PHIL 408Q/PHPE 308D
Fairness

Eric Pacuit, University of Maryland

February 15, 2024

Local Interaction

J. McKenzie Alexander and B. Skyrms (1999). *Bargaining with Neighbors: Is Justice Contagious*. Journal of Philosophy, 96(11).

Two principles of distributional justice:

Optimality: a distribution is not just if, under an alternative distribution, all recipients would be better off.

Equity: if the position of the recipients is symmetric, then the distribution should be symmetric. That is to say, it does not vary when we switch the recipients.

If you ask people to judge the just distribution, their answers show that optimality and equity are powerful operative principles.

Menachem Yaari and Maya Bar-Hillel (1981). *On Dividing Justly*. Social Choice and Welfare, I, pp. 1-24.

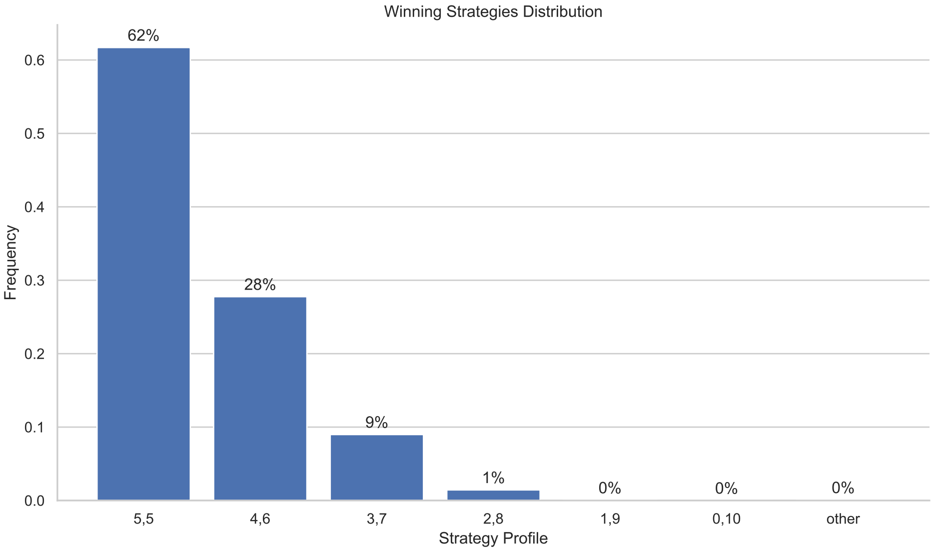
Divide-the-Dollar Game

Two players are faced with a windfall of \$10. They each can demand any integer between 0 and 10 (i.e., $0 \leq D \leq 10$). If the sum of the two demands is less than or equal to 10, then each receives what they demand. Otherwise, the each receive 0.

Divide-the-Dollar Game

Two players are faced with a windfall of \$10. They each can demand any integer between 0 and 10 (i.e., $0 \leq D \leq 10$). If the sum of the two demands is less than or equal to 10, then each receives what they demand. Otherwise, the each receive 0.

Evolutionary game theory (reading 'evolution' as cultural evolution) promises an explanation, but the promise is only partially fulfilled. Demand-half is the only evolutionarily stable strategy in divide-the-dollar. It is the only strategy such that, if the whole population played that strategy, no small group of innovators, or "mutants," playing a different strategy could achieve an average payoff at least as great as the natives. If we could be sure that this unique evolutionarily stable strategy would always take over the population, the problem would be solved.


Winning Strategies Distribution
62%
28%
9%
1%
0%
0%
0%
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Frequency
5,5
4,6
3,7
2,8
1,9
0,10
other
Strategy Profile

Fair Division	62,209
4-6 Polymorphism	27,469
3-7 Polymorphism	8,801
2-7 polymorphism	1,483
1-9 Polymorphism	38
0-10 Polymorphism	0

Table 1: Convergence results for replicator dynamics - 100,000 trials

The projected evolutionary explanation seems to fall somewhat short. The best we might say on the basis of pure replicator dynamics is that fixation of fair division is more likely than not, and that polymorphisms far from fair division are quite unlikely.

The projected evolutionary explanation seems to fall somewhat short. The best we might say on the basis of pure replicator dynamics is that fixation of fair division is more likely than not, and that polymorphisms far from fair division are quite unlikely.

Two solutions

- Inject some probability: Every once and a while a member of the population just picks a strategy at random and tries it out perhaps as an experiment, perhaps just as a mistake.
- Add correlation of players with the same strategy: There is a higher probability of playing the game with players of the same strategy.

Injecting some Probability

Suppose that we are at a 4-6 polymorphic equilibrium.

Injecting some Probability

Suppose that we are at a 4-6 polymorphic equilibrium.

If there is some fixed probability of an experiment (or mistake), and if experiments are independent, and if we wait long enough, there will be enough experiments of the right kind to kick the population out of the basin of attraction of the 4-6 polymorphism and into the basin of attraction of fair division and the evolutionary dynamics will carry fair division to the 5-5 split.

Injecting some Probability

Suppose that we are at a 4-6 polymorphic equilibrium.

If there is some fixed probability of an experiment (or mistake), and if experiments are independent, and if we wait long enough, there will be enough experiments of the right kind to kick the population out of the basin of attraction of the 4-6 polymorphism and into the basin of attraction of fair division and the evolutionary dynamics will carry fair division to the 5-5 split.

Eventually, experiments or mistakes will kick the population out of the basin of attraction of fair division, but we should expect to wait much longer for this to happen.

In the long run, the system will spend most of its time in the fair-division equilibrium.

In the long run, the system will spend most of its time in the fair-division equilibrium.

Peyton Young showed that, if we take the limit as the probability of someone experimenting gets smaller and smaller, the ratio of time spent in fair division approaches one. In his terminology, fair division is the stochastically stable equilibrium of this bargaining game.

H. Peyton Young (1993). *An Evolutionary Model of Bargaining*. Journal of Economic Theory, 59(1), pp. 145-168.

This explanation gets us a probability arbitrarily close to one of finding a fair-division equilibrium if we are willing to wait an arbitrarily long time.

This explanation gets us a probability arbitrarily close to one of finding a fair-division equilibrium if we are willing to wait an arbitrarily long time.

But one may well be dissatisfied with an explanation that lives at infinity.

This explanation gets us a probability arbitrarily close to one of finding a fair-division equilibrium if we are willing to wait an arbitrarily long time.

But one may well be dissatisfied with an explanation that lives at infinity.

Putting the limiting analysis to one side, pick some plausible probability of experimentation or mistake and ask yourself how long you would expect it to take in a population of 10,000, for 1,334 demand-6 types simultaneously to try out being demand-5 types and thus kick the population out of the basin of attraction of the 4-6 polymorphism and into the basin of attraction of fair division.

This explanation gets us a probability arbitrarily close to one of finding a fair-division equilibrium if we are willing to wait an arbitrarily long time.

But one may well be dissatisfied with an explanation that lives at infinity.

Putting the limiting analysis to one side, pick some plausible probability of experimentation or mistake and ask yourself how long you would expect it to take in a population of 10,000, for 1,334 demand-6 types simultaneously to try out being demand-5 types and thus kick the population out of the basin of attraction of the 4-6 polymorphism and into the basin of attraction of fair division.

The evolutionary explanation still seems less than compelling.

Correlated Strategies

Let us...replace the assumption of random encounters with one of positive correlation between like strategies.

Correlated Strategies

Let us...replace the assumption of random encounters with one of positive correlation between like strategies.

It is evident that in the extreme case of perfect correlation, stable polymorphisms are no longer possible. Strategies that demand more than 1/2 meet each other and get nothing. Strategies that demand less than 1/2 meet each other and get what they demand. The fittest strategy is that which demands exactly 1/2 of the cake....

Correlated Strategies

Let us...replace the assumption of random encounters with one of positive correlation between like strategies.

It is evident that in the extreme case of perfect correlation, stable polymorphisms are no longer possible. Strategies that demand more than 1/2 meet each other and get nothing. Strategies that demand less than 1/2 meet each other and get what they demand. The fittest strategy is that which demands exactly 1/2 of the cake....

In the real world, both random meeting and perfect correlation are likely to be unrealistic assumptions. The real cases of interest lie in between.

(Skyrms, p. 18)

Example: Correlating Strategies

Let $0 \leq \epsilon \leq 1$ be a **level of correlation**.

- $Pr_t(Modest \mid Modest)$ is the proportion playing *Modest* against *Modest*
  $Pr_t(Modest \mid Modest) = Pr_t(Modest) + \epsilon * (1 - Pr_t(Modest))$

Example: Correlating Strategies

Let $0 \leq \epsilon \leq 1$ be a **level of correlation**.

- $Pr_t(Modest \mid Modest)$ is the proportion playing *Modest* against *Modest*
  $Pr_t(Modest \mid Modest) = Pr_t(Modest) + \epsilon * (1 - Pr_t(Modest))$

- $Pr_t(Fair \mid Modest)$ is the proportion playing *Fair* against *Modest*
  $Pr_t(Fair \mid Modest) = Pr_t(Fair) - \epsilon * (Pr_t(Fair))$

Example: Correlating Strategies

Let $0 \leq \epsilon \leq 1$ be a **level of correlation**.

- $Pr_t(Modest \mid Modest)$ is the proportion playing *Modest* against *Modest*
  $Pr_t(Modest \mid Modest) = Pr_t(Modest) + \epsilon * (1 - Pr_t(Modest))$

- $Pr_t(Fair \mid Modest)$ is the proportion playing *Fair* against *Modest*
  $Pr_t(Fair \mid Modest) = Pr_t(Fair) - \epsilon * (Pr_t(Fair))$

- $Pr_t(Greedy \mid Modest)$ is the proportion playing *Greedy* against *Modest*
  $Pr_t(Greedy \mid Modest) = Pr_t(Greedy) - \epsilon * (Pr_t(Greedy))$

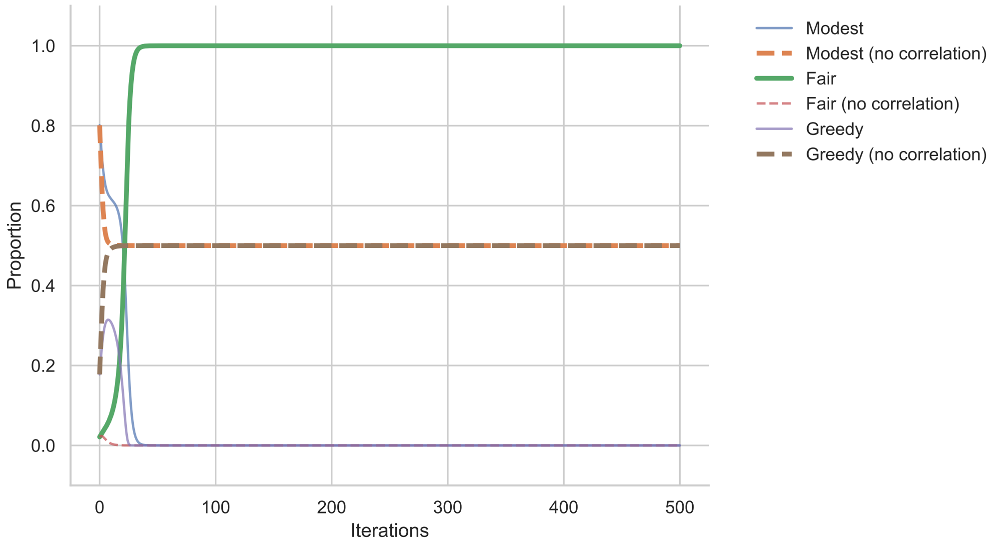
Example: Correlating Strategies

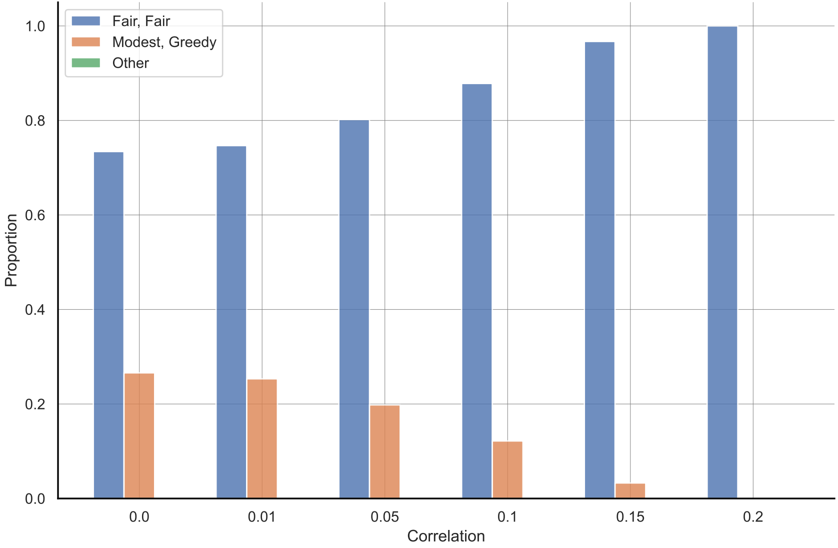
$$f\_Modest_t = Pr_t(Modest) * \frac{1}{3} + Pr_t(Fair) * \frac{1}{3} + Pr_t(Greedy) * \frac{1}{3}$$

Example: Correlating Strategies

$$f\_Modest_t = Pr_t(Modest) * \frac{1}{3} + Pr_t(Fair) * \frac{1}{3} + Pr_t(Greedy) * \frac{1}{3}$$

$$Pr_t(Modest \mid Modest) * \frac{1}{3} + Pr_t(Fair \mid Modest) * \frac{1}{3} + Pr_t(Greedy \mid Modest) * \frac{1}{3}$$


1.0
0.8
0.6
0.4
0.2
0.0
Proportion
0
100
200
300
400
500
Iterations
Modest
Modest (no correlation)
Fair
Fair (no correlation)
Greedy
Greedy (no correlation)


Fair, Fair
Modest, Greedy
Other
1.0
0.8
0.6
0.4
0.2
0.0
Proportion
0.0
0.01
0.05
0.1
0.15
0.2
Correlation

What is the justification for adding a correlation factor, though? Once Skyrms relaxes the requirement of random interactions in the population, and allows some degree of assortative interactions, we need to hear a justification for assuming that the likely departure from random interactions will be toward correlation in particular. Why think that individuals are especially likely to meet others playing the same strategy as they play? (D'Arms, Batterman, and Gorny, p. 92)

Justin D'Arms, Robert Batterman, and Krzyzstof Gorny (1998). *Game Theoretic Explanations and the Evolution of Justice*. Philosophy of Science, 65(1), pp. 76-102.

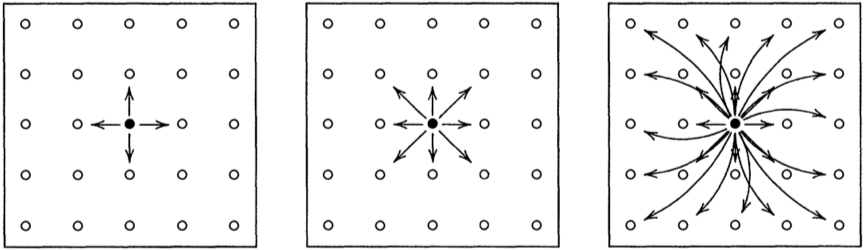
J. McKenzie Alexander and B. Skyrms (1999). *Bargaining with Neighbors: Is Justice Contagious*. Journal of Philosophy, 96(11), pp. 588 - 598.

J. McKenzie Alexander (2000). *Evolutionary Explanations of Distributive Justice*. Philosophy of Science, 67(3), pp. 490 - 516.

The dynamics is driven by imitation. Individuals imitate the most successful person in the neighborhood. A generation an iteration of the discrete dynamics has two stages:

1. Each individual plays the Nash bargaining game with each of her neighbors using her current strategy. Summing the payoffs gives her current success level.
2. Each player looks around her neighborhood and changes her current strategy by imitating her most successful neighbor, providing that her most successful neighbor is more successful than she is; otherwise, she does not switch strategies. (Ties are broken by a coin flip.)

Neighborhoods



(a) von Neumann (b) Moore (8) (c) Moore (24)

Figure 1. Three common neighborhoods defined on a square lattice.

Dynamics

1. Imitate the best neighbor: Each player looks at her neighbors and adopts the strategy of the neighbor who did the best, where "best" means "earned the highest score."

Dynamics

1. Imitate the best neighbor: Each player looks at her neighbors and adopts the strategy of the neighbor who did the best, where "best" means "earned the highest score."

2. Imitate with probability proportional to success: Assigns to every neighbor q who did better than the player p a nonzero probability that p will adopt q's strategy.

Dynamics

1. Imitate the best neighbor: Each player looks at her neighbors and adopts the strategy of the neighbor who did the best, where "best" means "earned the highest score."

2. Imitate with probability proportional to success: Assigns to every neighbor q who did better than the player p a nonzero probability that p will adopt q's strategy.

3. Imitate best average payoff: Calculate the average payoff of each strategy in their neighborhood and select the one with the highest value.

Dynamics

1. Imitate the best neighbor: Each player looks at her neighbors and adopts the strategy of the neighbor who did the best, where "best" means "earned the highest score."

2. Imitate with probability proportional to success: Assigns to every neighbor q who did better than the player p a nonzero probability that p will adopt q's strategy.

3. Imitate best average payoff: Calculate the average payoff of each strategy in their neighborhood and select the one with the highest value.

The general question of how one's choice of the update rule affects the limit form of the model remains an open and difficult problem.

	Bargaining with Neighbors		Bargaining with Strangers	
	A	*B*	*C*	*D*
0-10	0	0	0	0
1-9	0	0	0	0
2-8	0	0	54	57
3-7	0	0	550	556
4-6	26	26	2560	2418
fair	9972	9973	6833	6964

Table 2: Convergence results for five series of 10,000 trials

Sometimes we bargain with neighbors, sometimes with strangers. The dynamics of the two sorts of interaction are quite different.

Sometimes we bargain with neighbors, sometimes with strangers. The dynamics of the two sorts of interaction are quite different.

Bargaining with neighbors almost always converges to fair division and convergence is remarkably rapid.

Sometimes we bargain with neighbors, sometimes with strangers. The dynamics of the two sorts of interaction are quite different.

Bargaining with neighbors almost always converges to fair division and convergence is remarkably rapid.

Both bargaining with strangers and bargaining with neighbors are artificial abstractions. In initial phases of human cultural evolution, bargaining with neighbors may be a closer approximation to the actual situation than bargaining with strangers. The dynamics of bargaining with neighbors strengthens the evolutionary explanation of the norm of fair division.