

Social Learning in the Ultimatum Game

Boyu Zhang*

School of Mathematical Sciences, Beijing Normal University, Beijing, P.R. China

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 humans prefer fair share. In this paper, we study the ultimatum game by a learning-mutation process based on quantal response equilibrium, where players are assumed boundedly rational and make mistakes when estimating the payoffs of strategies. Social learning is never stabilized at the fair outcome or the rational outcome, but leads to oscillations from offering 40 percent to 50 percent. To be precise, 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. If mutations occur rarely, fair behavior is favored in the limit of local mutation. If mutation rate is sufficiently high, fairness can evolve for both local mutation and global mutation.

Citation: Zhang B (2013) Social Learning in the Ultimatum Game. PLoS ONE 8(9): e74540. doi:10.1371/journal.pone.0074540

Editor: Matjaz Perc, University of Maribor, Slovenia

Received: June 8, 2013; **Accepted:** August 5, 2013; **Published:** September 4, 2013

Copyright: © 2013 Boyu Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The author is thankful for the support of the Chinese Scholarship Council (CSC). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The author has declared that no competing interests exist.

* E-mail: zhangboyu5507@gmail.com

Introduction

Ultimatum game introduced by Guth et al. [1] 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 on 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 the 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 [2–7].

How to understand people rejecting positive offers? One well known economic explanation is that irrational individuals have preference on fairness [8–9]. 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 closer to the fair share [9–12]. However, this contradicts the equilibrium analysis since the only subgame perfection is not to reject.

From the perspective of evolutionary game theory, replicator dynamics, which assumes that successful strategies will always come to dominate the population, favors low offers and demands [13–16]. Therefore, to explain fairness using evolutionary

dynamics, additional mechanisms are necessary. One approach involves reputation system. If proposers can obtain information about responders' demands and believe that responders reject offers lower than their aspiration level, high offers and demands will prevail the population [14,16]. Spatial structure was also found to play a key role in the evolution of fairness. Pioneering work by Page et al. [17] pointed out that ultimatum game on ring and lattice may evolve to much fairer outcomes compared with random encounter setting. Subsequent research confirmed that fairness is enhanced in heterogenous graphs [18]. Furthermore, spatial ultimatum game with discrete strategies exhibits fascinatingly rich dynamical behavior such as traveling waves and cyclic dominance [19], and the fair solution is obtained in the limit of a continuous strategy set [20]. A third approach emphasizes the importance of randomness [13,21]. By using stochastic evolutionary game theory, Rand et al. [21] demonstrated that natural selection favors fair behavior when sufficient randomness is present.

In this paper, we study the iterated ultimatum game by a learning-mutation process. 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 a 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 a responder, where $0 \leq x, y \leq 1$ and $0 < p \leq 1$. The parameter y can be seen as proposer's demand or aspiration level. Following this definition, the rational strategy is written as $S(\varepsilon, \varepsilon, 1)$, where ε is the minimum offer greater than 0, and the fair strategy is $S(\frac{1}{2}, \frac{1}{2}, 1)$. In our model, individuals update their strategies through quantal response learning. We will show that learning-mutation process leads to oscillations from offering 0.4 to 0.5 and rejecting lower offers with probability one half.

Methods

Before study the ultimatum game with continuous strategies, we first consider the iterated mini ultimatum game with only two possible offers h and l , with $0 \leq l < h \leq 1$ [13–14,16,22–23]. In each round, the proposer has to choose 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

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

where the proposer plays rows and the responder plays columns. The mini game has a strict Nash equilibrium (L, L) , and non-isolated Nash equilibria $(H, sH + (1-s)L)$, where $\frac{h-l}{1-l} \leq s \leq 1$. Notice that 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.

Payoff matrix (1) can also be interpreted as the Prisoner’s Dilemma game with punishment, where higher offer and low offer correspond to cooperation and defection, respectively, and rejecting the low offer means paying l to punish defector $1-l$ [16,22]. Similarly as the mini ultimatum game, (L, L) is the only subgame perfection.

There are many ways to model social learning [10,13,24–26]. In this paper, we apply the well-known quantal response equilibrium (QRE) introduced by McKelvey and Palfrey [27–29]. In a QRE, players are assumed boundedly rational and do not always choose best responses. Instead, they make decisions based on probabilistic choice functions and believe 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. The most common specification of QRE is the logit equilibrium, where noises follow the extreme value distribution [30–32]. Let u_{ij} denotes the expected payoff of player i using strategy j ($j \in \{1, \dots, J_i\}$). The logistic response function is defined as

$$\sigma_{ij}(\vec{u}_i) = \frac{e^{\lambda u_{ij}}}{\sum_{k=1}^{J_i} e^{\lambda u_{ik}}}, \tag{2}$$

where σ_{ij} is the probability that player i adopts strategy j and $\vec{u}_i = (u_{i1}, \dots, u_{iJ_i})$. If each player uses a logistic response function, QRE or logit equilibria are the solutions of $\pi_{ij} = \sigma_{ij}$, where π_{ij} is the frequency of strategy j in player i . The logistic response function has one free parameter 2λ , who has been interpreted as the intensity of selection [16,21]. At $2\lambda = 0$, players have no information about the game and each strategy is chosen with equal probability. As 2λ 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. For instance, Yi [29] applied QRE to fit data of high stakes ultimatum games [33]. In iterated games, estimates of 2λ usually increase as the game progresses [27–29]. 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 2λ . When $2\lambda = 0$, the QRE is at the centroid of the strategy

simplex, and when $2\lambda \rightarrow +\infty$, the QRE set consists of Nash equilibria only. As pointed out by McKelvey and Palfrey [27], for almost all normal form games, the graph of logit equilibria correspondence contains a unique branch which starts for $2\lambda = 0$ at the centroid and converges to a unique Nash equilibrium as 2λ goes to infinity. This then defines a unique selection from the set of Nash equilibria 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.

Results

For the mini ultimatum game Eq.(1), logit equilibria are solutions of

$$\begin{aligned}
 \pi_1 &= \frac{1}{1 + e^{\lambda(1-l)(q-\pi_2)}}, \\
 \pi_2 &= \frac{1}{1 + e^{\lambda(1-\pi_1)}}, \tag{3}
 \end{aligned}$$

where π_i denotes the probability of player i using his first strategy and $q = \frac{h-l}{1-l}$. When $2\lambda = 0$, Eq.(3) has a unique solution $(\frac{1}{2}, \frac{1}{2})$, and when $2\lambda \rightarrow +\infty$, the QRE set consists of three Nash equilibria only, $(0, 0)$ $(1, \frac{1}{2})$ and $(1, q)$ (see Figures S1 and S2). In general, the LLE is one of two Nash equilibria, 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, the LLE is $(H, \frac{H}{2} + \frac{L}{2})$ if and only if

$$2h < l + 1 - \sqrt{l(1-l)}. \tag{4}$$

(See Text S1).

If the high offer is the fair offer, i.e., $h = \frac{1}{2}$, Eq.(4) tells us that social learning chooses the low offer. In fact, any high offer equal or greater than $\frac{1}{2}$ is unfavored. On the other hand, if the low offer is the rational decision, i.e., $l = \varepsilon$, any high offer smaller than $\frac{1-\sqrt{\varepsilon}}{2}$ is selected. Therefore, social learning does not always prefer 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 two offers x_1 and x_2 . Dominant regions of x_1 and x_2 are shown in Figure 1. Offers lower than $\frac{1}{2}$ are dominated by slightly higher offers. For $l < \frac{1}{2}$, the right side of Eq.(4) 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^* , see the red point in Figure 1).

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 paired and play the iterated mini ultimatum game. We begin by considering the single role mode, in which roles of two members in a group are decided randomly before the game starts and do not change in an interaction, and discuss later that main results are qualitatively unchanged if they play both roles. Two players update their strategies by the quantal response learning and the iterated game 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

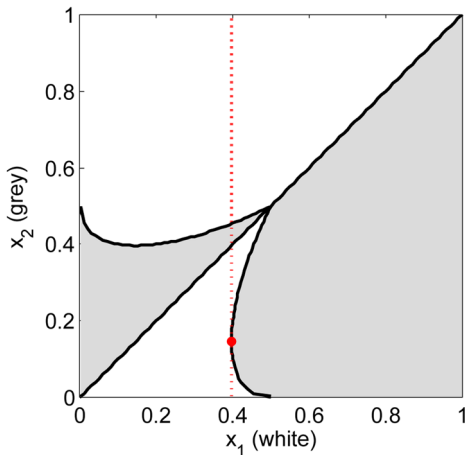


Figure 1. Pairwise invisibility plot. x_1 and x_2 are dominant in white and gray regions, respectively. The red dash line denotes $x_1 = 0.4$ and the red point denotes $(x_1, x_2) = (0.4, 0.15)$. Every offer x_1 lower than 0.5 is dominated by some higher offers and x_1 equal or greater than 0.4 is also dominated by lower offers. In particular, $x_1 = 0.4$ dominates almost all lower offers, the only exception is $x_2 = 0.15$ (the red point). doi:10.1371/journal.pone.0074540.g001

new strategy. In local mutation scenario, they plus or minus a small random value on their former strategies, and in global mutation scenario, all strategies are drawn with equal probability (See Figure 2).

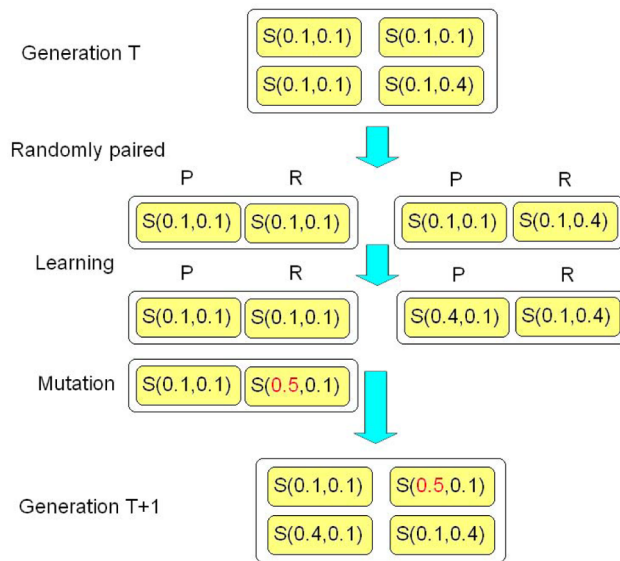


Figure 2. An example for the learning-mutation process. An example for the learning-mutation process in a population of four players from generation T to generation $T + 1$. In generation T , three players adopt $S(0.1, 0.1)$ and one adopts a mutant strategy $S(0.1, 0.4)$. At the beginning, they are 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 do not change their original strategies, while in the second group, the proposer will change his 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 $S(0.1, 0.4)$. doi:10.1371/journal.pone.0074540.g002

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 offer lower than y_2 with probability p_2 . The payoff of this game is described by (1), where $h = \max\{x_1, y_2\}$ and $l = \min\{x_1, y_2\}$. At the beginning, both players have the motivation to adjust their strategies. If $y_2 > x_1$, y_2 is the high offer h and x_1 is the low offer l . 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 demand from y_2 to x_1 . Conversely, if $x_1 > y_2$, the proposer wants to decrease his offer from x_1 to y_2 and the responder wants to increase his demand from y_2 to x_1 .

According to the quantal response learning, both players choose their initial action randomly, i.e., the proposer suggests the lower offer with probability one half and the responder rejects this offer with the same probability. An approximated formula to decide the LLE is provided by Eq.(4). At the LLE, the responder either accepts the low offer or rejects the low offer with probability one half. This implies that the rejection rate p will converge to $\frac{1}{2}$ in the long run. For simplicity, we write $S(x, y) = S(x, y, \frac{1}{2})$ in later discussions. At the end of 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 $S(x_2, x_1)$. However, 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 the diversity of offers decreases from one generation to another since in each mini game, one dominated offer is eliminated. In fact, for any (heterogenous) population, quantal response learning will eventually lead to a homogenous self-consistent population in which each player's offer equals to his demand.

The same result arises if players act both roles. To see this, suppose that two players using strategies $S_1(x_1, y_1)$ and $S_2(x_2, y_2)$ are randomly paired. In the interaction, they will play two games $UG(S_1(x_1, y_1), S_2(x_2, y_2))$ and $UG(S_2(x_2, y_2'), S_1(x_1, y_1))$, where player 1 acts as a proposer in the first game and acts as a responder in the second game, and x'_1 and y'_2 denote respectively the offer of player 1 and the demand of player 2 after the first game. Notice that x_1 competes with y_2 in the first game and x_2 competes with y_1 in the second game, the final outcome does not change with the order of two games, and two dominated offers are eliminated at the end of the interaction. Thus, similarly as the single role mode, the diversity of offers decreases and the population will converge to a homogeneous state.

Let us now add the possibility of mutation. We first look at the limit of weak mutation rate $\mu \rightarrow 0$. As in the adaptive dynamics model, mutations occur rarely so that a mutant will either vanish or has taken over the population before the next mutation happens [15,34–35]. Under weak mutation, the population is homogeneous in most of time, and we represent the strategy of the residents $S(x_r, x_r)$ by x_r . Eq.(4) indicates that (a) if $x_r < h^*$, the population could only be replaced by mutants using offers higher than x_r , (b) if $h^* \leq x_r < \frac{1}{2}$, both higher and lower offers may invade, (c) if $\frac{1}{2} \leq x_r$, any lower offer could take over the population and higher offer can not invade (See Figure 1). Generally speaking, the learning-mutation process leads to oscillations in 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 (see Figure 3).

The long-term mean offer of the population depends significantly on the range of mutation. In local mutation scenario where the mutational jumps are small such that the resident strategy

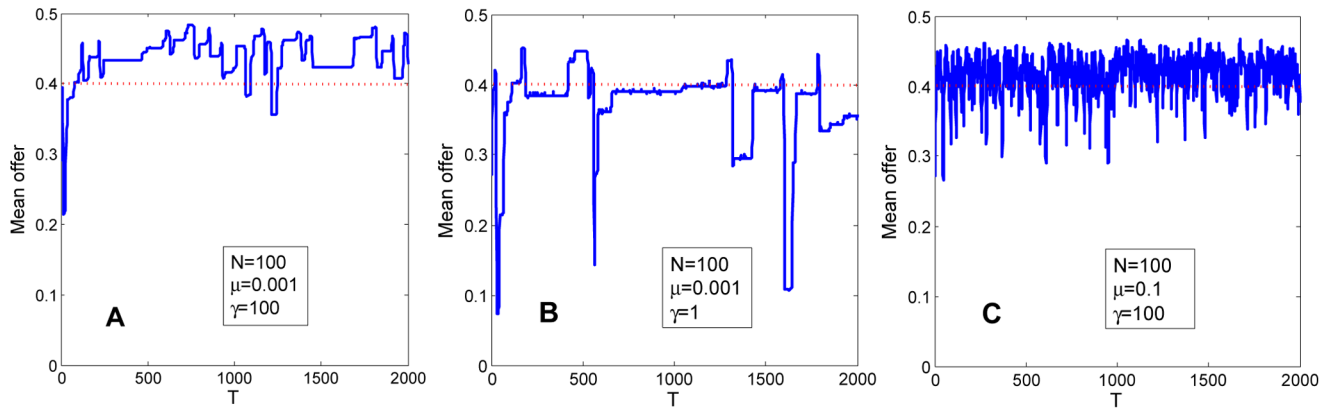


Figure 3. Time evolution of the population mean offer. The population size is 100 and evolves under the learning-mutation process. Mutation rates are taken as $\mu=0.001$ (weak mutation rate) in Figures 3A and 3B and $\mu=0.1$ (intermediate mutation rate) in Figure 3C. Mutant strategies follow β -distributions, where $2\gamma=100$ (local mutation) in Figures 3A and 3C, and $2\gamma=1$ (global mutation) in Figure 3B. Red dash lines denote $x=0.4$. In all figures, the population mean offer increases if it is smaller than 0.4 but oscillates if it is between 0.4 to 0.5. doi:10.1371/journal.pone.0074540.g003

changes continuously, the time evolution of x_r could be described by the adaptive dynamics [15,34–35]. It is easy to verify that $\frac{dx_r}{dt} > 0$ if $x_r < \frac{1}{2}$ but $\frac{dx_r}{dt} < 0$ if $x_r \geq \frac{1}{2}$. $x_r = \frac{1}{2}$ is a degenerate point of the adaptive dynamics, i.e., the resident strategy will decrease when it reaches the fair offer. In this case, the long-term mean offer is slightly lower than 0.5 (see Figure 4A). In contrast, if mutations are global that are picked from the uniform distribution $[0,1]$, x_r may drop dramatically when it is greater than 0.4 (see Figure 3B). Thus, broader range of mutation leads to lower mean offer. To illustrate this intuition, suppose that mutant strategies in x_r population are picked from a β -distribution $\text{Beta}(\alpha, \beta)$, where $\alpha = \frac{x_r 2\gamma - 2x_r + 1}{1 - x_r}$ and $\beta = 2\gamma$ [21]. The particular values of α and β are chosen such that the modal value of the distribution is x_r . The inverse of 2γ measures the range of mutation, where $2\gamma=1$ means uniform distribution on $[0,1]$ and $2\gamma \rightarrow +\infty$ means local mutation. Figure 4A shows clearly that the long-term mean offer of the population is monotonically increasing in 2γ . When $2\gamma=1$, the mean offer is about 0.32, and when $2\gamma \rightarrow +\infty$, it converges to the fair outcome 0.5.

In the high mutation limit $\mu \rightarrow 1$ with global mutation $2\gamma=1$, all strategies are present in the population simultaneously with

approximately equal frequency and therefore the long-term mean offer is exactly 0.5 [21]. As 2γ increase, the mean offer will first decline to below 0.4. This follows from the fact that a successful strategy should maximize the winning probability when playing against a randomly chosen strategy. As shown in Figure 1, offers around 0.4 have smaller dominated regions therefore are preferred in a random world. However, when the mutation range becomes narrow, the mean offer will increase since any offer lower than 0.5 are dominated by slightly higher offers. Finally, it converges to 0.5 as $2\gamma \rightarrow +\infty$ (see Figure 4C).

At intermediate mutation rates $1 < N\mu < N$ where mutations happen frequently and the population has a high diversity of strategies, the time evolution of the one generation population mean offer could be characterized by the theoretical predictions of the weak mutation limit [15]. That is, the one generation mean offer increases if it is smaller than 0.4 but oscillates if it is between 0.4 to 0.5 (see Figure 3C). Although a smaller population converges to homogeneous state easier than a larger population, numerical simulations suggest that the long-term mean offer is robust to different population sizes (see Figure 4). In particular, intermediate mutation rates result in intermediate mean offer. For global mutation $2\gamma=1$, the long-term mean offer is higher than

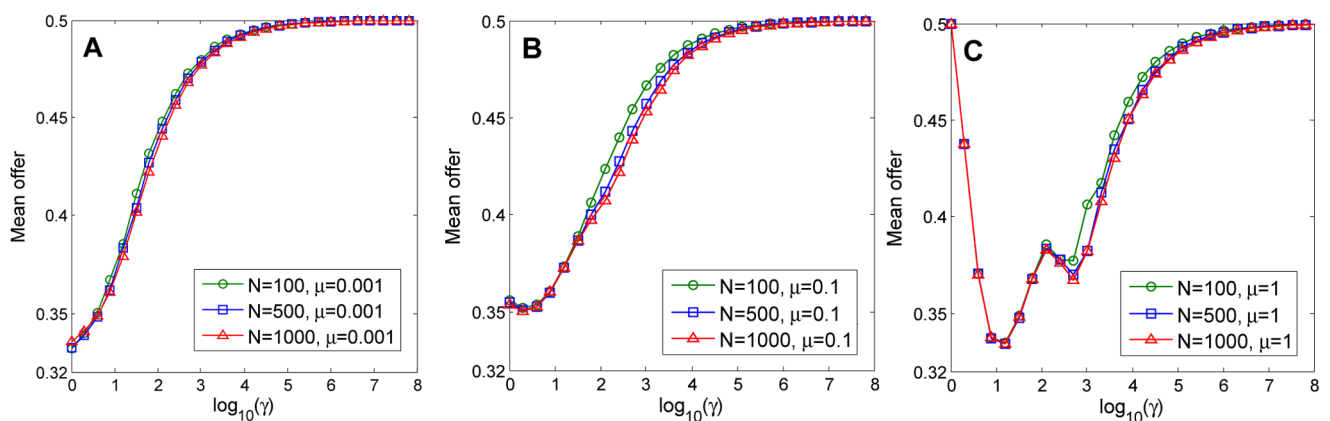


Figure 4. Effects of mutations and population sizes on the long-term mean offer. Mutation rates are taken as $\mu=0.001$ (weak mutation rate), $\mu=0.1$ (intermediate mutation rate) and $\mu=1$ (high mutation rate) in Figures 3A, 3B and 3C, respectively. The long-term mean offer depends significantly on the mutation rate and the mutation range, but is robust to different population sizes. doi:10.1371/journal.pone.0074540.g004

that of weak mutation limit 0.32 but lower than that of high mutation limit 0.5, and when $2\gamma \rightarrow +\infty$, the long-term mean offer goes to 0.5 as the two extremes (see Figure 4B).

Discussion

Similarly as [14], we assume that a proposer knows what offers have been accepted by the responder in the last generation. This information can be easily obtained since every group reaches an agreement on a certain offer at the end of a generation. Given this information, the proposer should either to stay unchanged or to adopt the responder's strategy. In fact, switching to any other offer is unreasonable: giving offers lower than the demand will face the risk of being rejected and giving offers higher than the demand is inefficient. If the ultimatum game is anonymous such that the responder's demand is unknown to the proposer, the proposer then chooses offers from interval $[0,1]$. In this case, the LLE of the ultimatum game is the rational equilibrium, where the proposer offers the minimal and the responder accepts any non-zero offer [29].

A difference between our model and [14] is that we include the possibility that responders accept lower offers. Consequently, proposers have to evaluate whether to satisfy responders' demands. According to quantal response learning, two players in a group choose their initial strategies randomly. The motivation is twofold. On the one hand, each player faces a new game in a new interaction since the payoff matrix of the mini ultimatum game is decided by the strategies of both players. On the other hand, empirical evidences from the repeated Prisoner's Dilemma games (with punishment) support this consideration. The frequency of cooperation (which corresponds to the frequency of high offer in the mini ultimatum game) in the initial round of each interaction is nearly the same and decreases over rounds [3,25,36–37]. We can then expect that players are affected little by past experience and update their strategies entirely by social learning.

We consider that players are boundedly rational and choose the best response according to noisy observations. In an interaction, two players update their strategies simultaneously. At the beginning, the proposer is inclined to make the high offer due to the high rejection rate and the responder tends to accept the lower offer since rejecting is costly. 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 payoffs of strategies lead to fairer outcome.

We note that there exist at least two theoretical papers investigated the effects of randomness on ultimatum game [13,21]. Gale et al. [13] studied noisy replicator dynamics with asymmetric mutation structure in which responders attempt to adopt new strategies more often than proposers. The greater variation in responder behavior then forces proposers to make higher offers. Rand et al. [21] analyzed the stochastic evolutionary

dynamics under weak selection. They showed that larger mutation rate could lead to a heterogeneous population with higher average offer and acceptance level. In the high mutation limit, as has been shown in [17], the prevailing strategy is $(x,y) = (\frac{1}{2}, 0)$, while in the low mutation limit, the prevailing strategy is $(x,y) = (\frac{1}{3}, \frac{1}{3})$. Furthermore, their results are robust to the range of mutation.

In our paper, the intensity of selection is weak at the beginning of an interaction and increases as players gain experience from repeated observations. At the end of the interaction, the selection is strong and players only choose the strategies with the highest payoff. Different from [21], the long-term mean offer is affected by both the mutation rate and the mutation range. For low and intermediate mutation rates, the mean offer increases with decreasing mutation range, and fair behavior is favored in the limit of local mutation. In contrast, for high mutation rates, both global mutation and local mutation lead to fair outcome.

The emergence of equity is as complicated as the evolution of human society. Our model excluded many important issues, such as preference on fairness [8–9] or punishment [38–40], social networks [17–20], emotions [41–42] and culture difference [4–6,43]. Based on learning and mutation, we show that individuals entirely motivated by self interests can evolve toward fairness.

Supporting Information

Figure S1 Graph of QRE correspondence. Parameters are taken as $l=0.05$, $h=0.4$ in Figure S1A and $h=0.5$ in Figure S1B. Blue curve and red curve are π_1 and π_2 , respectively. In Figure S1A, since 0.4 dominates 0.05, $(1, \frac{1}{2})$ is the LLE. In Figure S1B, since 0.5 is dominated by 0.05, $(0,0)$ is the LLE.
(TIF)

Figure S2 Graph of Eq.(S7). Parameters are taken as $h = \frac{3}{8}$ and $l = \frac{1}{6}$, i.e., $q = \frac{1}{4}$. $2\lambda \geq 0$ on solid curves but $2\lambda < 0$ on the dashed curve. Black points are Nash equilibria and the red point is the LLE. The graph of Eq.(S7) consists of two branches, where one passes through the Nash equilibrium $(1, \frac{1}{4})$ and the other passes through the centroid $(\frac{1}{2}, \frac{1}{2})$. Since $(1, \frac{1}{2})$ and $(\frac{1}{2}, \frac{1}{2})$ are on the same branch, higher offer is the LLE.
(TIF)

Text S1 Supporting Information for Social learning in the ultimatum game.
(PDF)

Acknowledgments

This paper is a part of my PhD thesis, written under the guidance of Josef Hofbauer and Karl Sigmund. I am very grateful to them. I also thank Claus Ruffler and two anonymous reviewers for useful comments.

Author Contributions

Conceived and designed the experiments: BZ. Performed the experiments: BZ. Analyzed the data: BZ. Contributed reagents/materials/analysis tools: BZ. Wrote the paper: BZ.

References

- Guth W, Schmittberger R, Schwartz B (1982) An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization* 3: 367–388.
- Roth AE, Prasnikar V, Okuno-Fujiwara M, Zamir S (1991) Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo. *The American Economic Review* 81: 1068–1095.
- Kagel JH, Roth AE (1995) *The Handbook of Experimental Economics*. Princeton: Princeton University Press.
- Henrich J, Boyd R, Bowles S, Camerer C, Fehr E, et al. (2001) In search of homo economicus: Behavioral experiments in 15 small-scale societies. *The American Economic Review* 91: 73–78.

5. Henrich J, Ensminger J, McElreath R, Barr A, Barrett C, et al. (2010) Markets, religion, community size, and the evolution of fairness and punishment. *Science* 327: 1480–1484.
6. Oosterbeek H, Sloof R, Kuilen GVD (2004) Cultural differences in ultimatum game experiments: Evidence from a meta-analysis. *Experimental Economics* 7: 171–188.
7. Cooper DJ, Dutcher EG (2010) The dynamics of responder behavior in ultimatum games: A meta-study. *Experimental Economics* 14: 519–546.
8. Fehr E, Schmidt KM (1999) A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics* 114: 817–868.
9. Bolton GE, Ockenfels A (2000) ERC: A theory of equity, reciprocity, and competition. *The American Economic Review* 90: 166–193.
10. Roth AE, 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.
11. Brenner T, Vriend NJ (2006) On the behavior of proposers in ultimatum games. *Journal of Economic Behavior and Organization* 61: 617–631.
12. Fischbacher U, Fong CM, Fehr E (2009) Fairness, errors and the power of competition. *Journal of Economic Behavior and Organization* 72: 527–545.
13. Gale J, Binmore K, Samuelson L (1995) Learning to be imperfect: The ultimatum game. *Games and Economic Behavior* 8: 56–90.
14. Nowak MA, Page KM, Sigmund K (2000) Fairness versus reason in the ultimatum game. *Science* 289: 1773–1775.
15. Pge KM, Nowak MA (2001) A generalized adaptive dynamics framework can describe the evolutionary ultimatum game. *Journal of Theoretical Biology* 209: 173–179.
16. Sigmund (2010) *Calculus of selfishness*. Princeton: Princeton University Press.
17. Page KM, Nowak MA, Sigmund K (2000) The spatial ultimatum game. *Proceedings of the Royal Society B: Biological Sciences* 267: 2177–2182.
18. Sinatra R, Iranzo J, Gomez-Gardeñes J, Floría LM, Latora V, et al. (2009) The ultimatum Game in complex networks. *Journal of Statistical Mechanics: Theory and Experiment* 09: P09012.
19. Szolnoki A, Perc M, Szabó G (2012) Defense mechanisms of empathetic players in the spatial ultimatum game. *Physical Review Letters* 109: 078701.
20. Szolnoki A, Perc M, Szabó G (2012) Accuracy in strategy imitations promotes the evolution of fairness in the spatial ultimatum game. *Europhysics Letters* 100: 28005.
21. Rand DG, Tarnita CE, Ohtsuki H, Nowak MA (2013) Evolution of fairness in the oneshot anonymous ultimatum game. *Proceedings of the National Academy of Sciences* 110: 2581–2586.
22. Sigmund K, Hauert C, Nowak MA (2001) Reward and punishment. *Proceedings of the National Academy of Sciences* 98: 10757–10762.
23. Falk A, Fehr E, Fischbacher U (2003) On the nature of fair behavior. *Economic Inquiry* 41: 20–26.
24. Abbink K (2001) Adaptive learning versus punishment in ultimatum bargaining. *Games and Economic Behavior* 37: 1–25.
25. Selten R, 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.
26. Kirman AP, Vriend NJ (2001) Evolving market structure: An ACE model of price dispersion and loyalty. *Journal of Economic Dynamics and Control* 25: 459–502.
27. Mckelvey RD, Palfrey TA (1995) Quantal response equilibria for normal form games. *Games and Economic Behavior* 10: 6–38.
28. Mckelvey RD, Palfrey TA (1998) Quantal response equilibria for extensive form games. *Experimental Economics* 1: 9–41.
29. Yi K (2005) Quantal response equilibrium models of the ultimatum bargaining game. *Games and Economic Behavior* 51: 324–348.
30. Blume LE (1993) The statistical mechanics of strategic interaction. *Games and Economic Behavior* 5: 387–424.
31. Blume LE (1995) The statistical mechanics of best-response strategy revision. *Games and Economic Behavior* 11: 111–145.
32. Turocy TL (2005) A Dynamic homotopy interpretation of the logistic quantal response equilibrium correspondence. *Games and Economic Behavior* 51: 243–263.
33. Slonim R, Roth A (1998) Learning in high stakes ultimatum games: An experiment in the Slovak Republic. *Econometrica* 66: 569–596.
34. Hofbauer J, Sigmund K (1998) *Evolutionary Games and Population Dynamics*. Cambridge: Cambridge University Press.
35. Geritz SAH, Kisdi E, Meszena G, Metz JAJ (1998) Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evolutionary Ecology* 12: 35–57.
36. Dreber A, Rand DG, Fudenberg D, Nowak MA (2008) Winners don't punish. *Nature* 452: 348–351.
37. Wu J, Zhang B, Zhou Z, He Q, Zheng X, et al. (2009) Costly punishment does not always increase cooperation. *Proceedings of the National Academy of Sciences* 106: 17448–17451.
38. Charness G, Rabin M (2002) Understanding social preferences with simple tests. *Quarterly Journal of Economics* 117: 817–869.
39. Falk A, Fischbacher U (2006) A theory of reciprocity. *Games and Economic Behavior* 54: 293–315.
40. Levine DK (1998) Modeling altruism and spitefulness in experiments. *Review of Economic Dynamics* 1: 593–622.
41. Pillutla MM, Murnighan JK (1996) Unfairness, anger, and spite: Emotional rejections of ultimatum offers. *Organizational Behavior and Human Decision Processes* 68: 208–224.
42. Sanfey AG, Rilling JK, Aronson JA, Nystrom LE, Cohen JD (2003) The neural basis of economic decision-making in the Ultimatum Game. *Science* 300: 1755–1758.
43. Stirlin WC, Feli T (2013) Game theory, conditional preferences, and social influence. *PLoS ONE* 8: e56751.