BARGAINING WITH NEIGHBORS: IS JUSTICE CONTAGIOUS?

What is justice? The question is harder to answer in some cases than in others. We focus on the easiest case of distributive justice. Two individuals are to decide how to distribute a windfall of a certain amount of money. Neither is especially entitled, or especially needy, or especially anything—their positions are entirely symmetric. Their utilities derived from the distribution may be taken, for all intents and purposes, simply as the amount of money received. If they cannot decide, the money remains undistributed and neither gets any. The essence of the situation is captured in the simplest version of a bargaining game devised by John Nash.[1] Each person decides on a bottom-line demand. If those demands do not jointly exceed the windfall, then each person gets his demand; if not, no one gets anything. This game is often simply called *divide-the-dollar.*

In the ideal simple case, the question of distributive justice can be decided by two principles:

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.

Since we stipulate that the position of the two individuals is symmetric, equity requires that the just distribution must give them the same amount of money. Optimality then rules out such unlikely schemes as giving each one dime and throwing the rest away—each must get half the money.

There is nothing new about our two principles. Equity is the simplest consequence of the theory of distributive justice in Aristotle's *Politics.* It is a consequence of Immanuel Kant's categorical imperative. Utilitarians tend to stress optimality, but are not completely insensitive to equity. Optimality and equity are the two most uncontroversial requirements in Nash's axiomatic treatment of bargaining. If you ask people to judge the just distribution, their answers show that optimality and equity are powerful operative principles.[2] So, although nothing much hangs on it, we shall feel

[1] "The Bargaining Problem," *Econometrica,* XVIII (1950): 155-62.

[2] Menachem Yaari and Maya Bar-Hillel, "On Dividing Justly," *Social Choice and Welfare,* I (1981): 1-24.

0022-362X/99/9611/588–98

free to use moral language and to call the equal split *fair division* in divide-the-dollar.

I. RATIONALITY, BEHAVIOR, EVOLUTION

Two rational agents play the divide-the-dollar game. Their rationality is common knowledge. What do they do? The answer that game theory gives us is that *any* combination of demands is compatible with these assumptions. For example, Jack may demand ninety percent thinking that Jill will only demand ten percent on the assumption that Jill thinks that Jack will demand ninety percent and so forth, while Jill demands seventy-five percent thinking that Jack will demand twenty-five percent on the assumption that Jack thinks that Jill will demand seventy-five percent and so forth. *Any* pair of demands is *rationalizable,* in that it can be supported by a hierarchy of conjectures for each player, compatible with common knowledge of rationality. In the example given, these conjectures are quite mistaken.

Suppose we add the assumption that each agent somehow knows what the other will demand. Then any combination of demands that total the whole sum to be divided is still possible. For example, suppose that Jack demands ninety percent knowing that Jill will demand ten percent and Jill demands ten percent knowing that Jack will demand ninety percent. Then each player is maximizing payoff given the demand of the other. That is to say that this is a Nash equilibrium of divide-the-dollar. If the dollar is infinitely divisible, then there are an infinite number of such equilibria.

If experimental game theorists have people actually play divide-the-dollar, they *always* split equally.[3] This is not always true in more complicated bargaining experiments where there are salient asymmetries, but it is true in divide-the-dollar. Rational-choice theory has no explanation of this phenomenon. It appears that the experimental subjects are using norms of justice to select a particular Nash equilibrium of the game. But what account can we give for the existence of these norms?

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

[3] Rudy V. Nydegger and Guillermo Owen, "Two-Person Bargaining: An Experimental Test of the Nash Axioms," *International Journal of Game Theory*, III (1974): 239-50; Alvin Roth and Michael Malouf, "Game Theoretic Models and the Role of Information in Bargaining," *Psychological Review*, LXXXVI (1979): 574-94; John Van Huyck, Raymond Batallio, Sondip Mathur, Patsy Van Huyck, and Andreas Ortmann, "On the Origin of Convention: Evidence From Symmetric Bargaining Games," *International Journal of Game Theory*, XXIV (1995): 187-212.

divide-the-dollar.[4] 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.

But we cannot be sure that this will happen. There are states of the population which are evolutionarily stable where some fraction of the population makes one demand and some fraction makes another. The state where half the population demands one third and half the population demands two thirds is such an evolutionarily stable polymorphism of the population. So is the state where two thirds of the population demands forty percent and one third of the population demands sixty percent We can think of these as pitfalls along the evolutionary road to justice.

How important are these polymorphisms? To what extent do they compromise the evolutionary explanation of the egalitarian norm? We cannot begin to answer these questions without explicitly modeling the evolutionary dynamics and investigating the size of their basins of attraction.

II. BARGAINING WITH STRANGERS

The most widely studied dynamic evolutionary model is a model of interactions with strangers. Suppose that individuals are paired at random from a very large population to play the bargaining game. We assume that the probability of meeting a strategy can be taken as the proportion of the population that has that strategy. The population proportions evolve according to the replicator dynamics. The proportion of the population using a strategy in the next generation is the proportion playing that strategy in the current generation mutiplied by a *fitness factor*. This fitness factor is just the ratio of the average payoff to this strategy to the average payoff in the whole population.[5] Strategies that do better than average grow; those which do worse than average shrink. This dynamic arose in biology

[4] Robert Sugden, *The Economics of Rights, Cooperation, and Welfare* (New York: Blackwell, 1986).

[5] This is the discrete time version of the replicator dynamics, which is most relevant in comparison to the alternative *bargaining-with-neighbors* dynamics considered here. There is also a continuous time version. As comprehensive references, see Josef Hofbauer and Karl Sigmund, *The Theory of Evolution and Dynamical Systems* (New York: Cambridge, 1988); Jorgen W. Weibull, *Evolutionary Game Theory* (Cambridge: MIT, 1995); Larry Samuelson, *Evolutionary Games and Equilibrium Selection* (Cambridge: MIT, 1997).

as a model of asexual reproduction, but more to the point here, it also has a cultural evolutionary interpretation where strategies are imitated in proportion to their success.[6]

The basins of attraction of these polymorphic pitfalls are not negligible. A realistic version of divide-the-dollar will have some finite number of strategies instead of the infinite number that we get from the idealization of infinite divisibility. For a finite number of strategies, the size of a basin of attraction of a population state makes straightforward sense. It can be estimated by computer simulations. We can consider coarse-grained or fine-grained versions of divide-the-dollar; we can divide a stack of quarters, or of dimes, or of pennies. Some results of simulations persist across a range of different granularities. Equal division always has the largest basin of attraction and it is always greater than the basins of attractions of all the polymorphic pitfalls combined. If you choose an initial population state at random, it is more probable than not that the replicator dynamics will converge to a state of fixation of demand-half. Simulation results range between fifty-seven and sixty-three percent of the initial points going to fair division. The next largest basin of attraction is always that closest to the equal split: for example, the four-six polymorphism in the case of dividing a stack of ten dimes and the forty-nine/fifty-one polymorphism in the case of dividing a stack of one-hundred pennies. The rest of the polymorphic equilibria follow the general rule—the closer to fair division, the larger the basin of attraction.

For example, the results running the discrete replicator dynamics to convergence and repeating the process 100,000 times on the game of dividing ten dimes are given in table 1.

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

[6] Jonas Björnerstedt and Jörgen Weibull, "Nash Equilibrium and Evolution by Imitation," in Kenneth J. Arrow et alia, eds., *The Rational Foundations of Economic Behavior* (New York: Macmillan, 1996), pp. 155-71; Karl Schlag, "Why Imitate, and If So How?" Discussion Paper B-361 (University of Bonn, Germany, 1996).

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.

We can say something more if we inject a little bit of probability into the model. Suppose that 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. Suppose we are at a polymorphic equilibrium—for instance, the four-six equilibrium in the problem of dividing ten dimes. 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 four-six polymorphism and into the basin of attraction of fair division and the evolutionary dynamics will carry fair division to fixation. 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. Peyton Young[7] 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.

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-six types simultaneously to try out being demand-five types and thus kick the population out of the basin of attraction of the four-six polymorphism and into the basin of attraction of fair division.[8]) The evolutionary explanation still seems less than compelling.

[7] "An Evolutionary Model of Bargaining," *Journal of Economic Theory*, LIX (1993): 145-68, and "The Evolution of Conventions," *Econometrica*, LXI (1993): 57-94; and Dean Foster and Young, "Stochastic Evolutionary Game Dynamics," *Theoretical Population Biology*, XXXVIII (1990): 219-32.

[8] For discussion of expected waiting times, see Glenn Ellison, "Learning, Local Interaction and Coordination," *Econometrica*, LXI (1993): 1047-71; and Robert Axtell, Joshua M. Epstein, and H. Peyton Young, "The Emergence of Economic Classes in an Agent-Based Bargaining Model," preprint (Brookings Institution, 1999).

III. BARGAINING WITH NEIGHBORS

The model of random encounters in an infinite population that motivates the replicator dynamics may not be the right model. Suppose interactions are with neighbors. Some investigations of cellular automaton models of prisoner's dilemma and a few other games show that interactions with neighbors may produce dynamical behavior quite different from that generated by interactions with strangers.[9] Bargaining games with neighbors have not, to the best of our knowledge, previously been studied.

Here, we investigate a population of 10,000 arranged on a one hundred by one hundred square lattice. As the neighbors of an individual in the interior of the lattice, we take the eight individuals to the N, NE, E, SE, S, SW, W, NW. This is called the Moore(8) neighborhood in the cellular automaton literature.[10] 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. First, each individual plays the divide-ten-dimes game with each of her neighbors using her current strategy. Summing the payoffs gives her current success level. Then 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.)

In initial trials of this model, fair division *always* went to fixation. This cannot be a universal law, since you can design "rigged" configurations where a few demand-one-half players are, for example, placed in a population of demand-four and demand-six players with the latter so arranged that there is a demand-six type who is the most successful player in the neighborhood of every demand-one-half player. Start enough simulations at random starting points and sooner or later you will start at one of these.

We ran a large simulation starting repeatedly at randomly chosen starting points. Fair division went to fixation in more than ninety-nine point five percent of the trials. The cases where it did not were all cases where the initial population of 10,000 contained fewer than seventeen

[9] Gregory B. Pollack, "Evolutionary Stability on a Viscous Lattice," *Social Networks*, XI (1989): 175-212; Martin A. Nowak and Robert M. May, "Evolutionary Games and Spatial Chaos," *Nature*, CCCLIX (1992): 826-29; Kristian Lindgren and Mats Nordahl, "Evolutionary Dynamics in Spatial Games," *Physica D*, LXXV (1994): 292-309; Luca Anderlini and Antonella Ianni, "Learning on a Torus," in Cristina Bicchieri, Richard Jeffrey, and Brian Skryms, eds., *The Dynamics of Norms* (New York: Cambridge, 1997), pp. 87-107.

[10] We find that behavior is not much different if we use the von Neumann neighborhood: N, S, E, W, or a larger Moore neighborhood.

demand-one-half players. Furthermore, convergence was remarkably quick. Mean time to fixation of fair division was about sixteen generations. This may be compared with a mean time to convergence[11] in discrete replicator dynamics of forty-six generations, and with the ultra-long-run character of stochastically stable equilibrium.

It is possible to exclude fair division from the possible initial strategies in the divide-ten-dimes game and start at random starting points that include the rest. If we do this, all strategies other than demand-four dimes and demand-six dimes are eliminated and the four-six polymorphic population falls into a "blinking" cycle of period two. If we then turn on a little bit of random experimentation or "mutation" allowing the possibility of demand-five, we find that as soon as a very small clump of demand-five players arises, it systematically grows until it takes over the whole population—as illustrated in figure 1. *Justice is contagious.*[12]

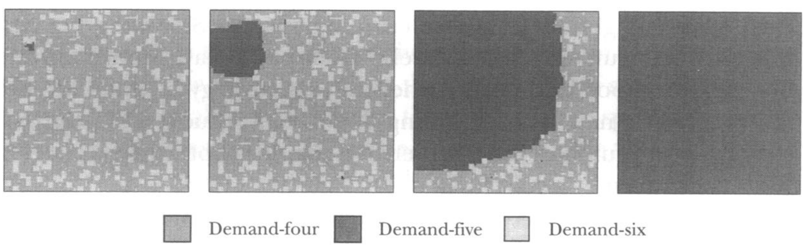

Figure 1: The steady advance of fair division

IV. ROBUSTNESS

The bargaining-with-neighbors model of the last section differs from the bargaining with strangers model in more than one way. Might the difference in behavior that we have just described be due to the imitate-the-most-successful dynamics rather than the neighbor effect? To answer this question, we ran simulations varying these factors independently.

We consider both fixed and random neighborhoods. The models using fixed neighborhoods use the Moore (8) neighborhood described above. In the alternative random-neighborhood model, each generation a new set of "neighbors" is chosen at random from

[11] At .9999 level to keep things comparable.

[12] Ellison (*op. cit.*) found such contagion effects in local interaction of players arranged on a circle and playing pure coordination games.

the population for each individual. That is to say, these are neighborhoods of *strangers*.

We investigated two alternative dynamics. One imitates the most successful neighbor as in our bargaining-with-neighbors model. The other tempers the all-or-nothing character of imitate-the-best. Under it, an individual imitates one of the strategies in its neighborhood that is more successful than it (if there are any) with relative probability proportional to their success in the neighborhood. This is a move in the direction of the replicator dynamics.

	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

In table 2, *A* and *B* are bargaining with neighbors, with imitate-the-best-neighbor and imitate-with-probability-proportional-to-success dynamics, respectively. The results are barely distinguishable. *C* and *D* are the random-neighborhood models corresponding to *A* and *B*, respectively. These results are much closer to those given for the replicator dynamics in table 1. The dramatic difference in convergence to fair division between our two models is due to the structure of interaction with neighbors.

V. ANALYSIS

Why is justice contagious? A strategy is contagious if an initial "patch" of that strategy will extend to larger and larger patches. The key to contagion of a strategy is interaction along the edges of the patch, since in the interior the strategy can only imitate itself.[13]

[13] For this reason, "frontier advantage" is used to define an unbeatable strategy in Illan Eshel, Emilia Sansone, and Avner Shaked, "Evolutionary Dynamics of Populations with a Local Interaction Structure," working paper (University of Bonn, 1996).

Consider an edge with demand-five players on one side, and players playing the complementary strategies of one of the polymorphisms on the other. Since the second rank of demand-five players always meet their own kind, they each get a total payoff of forty from their eight neighbors. Players in the first rank will therefore imitate them unless a neighbor from the polymorphism gets a higher payoff. The low strategy in a polymorphic pair cannot get a higher payoff. So if demand-five is to be replaced at all, it must be by the high strategy of one of the polymorphic pairs.

In the four-six polymorphism—the polymorphism with the greatest basin of attraction in the replicator dynamics—this simply cannot happen, even in the most favorable circumstances. Suppose that we have someone playing demand-six in the first rank of the polymorphism, surrounded on his own side by compatible demand-four players to boost his payoff to the maximum possible.[14] Since he is in the first rank, he faces three incompatible demand-five neighbors. He has a total payoff of thirty while his demand-five neighbors have a total payoff of thirty-five. Demand-five begins an inexorable march forward as illustrated in figure 2. (The pattern is assumed to extend in all directions for the computation of payoffs of players at the periphery of what is shown in the figure.)

Initial		Iteration 1
5544		5554
5544		5554
5564	=>	5554
5544		5554
5544		5554

Figure 2: fair division versus four-six polymorphism

[14] In situating the high strategy of the polymorphic pair in a sea of low-strategy players, we are creating the best-case scenario for the advancement of the polymorphism into the patch of demand-five players.

If we choose a polymorphism that is more extreme, however, it is possible for the high strategy to replace some demand-five players for a while. Consider the one-nine polymorphism, with a front line demand-nine player backed by compatible demand-one neighbors. The demand-nine player gets a total payoff of forty-five—more than anyone else—and thus is imitated by all his neighbors. This is shown in the first transition in figure 3.

Initial		Iteration 1		Iteration 2		Iteration 3
55111		55511		55551		55555
55111		55511		55559		55559
55111	=>	59991	=>	55999	=>	55599
55911		59991		55999		55599
55111		59991		55999		55599
55111		55511		55559		55559
55111		55511		55551		55555

Figure 3: fair division versus one-nine polymorphism

But the success of the demand-nine strategy is its own undoing. In a cluster of demand-nine strategies, it meets itself too often and does not do so well. In the second transition, demand-five has more than regained its lost territory, and in the third transition it has solidly advanced into one-nine territory.

Analysis of the interaction along an edge between demand-five and other polymorphisms is similar to one of the cases analyzed here.[15] Either the polymorphism cannot advance at all, or the advance creates the conditions for its immediate reversal. A complete analysis of this complex system is something that we cannot offer. But the foregoing does offer some analytic insight into the contagious dynamics of equal division in "bargaining with neighbors."

VI. CONCLUSION

Sometimes we bargain with neighbors, sometimes with strangers. The dynamics of the two sorts of interaction are quite different. In the bargaining game considered here, bargaining with strangers—modeled by the replicator dynamics—leads to fair division from a randomly chosen starting point about sixty percent of the time. Fair division becomes the unique answer in bargaining with strangers if we change the ques-

[15] With some minor complications involving ties.

tion to that of stochastic stability in the ultra-long-run. But long expected waiting times call the explanatory significance of the stochastic stability result into question.

Bargaining with neighbors almost always converges to fair division and convergence is remarkably rapid. In bargaining with neighbors, the local interaction generates clusters of those strategies which are locally successful. Clustering and local interaction together produce positive correlation between like strategies. As noted elsewhere,[16] positive correlation favors fair division over the polymorphisms. In bargaining with neighbors, this positive correlation is not something externally imposed but rather an unavoidable consequence of the dynamics of local interaction. As a consequence, once a small group demand-half players is formed, justice becomes contagious and rapidly takes over the entire population.

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.

JASON ALEXANDER
BRIAN SKYRMS

University of California/Irvine

[16] Skyrms, "Sex and Justice," this JOURNAL, XCI, 6 (June 1994): 305-20 and *Evolution of the Social Contract* (New York: Cambridge, 1996).