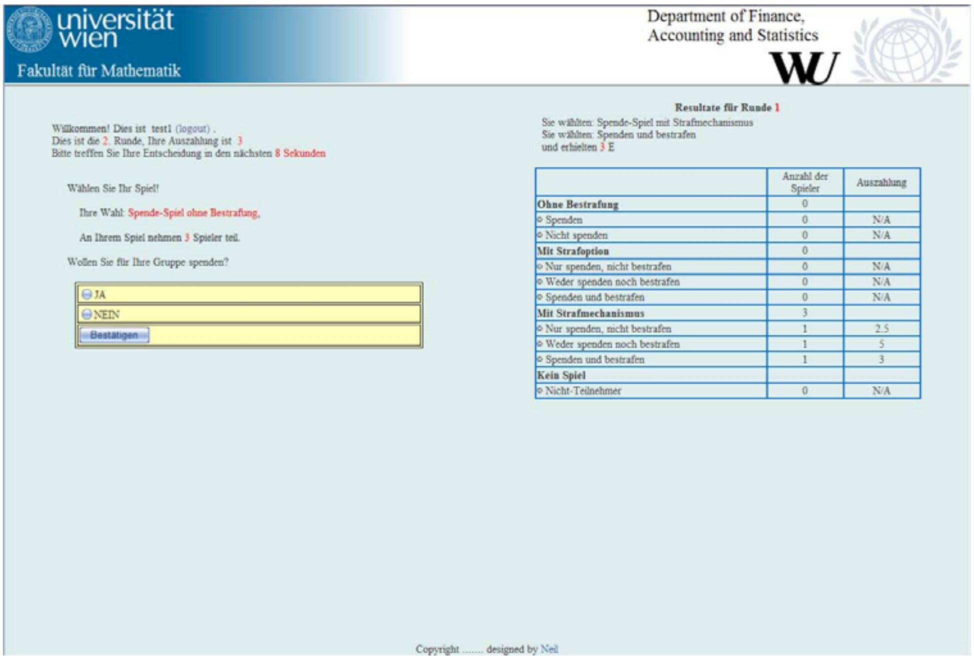


Figure B.4.3: Practice rounds, game (A)

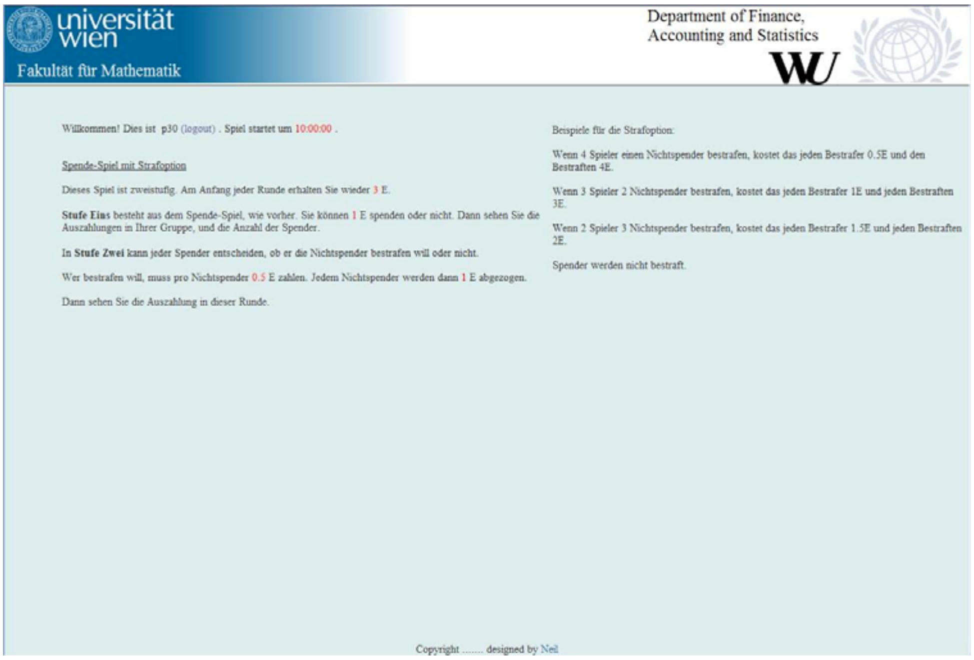


Figure B.4.4: Practice rounds, instruction, game (B)

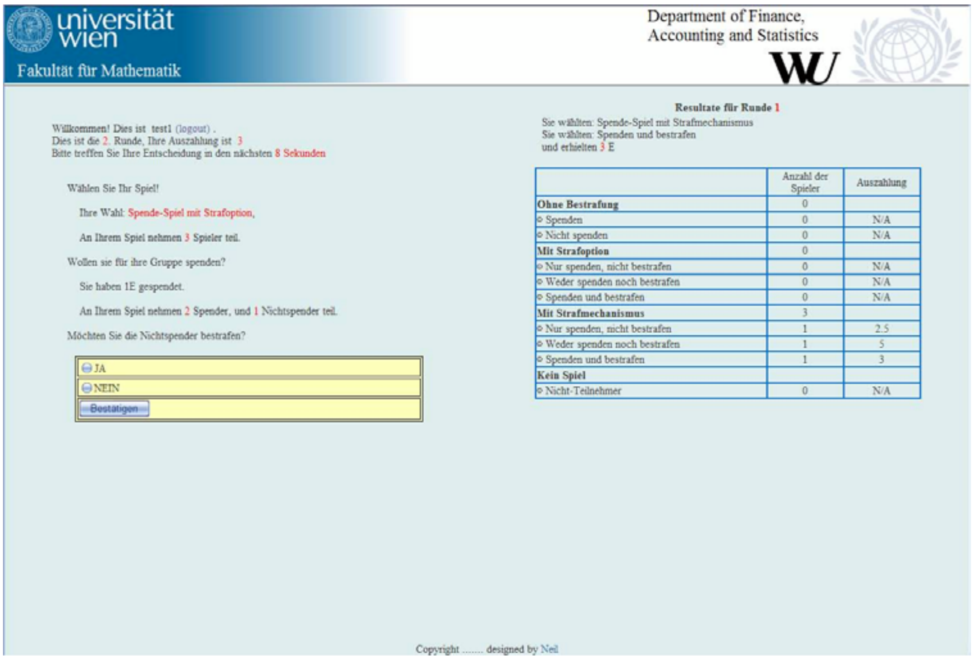


Figure B.4.5: Practice rounds, game (B)

Figure B.4.6: Practice rounds, instruction, game (C), first-order variant

Figure B.4.7: Practice rounds, instruction, game (C), second-order variant

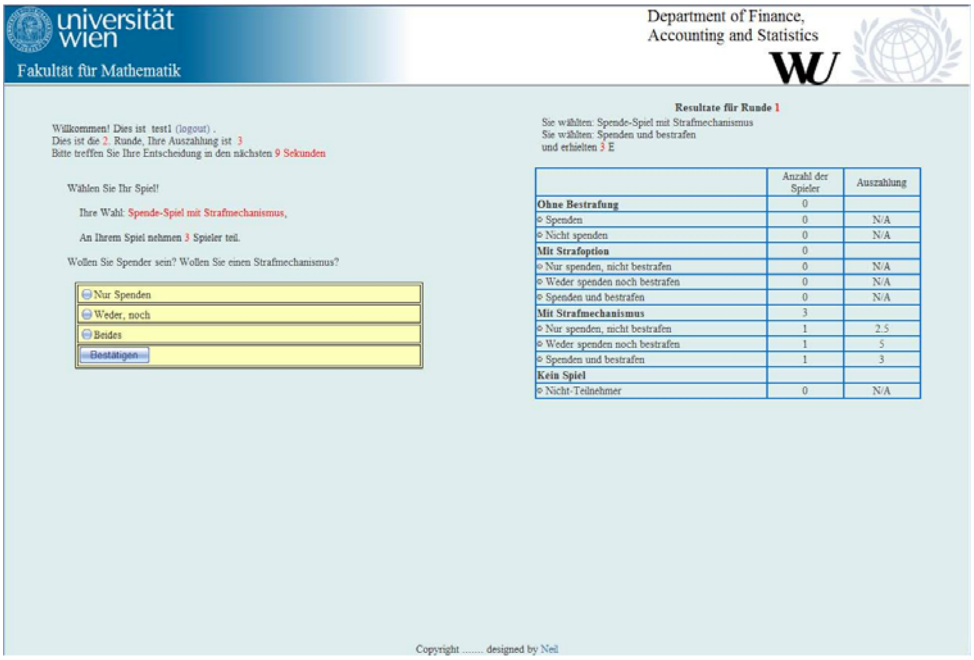


Figure B.4.8: Practice rounds, game (C)

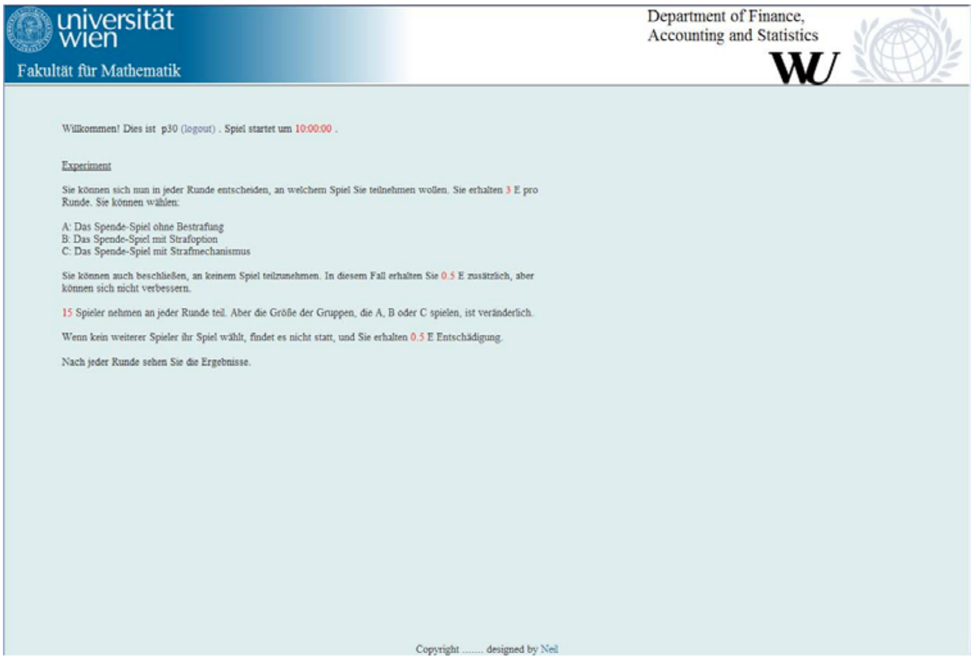


Figure B.4.9: Practice rounds, instruction, full game with option to choose a game

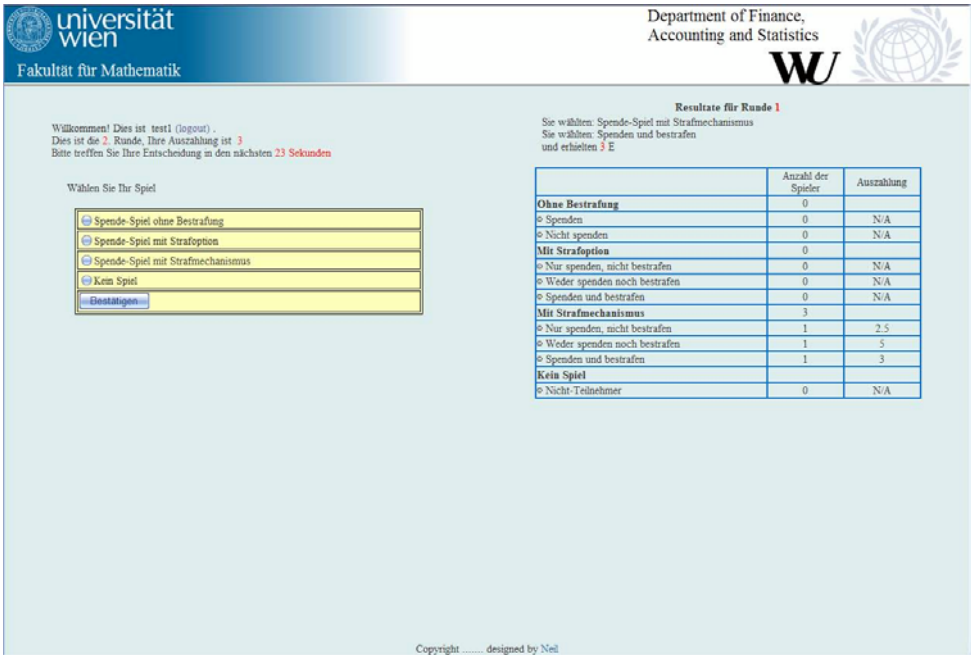
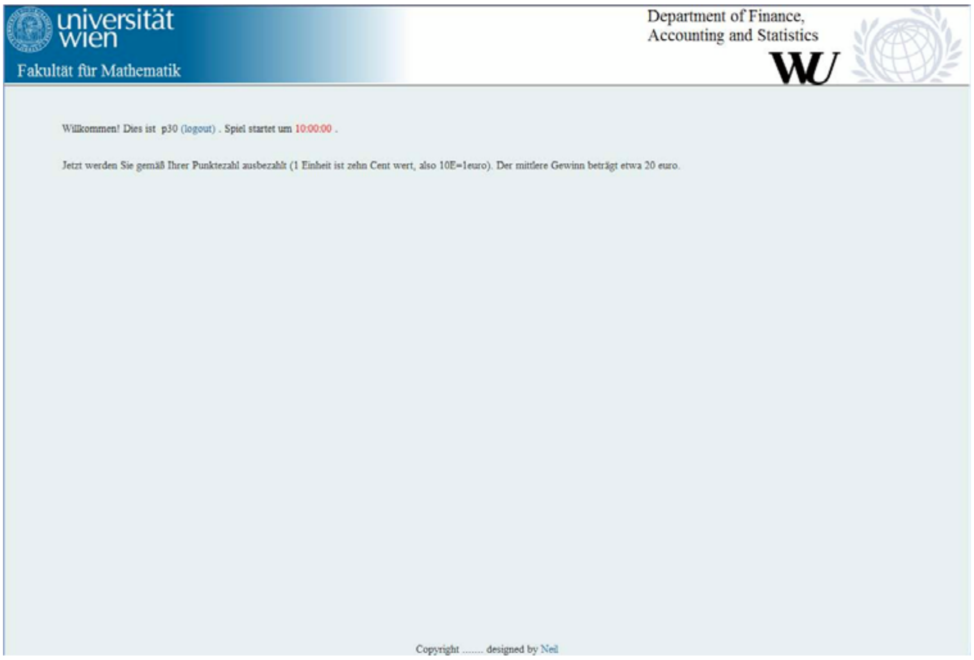


Figure B.4.10: Practice rounds, full game with option to choose a game


universität wien
Fakultät für Mathematik
Department of Finance, Accounting and Statistics
WU

Willkommen! Dies ist p30 (logout) . Spiel startet um 10:00:00 .

Jetzt werden Sie gemäß Ihrer Punktezahl ausbezahlt (1 Einheit ist zehn Cent wert, also 10E=1euro). Der mittlere Gewinn beträgt etwa 20 euro.

Copyright .... designed by Neil

Figure B.4.11: Experiment, instruction


universität wien
Fakultät für Mathematik
Department of Finance, Accounting and Statistics
WU

Willkommen! Dies ist p30 (logout) .

Spiel beendet, Ihre Auszahlung beträgt:227.87

	Anzahl der Teilnehmer	Mittlerer Gewinn
Ohne Bestrafung	0	N/A
○ Spenden	0	N/A
○ Nicht spenden	0	N/A
Mit Strafoption	356	4.51
○ Nur spenden, nicht bestrafen	265	4.81
○ Weder-noch	35	3.48
○ Spenden und bestrafen	56	3.74
Mit Strafmechanismus	238	4.41
○ Nur spenden, nicht bestrafen	3	-1.33
○ Weder-noch	2	5
○ Spenden und bestrafen	233	4.47
Kein Spiel		
○ Nicht-Teilnehmer	0	N/A

Sie haben 22.8 Euro gewonnen

Resultate für Runde 50

Sie wählten: Spende-Spiel mit Strafoption
Sie wählten: Nur spenden, nicht bestrafen
und erhielten 5 E

	Anzahl der Spieler	Auszahlung
Ohne Bestrafung	0	
○ Spenden	0	N/A
○ Nicht spenden	0	N/A
Mit Strafoption	12	
○ Nur spenden, nicht bestrafen	12	5
○ Weder spenden noch bestrafen	0	N/A
○ Spenden und bestrafen	0	N/A
Mit Strafmechanismus	0	
○ Nur spenden, nicht bestrafen	0	N/A
○ Weder spenden noch bestrafen	0	N/A
○ Spenden und bestrafen	0	N/A
Kein Spiel		
○ Nicht-Teilnehmer	0	N/A

Copyright .... designed by Neil

Figure B.4.12: Experiment, resulting page

Bibliography

[1] Abbink, K., 2001. Adaptive learning versus punishment in ultimatum bargaining. Games and Economic Behavior, 37: 1-25.

[2] Anderson, S. P., Goeree, J. K., and Holt, C. A., 1998. Rent seeking with bounded rationality: an analysis of the all-pay auction. Journal of Political Economy, 106: 828-853.

[3] Anderson, S. P., Goeree, J. K., and Holt, C. A., 2001. Minimum effort coordination games: stochastic potential and the logit equilibrium. Games and Economic Behavior, 34: 177-199.

[4] Anderson, S. P., Goeree, J. K., and Holt, C. A., 2004. Noisy directional learning and the logit equilibrium. The Scandinavian Journal of Economics, 106: 581-602.

[5] Andreoni, J. and Gee, L. L., 2011. Gun for hire: does delegated enforcement crowd out peer punishment in giving to public goods? NBER Working Paper, No. 17033. http://www.nber.org/papers/w17033. pdf

[6] Andreoni, J., Harbaugh, W. and Vesterlund, L., 2003. The carrot or the stick: rewards, punishments, and cooperation. American Economic Review, 93: 893-902.

[7] Baldassarri, D. and Grossman, G., 2011. Centralized sanctioning and legitimate authority promote cooperation in humans. Proceedings of the National Academy of Sciences, 108: 11023-11027.

[8] Berger, U., 2002. Best response dynamics for role games. International Journal of Game Theory, 4: 527-538.

[9] Binmore, K. and Samuelson, L., 1997. Muddling through: noisy equilibrium selection. Journal of Economic Theory, 74: 235-265.

[10] Blume, L. E., 1993. The statistical mechanics of strategic interaction. Games and Economic Behavior, 5: 387-424.

[11] Blume, L. E., 1995. The statistical mechanics of best-response strategy revision. Games and Economic Behavior, 11: 111-145.

[12] Bochet, O., Page, T. and Putterman, L., 2006. Communication and punishment in voluntary contribution experiments. Journal of Economic Behavior and Organization, 60: 11-26.

[13] Bolton, G.E. and Ockenfels, A., 2000. ERC: a theory of equity, reciprocity, and competition. The American Economic Review, 90: 166-193.

[14] Borgers, T. and Sarin, R., 1997. Learning through reinforcement and replicator dynamics. Journal of Economic Theory, 77: 1-14.

[15] Boyd, R., Gintis, H. and Bowles, S., 2010. Coordinated punishment of defectors sustains cooperation and can proliferate when rare. Science, 328: 617-620.

[16] Boyd, R., and Richerson, P. J., 1992. Punishment allows the evolution of cooperation (or anything Else) in sizeable groups. Ethnology and Sociobiology, 113: 171-195.

[17] Breitmoser, Y., Tan, J. and Zizzo, D., 2010. On the Beliefs off the Path: Equilibrium Refinement due to Quantal Response and Level-k. Mimeo.

[18] Brenner, T. and Vriend, N.J., 2006. On the behavior of proposers in ultimatum games. Journal of Economic Behavior and Organization, 61: 617-631.

[19] Cabrales, A., 2000. Stochastic replicator dynamics. International Economic Review, 41: 451-481.

[20] Camerer, C., 2003. Behavioral Game Theory: Experiments in Strategic Interaction, Princeton University Press, Princeton.

[21] Capra, C. M., Goeree, J. K., Gomez, R. and Holt, C. A., 1999. Anomalous behavior in a travelers dilemma? American Economic Review, 89: 678-690.

[22] Carlsson, H. and Damme, E. v., 1993. Global games and equilibrium selection. Econometrica, 61: 989-1018.

[23] Carpenter, J. P., 2007. The demand for punishment. Journal of Economic Behavior and Organization, 62: 522-542.

[24] Casari, M., 2007. On the design of peer punishment experiments. Experimental Economics, 8: 107-115.

[25] Casari, M. and Luini, L., 2009. Cooperation under alternative punishment institutions: an experiment. Journal of Economic Behavior and Organization, 71: 273-282.

[26] Charness, G. and Rabin, M., 2002. Understanding social preferences with simple tests. Quarterly Journal of Economics, 117: 817-869.

[27] Chaudhuri, A., 2011. Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. Experimental Economics, 14: 47-83.

[28] Choi, J-K. and Bowles, S., 2007. The coevolution of parochial altruism and war. Science, 318: 636-640.

[29] Choi, S., Gale, D. and Kariv, S., 2009. Social learning in networks: a quantal response equilibrium analysis of experimental data. Mimeo.

[30] Cinyabuguma, M., Page, T. and Putterman, L., 2006. Can second-order punishment deter perverse punishment. Experimental Economics, 9: 265-279.

[31] Clutton-Brock, T. H. and Parker, G. A., 1995. Punishment in animal societies. Nature, 373: 209-216.

[32] Cooper, R.W., DeJong, D.V., Forsythe, R. and Ross, T.W., 1990. Selection criteria in coordination games: Some experimental results. The American Economic Review, 80: 218-233.

[33] Cooper, D.J. and Dutcher, E.G., 2010. The dynamics of responder behavior in ultimatum games: a meta-study. Experimental Economics, 14: 519-546.

[34] Dasgupta, P. and Maskin, K., 1986. The existence of equilibrium in discontinuous economic games, I: Theory. The Review of Economic Studies, 53 :1-26.

[35] De Quervain, D. J-F., Fischbacher, U., Treyer, T., Schellhammer, M., Schnyder, U., Buck, A. and Fehr, E., 2004. The neural basis of altruistic punishment. Science 305: 1254-1258.

[36] Dreber, A., Rand, D.G., Fudenberg, D. and Nowak M.A., 2008. Winners don't punish. Nature, 452: 348-351.

[37] Egas, M., and Riedl, A., 2008. The economics of altruistic punishment and the maintenance of cooperation. Proceedings of the Royal Society B: Biological Sciences, 275: 871-878.

[38] Ertan, A., Page, T. and Putterman, L., 2009. Who to punish? Individual decisions and majority rule in mitigating the free rider problem. European Economic Review, 53: 495-511.

[39] Falk, A., Fehr, E. and Fischbacher, U., 2003. On the nature of fair behavior. Economic Inquiry, 41: 20-26.

[40] Falk, A. and Fischbacher, U., 2006. A theory of reciprocity. Games and Economic Behavior, 54: 293-315.

[41] Fehr, E. and Gachter, S., 2000. Cooperation and punishment in public goods experiments. American Economic Review, 90: 980-994.

[42] Fehr, E. and Gachter, S., 2002. Altruistic punishment in humans. Nature, 415: 137-140.

[43] Fehr, E. and Rockenbach, B., 2003. Detrimental effects of sanctions on human altruism. Nature, 422: 137-140.

[44] Fehr, E. and Schmidt, K. M., 1999. A Theory of fairness, competition, and cooperation. Quarterly Journal of Economics, 114: 817-868.

[45] Fischbacher, U., 2007. Z-tree: Zurich toolbox for ready-made economic experiments. Experimental Economics, 10: 171-78.

[46] Fischbacher, U., Fong, C.M. and Fehr, E., 2009. Fairness, errors and the power of competition. Journal of Economic Behavior and Organization, 72: 527-545.

[47] Fletcher, J.A. and Zwick, M., 2004. Strong altruism can evolve in randomly formed groups. Journal of Theoretical Biology, 228: 303-313.

[48] Fowler, J. H., 2005. Altruistic punishment and the origin of cooperation. Proceedings of the National Academy of Sciences, 102: 7047-7049.

[49] Foster, D. and Young, H. P., 1990. Stochastic evolutionary game dynamics. Theoretical Population Biology, 38: 219-232.

[50] Frank, S. A., 1995. Mutual policing and repression of competition in the evolution of cooperative groups. Nature, 377: 520-522.

[51] Fudenberg, D. and Harris, C., 1992. Evolutionary dynamics with aggregate shocks. Journal of Economic Theory, 57: 420-441.

[52] Gachter, S., Renner, E. and Sefton, M., 2008. The long-run benefits of punishment. Science, 322: 1510-1512.

[53] Gale, J., Binmore, K. and Samuelson, L., 1995. Learning to be imperfect: the ultimatum game. Games and Economic Behavior 8: 56–90.

[54] Gaunersdorfer, A., Hofbauer, J. and Sigmund, K., 1991. On the dynamics of asymmetric games. Theoretical Population Biology, 39: 345-357.

[55] Geritz, S.A.H., Kisdi, E., Meszena, G. and Metz, J.A.J, 1998. Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. Evolutionary Ecology, 12: 35-57.

[56] Goeree, J. K., and Holt, C. A., 2000. Asymmetric inequality aversion and noisy behavior in alternating-offer bargaining games. European Economic Review, 44: 1079-1089.

[57] Goeree, J. K., and Holt, C. A., 2001. Ten little treasures of game theory and ten intuitive contradictions. American Economic Review, 91: 1402-1422.

[58] Goeree, J. K., and Holt, C. A., 2004. A model of noisy introspection. Games and Economic Behavior, 46: 365-382.

[59] Goeree, J. K., Holt, C. A. and Palfrey, T. A., 2002. Quantal response equilibrium and overbidding in private-value auction. Journal of Economic Theory, 104: 247-272.

[60] Goeree, J. K., Holt, C. A. and Palfrey, T. A., 2005. Regular quantal response equilibrium. Experimental Economics, 8: 347-367.

[61] Golman, R. and Page, S.E., 2010. Basins of attraction and equilibrium selection under different learning rules. Journal of Evolutionary Economics, 20: 49-72.

[62] Guillen, P., Schwieren, C. and Staffiero, G., 2006. Why feed the Leviathan? Public Choice, 130: 115-128.

[63] Gurerk, O., Irlenbusch, B. and Rockenbach, B., 2006. The competitive advantage of sanctioning institutions. Science, 312: 108-111.

[64] Guth, W., Schmittberger, R. and Schwarze, B., 1982. An experimental analysis of ultimatum bargaining. Journal of Economic Behavior and Organization, 3: 367-388.

[65] Haile, P. A., Hortacsu, A. and Kosenok, G., 2008. On the empirical content of quantal response equilibrium. The American Economic Review, 98: 180-200.

[66] Harsanyi, J. C., 1973. Oddness of the number of equilibrium points: a new proof. International Journal of Game Theory, 2: 235-250.

[67] Harsanyi, J. C., 1975. The tracing procedure: a Bayesian approach to defining a solution for n-person noncooperative games. International Journal of Game Theory, 4: 61-94.

[68] Harsanyi, J. C. and Selten, R., 1988. A General Theory of Equilibrium Selection in Games, MIT Press, Cambridge.

[69] Henrich, J., 2006. Cooperation, punishment, and the evolution of human institutions. Science, 312: 60-61.

[70] Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H. and McElreath, R., 2001. In search of homo economicus: behavioral experiments in 15 small-scale societies. The American Economic Review, 91: 73-78.

[71] Henrich J, Ensminger, J. McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., Cardenas, J. C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D. and Ziker, J., 2010. Markets, religion, community size, and the evolution of fairness and punishment. Science, 327: 1480-1484.

[72] 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., 2006. Costly punishment across human societies. Science, 312: 1767-1770.

[73] Herrmann, B., Thoni, C. and Gachter, S., 2008. Antisocial punishment across societies. Science, 319: 1362-1367.

[74] Hilbe, C., 2011. Equilibrium notions and framing effects. Mimeo.

[75] Hilbe, C., 2011. Local replicator dynamics: a simple link between deterministic and stochastic models of evolutionary game theory. Bulletin of Mathematical Biology, 73: 2068-2087.

[76] Hilbe, C. and Sigmund, K., 2010. Incentives and opportunism: from the carrot to the stick. Proceedings of the Royal Society B: Biological Sciences, 277: 2427-2433.

[77] Hofbauer, J., 1999. The spatially dominant equilibrium of a game. Annals of Operations Research, 89: 233-251.

[78] Hofbauer, J., Hutson, V. and Vickers, G. T., 1997. Travelling waves for games in economics and biology. Nonlinear Analysis, Theory, Methods and Applications, 30: 1235-1244.

[79] Hofbauer, J. and Sandholm, W., 2002. On the global convergence of stochastic fictitious play. Econometrica, 70: 2265-2294.

[80] Hofbauer, J. and Sandholm, W., 2007. Evolution in games with randomly disturbed payoffs. Journal of Economic Theory, 132: 47-69.

[81] Hofbauer, J. and Sigmund, K., 1998. Evolutionary Games and Population Dynamics, Cambridge University Press, Cambridge.

[82] Hofbauer, J. and Sorger, G., 1999. Perfect foresight and equilibrium selection in symmetric potential games. Journal of Economic Theory, 85: 1-23.

[83] Honda, J., 2012. Equilibrium selection for symmetric coordination games with an application to the minimum-effort game. Mimeo.

[84] Imhof, L. A., 2005. The long-run behavior of the stochastic replicator dynamics. The Annals of Applied Probability, 15: 1019-1045.

[85] Isaac, R. M. and Walker, J. M., 1988. Group size effects in public goods provision: the voluntary contributions mechanism. Quarterly Journal of Economics, 103: 179-199.

[86] Jean-Jacques Herings, P., 2002. Universally converging adjustment processesa unifying approach. Journal of Mathematical Economics, 38: 341-370.

[87] Jong, S.D., Uyttendaele, S. and Tuyls, K., 2008. Learning to reach agreement in a continuous ultimatum game. Journal of Artificial Intelligence Research, 33: 551-574.

[88] Kagel, J.H. and Roth, A.E., 1995. The Handbook of Experimental Economics, Princeton University Press, Princeton.

[89] Kajii, A. and Morris, S., 1997. The robustness of equilibria to incomplete information. Econometrica, 65: 1283-1309.

[90] Kamei, K., Putterman, L., and Tyran, J-R., 2011. State or nature? Formal vs. informal sanctioning in the voluntary provision of public goods. Discussion Papers 11-05, University of Copenhagen. Department of Economics.

[91] Kandori, M., Mailath, G. J. and Rob, R., 1993. Learning, mutation, and long run equilibria in games. Econometrica, 61: 29-56.

[92] Kim, Y., 1996. Equilibrium selection in n-person coordination games. Games and Economic Behavior, 15: 203-227.

[93] Kim, C. and Wong, K.C., 2010. Long-run equilibria with dominated strategies. Games and Economic Behavior, 68: 242-254.

[94] Kirman, A.P. and Vriend, N.J., 2001. Evolving market structure: an ACE model of price dispersion and loyalty. Journal of Economic Dynamics and Control, 25: 459-502.

[95] Kiyonari, T. and Barclay, P., 2008. Free riding may be thwarted by second-order reward rather than by punishment. Journal of Personality and Social Psychology, 95: 826-842.

[96] Kojima, F. and Takahashi, S., 2007. Anti-coordination games and dynamic stability. International Game Theory Review, 9: 667-688.

[97] Kosfeld, M., Okada, A. and Riedl, A., 2009. Institution formation in public goods games. American Economic Review, 99: 1335-1355.

[98] Levine, D.K., 1998. Modeling altruism and spitefulness in experiments. Review of Economic Dynamics, 1: 593-622.

[99] Luce, D., 1959. Individual Choice Behavior, New York: Wiley, New York.

[100] Markussen, T., Putterman, L., and Tyran, J-R., 2011. Self-organization for collective action: an experimental study of voting on formal, informal, and no sanction regimes. Working Papers 2011-4, Brown University, Department of Economics.

[101] Mas-Colell, A., 1974. A note on a theorem of F. Browder. Mathematical Programming, 6: 229-233.

[102] Mas-Colell, A., 1990. The Theory of General Economic Equilibrium: A Differentiable Approach, Cambridge University Press, Cambridge.

[103] Masclet, D., Noussair, C., Tucker, S. and Villeval, M-C., 2003. Monetary and Non-monetary Punishment in the Voluntary Contributions Mechanism. American Economic Review, 93: 366-380.

[104] Matsui, A. and Matsuyama, K., 1995. An approach to equilibrium selection. Journal of Economic Theory, 65:415-434.

[105] Maynard Smith, J. and Szathmary, E., 1995. The Major Transitions in Evolution, New York: Oxford University Press.

[106] McFadden, D., 1976. Quantal choice analysis: a survey. Annals of Economic and Social Measurement, 5: 363-390.

[107] McKelvey, R. D., McLennan, A. M. and Turocy, T. L., 2010. Gambit: Software Tools for Game Theory, Version 0.2010.09.01. http://www.gambit-project.org.

[108] McKelvey, R. D. and Palfrey, T. A., 1995. Quantal response equilibria for normal form games. Games and Economic Behavior, 10: 6-38.

[109] McKelvey, R. D. and Palfrey, T. A., 1998. Quantal response equilibria for extensive form games. Experimental Economics, 1: 9-41.

[110] McKelvey, R. D., Palfrey, T. A. and Weber, R. A., 2000. The effects of payoff magnitude and heterogeneity on behavior in 2×2 games with unique mixed strategy equilibria. Journal of Economic Behavior and Organization, 42: 523-548.

[111] Milnor, J. W., 1965. Topology from the Differentiable Viewpoint, The University Press of Virginia, Charlottesville.

[112] Monderer, D. and Shapley, L. S., 1996. Potential games. Games and Economic Behavior, 14: 124-143.

[113] Morris, S., Rob, R. and Shin, H. S., 1995. p-dominance and belief potential. Econometrica, 63: 145-157.

[114] Morris, S. and Shin, H. S., 2003. Global games: theory and applications. Economics and Econometrics, edited by M. Dewatripont, L. Hansen and S. Turnovsky, Cambridge University Press: 56-114.

[115] Nash, J., 1950. Equilibrium points in n-person games. Proceedings of the National Academy of Sciences, 36: 48-49.

[116] Nowak M.A., Page K.M., and Sigmund K., 2000. Fairness versus reason in the ultimatum game. Science, 289: 1773-1775.

[117] Nowak, M. A. and Sigmund, K., 2004. Evolutionary dynamics of biological games. Science, 303: 793-799.

[118] Nikiforakis, N., 2008. Punishment and counter-punishment in public good games: can we really govern ourselves? Journal of Public Economics, 92: 91-112.

[119] Nikiforakis, N. and Normann, H-T., 2008. A comparative statics analysis of punishment. Experimental Economics, 11: 358-369.

[120] Ohtsuki, H., Iwasa, Y. and Nowak, M.A., 2009. Indirect reciprocity provides only a narrow margin of efficiency for costly punishment. Nature, 457: 79-82.

[121] Oliver, P., 1980. Rewards and punishments as selective incentives for collective action: theoretical investigations. American Journal of Sociology, 2: 493-535.

[122] Olson, M., 1965. The Logic of Collective Action: Public Goods and the Theory of Groups, Harvard: Harvard University Press.

[123] Oosterbeek, H., Sloof, R. and Kuilen, G.V.D., 2004. Cultural differences in ultimatum game experiments: evidence from a meta-analysis. Experimental Economics, 7: 171-188.

[124] Ostrom, E., 1990. Governing the Commons: The Evolution of Institutions for Collective Action, Cambridge: Cambridge University Press.

[125] Ostrom, E., 2005. Understanding Institutional Diversity, Princeton: Princeton University Press.

[126] Page, K.M., Nowak, M.A. and Sigmund, K., 2000. The spatial ultimatum game. Proceedings of the Royal Society B: Biological Sciences, 267: 2177-2182.

[127] Palfrey, T. R. and Prisbrey, J. E., 1997. Anomalous behavior in public goods experiments: how much and why? The American Economic Review, 87: 829-846.

[128] Rockenbach, B. and Milinski, M., 2006. The efficient interaction of indirect reciprocity and costly punishment. Nature, 444: 718-723.

[129] Roth, A.E. and Erev, I., 1995. Learning in extensive-form games: experimental data and simple dynamic models in the intermediate term. Games and Economic Behavior, 8: 164-212.

[130] Roth, A.E., Prasnikar, V., Okuno-Fujiwara, M. and Zamir, S., 1991. Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo. The American Economic Review, 81: 1068-1095.

[131] Samuelson, L., 1997. Evolutionary games and equilibrium selection, MIT Press, Cambridge.

[132] Samuelson, L., 2002. Evolution and game theory. The Journal of Economic Perspectives, 16: 47-66.

[133] Samuelson, L. and Zhang, J., 1992. Evolutionary stability in asymmetric games. Journal of Economic Theory, 57: 363-392.

[134] Sandholm, W. H., 2010. Population Games and Evolutionary Dynamics, MIT Press, Cambridge.

[135] Schlag, K., 1998. Why imitate, and if so, how? A boundedly rational approach to multi-armed bandits. Journal of Economic Theory, 78: 130-156.

[136] Sefton, M., Shupp, R. and Walker, J., 2007. The effect of rewards and sanctions in provision of public goods. Economic Inquiry, 45: 671-690.

[137] Selten, R., 1980. A note on evolutionarily stable strategies in asymmetric animal conflicts. Journal of Theoretical Biology, 84: 93-101.

[138] Selten, R. and Stoecker, R., 1986. End behavior in sequences of finite Prisoner's Dilemma supergames: a learning theory approach. Journal of Economic Behavior and Organization, 7: 47-70.

[139] Sigmund, K., 2007. Punish or perish? Retaliation and cooperation among humans. Trends in Ecology and Evolution, 22: 593-600.

[140] Sigmund, K., 2010. The Calculus of Selfishness, Princeton: Princeton University Press.

[141] Sigmund, K., De Silva, H., Traulsen, A. and Hauert, C., 2010. Social learning promotes institutions for governing the commons. Nature, 466: 861-863.

[142] Sigmund, K., Hauert, C. and Nowak, M. A., 2001. Reward and punishment. Proceedings of the National Academy of Sciences, 98: 10757-10762.

[143] Sigmund, K., Hauert, C., Traulsen, A. and De Silva, H., 2011. Social control and the social contract: the emergence of sanctioning systems for collective action. Dynamic Games and Applications, 1: 149-171.

[144] Staudigl, M., 2011. Stochastic stability in asymmetric binary choice coordination games. Games and Economic Behavior, in press.

[145] Sugden, R., 1986. The economics of rights, cooperation and welfare, Oxford: Blackwell.

[146] Sutter, M., Haigner, S. and Kocher, M. G., 2010. Choosing the carrot or the stick? Endogeneous institutional choice in social dilemma situations. Review of Economic Studies, 77: 1540-1566.

[147] Taylor, P. and Jonker, L., 1978. Evolutionary stable strategies and game dynamics. Mathematical Biosciences, 50: 145-156.

[148] Tiebout, C., 1956. A pure theory of local expenditures. Journal of Political Economy, 64: 416-424.

[149] Traulsen, A., Rohl, T. and Milinski, M., 2012. Humans prefer pool punishment to maintain the commons, mimeo.

[150] Traulsen, A., Semmann, D., Sommerfeld, R. D., Krambeck, H-J. and Milinski, M., 2010. Human Strategy Updating in Evolutionary Games. Proceedings of the National Academy of Sciences, 107: 2962-2966.

[151] Turocy, T. L., 2005. A dynamic homotopy interpretation of the logistic quantal response equilibrium correspondence. Games and Economic Behavior, 51: 243-263.

[152] Tyler, T. R. and Degoey, P., 1995. Collective restraint in social dilemmas: procedural justice and social identification effects on support for authorities. Journal of Personality and Social Psychology, 69: 482-497.

[153] Ui, T., 2001. Robust equilibria of potential games. Econometrica, 69: 1373-1380.

[154] Ui, T., 2002. Quantal response equilibria and stochastic best response dynamics. Working paper.

[155] Ui, T., 2006. Correlated quantal responses and equilibrium selection. Games and Economic Behavior, 57: 361-369.

[156] Van Vugt, M., Henrich, J. and O'Gorman, R., 2009. Constraining free riding in public good games: designated solitary punishers can sustain human cooperation. Proceedings of the Royal Society B: Biological Sciences, 276: 323-329.

[157] Walker, J. M., Gardner, R., Herr, A. and Ostrom, E., 2000. Collective choice in the commons: experimental results on proposed allocation rules and votes. The Economic Journal, 110: 212-34.

[158] Weibull, J., 1995. Evolutionary Game Theory, MIT Press, Cambridge.

[159] Wilson, D.S., 1975. A theory of group selection. Proceedings of the National Academy of Sciences, 72: 13-146.

[160] Wu, J., Zhang, B., Zhou, Z., He, Q., Zheng, X., Cressman, R. and Tao, Y., 2009. Costly punishment does not always increase cooperation. Proceedings of the National Academy of Sciences, 106: 17448-17451.

[161] Yamagishi, T., 1986. The provision of a sanctioning system as a public good. Journal of Personality and Social Psychology, 51: 110-116.

[162] Yi, K., 2002. A quantal response equilibrium model of order-statistic games. Journal of Economic Behavior and Organization, 51: 413-425.

[163] Yi, K., 2005. Quantal response equilibrium models of the ultimatum bargaining game. Games and Economic Behavior, 51: 324-348.

[164] Young, H. P., 1993. The evolution of conventions. Econometrica, 61: 57-84.

[165] Zeeman, E. C., 1980. Population dynamics from game theory. Global Theory of Dynamical Systems, 819: 471-497.

Curriculum Vitae

Personal Information

Name:	Boyu Zhang
Date of birth:	September 15th, 1986
Place of birth:	Beijing, P.R. China
Citizenship:	P.R. China

Education

11/2009 - present: University of Vienna

Major:	Mathematics
Supervisor:	Prof. Josef Hofbauer, Prof. Karl Sigmund
Degree:	Expecting a doctor's degree in August 2012
PhD thesis:	Evolutionary games and equilibrium selection.

09/2006 - 07/2009: Beijing Normal University

Major:	Mathematics
Supervisor:	Prof. Zhonglai Li
Degree:	M.A. (July 2009)
Master thesis:	Evolutionary game dynamics with impulsive effects.

05/2006 -10/2009: Institute of Zoology, Chinese Academy of Science (joint student)

Major:	Theoretical ecology
Supervisor:	Prof. Yi Tao

09/2002 - 07/2006: Beijing Normal University

Major:	Mathematics
Degree:	B.A. (July 2006)

Selected Publications

[1] Zhang, B., Zhang, Z., Li, Z. and Tao, Y.. 2007. Stability analysis of a two-species model with transitions between population interactions. Journal of Theoretical Biology,

248: 145-153.

[2] Wang, S., Zhang, B., Li, Z., Cressman, R. and Tao, Y.. 2008. Evolutionary game dynamics with impulsive effects. Journal of Theoretical Biology, 254: 384-389.

[3] Tao, Y., Cressman, R., Zhang, B. and Zheng, X.. 2008. Stochastic fluctuations in frequency-dependent selection: A one-Locus, two-allele and two-phenotype model. Theoretical Population Biology, 74: 263-272.

[4] Ji, T.*, Zhang, B.*, Sun, Y. and Tao, Y. (joint first authors). 2009. Evolutionary dynamics of fearfulness and boldness. Journal of Theoretical Biology, 256: 637-643.

[5] Wu, J., Zhang, B., Zhou, Z., He, Q., Zheng, X., Cressman, R. and Tao, Y.. 2009. Costly punishment does not always increase cooperation. Proceedings of the National Academy of Sciences, 106: 17448-17451.

[6] Zhang, B., Tao, Y. and Cressman, R.. 2010. Cooperation and stability through periodic impulses. PloS ONE, 5, doi: 10.1371/journal.pone.0009882.

[7] Cressman, R, Song, J., Zhang, B. and Tao, Y.. 2012. Cooperation and evolutionary dynamics in the public goods game with institutional incentives. Journal of Theoretical Biology, 299: 144-151.

[8] Zhang, D, Zhang, B., Lin, K., Tao, Y., Hubbell, S., He, F. and Ostling, A.. 2012. Demographic trade-offs determine species abundance and diversity. Journal of Plant Ecology, 5: 82-88.

[9] Zhang, B. and Hofbauer, J. Quantal response methods for equilibrium selection. Accepted by GAMES2012. (See Chapter 2)

[10] Zhang, B., Cong, L., De Silva, H., Bednarik, P. and Sigmund, K. The evolution of sanctioning institutions: an experimental approach. Submitted. (See Chapter 5)

[11] Wu, J.*, Li, C.*, Zhang, B.*, Xiao, H., He, Q,. and Tao, Y. and Cressman, R. (joint first authors). Cooperation through institutional reward and punishment: Group versus individual incentives. Submitted.

Selected Presentations

[1] Evolutionary games and impulsive differential equations. Seminar Series: Arbeitsgemeinschaft Biomathematik, November 2009, Vienna, Austria.

[2] Does punishment always increase cooperation? Seminar Series: Arbeitsgemeinschaft Biomathematik, January 2010, Vienna, Austria.

[3] Homotopy methods for finding quantal response equilibria. Equilibrium Computation, April 2010, Dagstuhl, Germany.

[4] Cooperation through institutional reward and punishment. Seminar Series: Arbeitsgemeinschaft Biomathematik, June 2010, Vienna, Austria.

[5] On the evolution of cooperation. Invited by Key Laboratory of Management, Decision and Information Systems, Chinese Academy of Sciences, March 2011, Beijing, China.

[6] Social learning in the ultimatum game. European Conference in Complex Systems, September 2011, Vienna, Austria.

[7] Fashion games in social networks. Seminar Series: Arbeitsgemeinschaft Biomathematik, June 2012, Vienna, Austria.

[8] Quantal response methods for equilibrium selection. 4th World Congress of the Game Theory Society, July 2012, Istanbul, Turkey.