Evolution of fairness in the one-shot anonymous Ultimatum Game

David G. Rand, Corina E. Tarnita, Hisashi Ohtsuki, and Martin A. Nowak

Classical economic models assume that people are fully rational and selfish, while experiments often point to different conclusions. A canonical example is the Ultimatum Game: one player proposes a division of a sum of money between herself and a second player, who either accepts or rejects. Based on rational self-interest, responders should accept any nonzero offer and proposers should offer the smallest possible amount. Traditional, deterministic models of evolutionary game theory agree: in the one-shot anonymous Ultimatum Game, natural selection favors low offers and demands. Experiments instead show a preference for fairness: often responders reject low offers and proposers make higher offers than needed to avoid rejection. Here we show that using stochastic evolutionary game theory, where agents make mistakes when judging the payoffs and strategies of others, natural selection favors fairness. Across a range of parameters, the average strategy matches the observed behavior: proposers offer between 30% and 50%, and responders demand between 25% and 40%. Rejecting low offers increases relative payoff in pairwise competition between two strategies and is favored when selection is sufficiently weak. Offering more than you demand increases payoff when many strategies are present simultaneously and is favored when mutation is sufficiently high. We also perform a behavioral experiment and find empirical support for these theoretical findings: uncertainty about the success of others is associated with higher demands and offers; and inconsistency in the behavior of others is associated with higher offers but not predictive of demands. In an uncertain world, fairness finishes first.

cooperation | prosociality | stochastic dynamics

Game theorists traditionally assume that people act fully rationally to maximize their own financial gains. A wealth of behavioral data, however, has demonstrated that many people are influenced by the payoffs of others, exhibiting so-called “other-regarding preferences” (1). The Ultimatum Game (UG) has been a particularly influential example of this phenomenon (2). In the UG, two players have to divide a certain sum of money between them. One player (the proposer) makes an offer. The other player (the responder) can either accept the offer, in which case each receives the money as proposed, or reject the offer, in which case neither player receives anything. In a one-shot anonymous UG, a rational self-interested proposer will offer the minimum amount that she believes will be acceptable to the responder, a rational self-interested responder will accept any nonzero offer. Thus, under common knowledge of the rationality of both players, the subgame perfect Nash equilibrium is for the proposer to make the minimum possible offer, and for the responder to accept it (2).

To evaluate these predictions, many behavioral experiments have been conducted using the UG (1–8). Although there is considerable quantitative variation across studies, two clear qualitative deviations from rational self-interest are robustly observed: (i) many responders choose to reject low (but nonzero) offers, and (ii) many proposers offer more than the minimum amount required to avoid rejection. One popular explanation of both of these findings is that people are motivated by a sense of fairness (or “inequity aversion”): Subjects prefer both players to receive equal payoffs, and are willing to pay a price to create more equitable outcomes (9). By this argument, responders who reject low offers incur a cost to avoid getting a smaller payoff than the proposer (disadvantageous inequity), and proposers who offer more than needed to avoid rejection incur a cost to avoid receiving a larger payoff than the responder (advantageous inequity). Additional evidence of this psychological principle is demonstrated by an experiment where subjects will pay to alter randomly assigned payoffs of others to induce greater equality (10).

Furthermore, it is typically observed that people are more averse to disadvantageous inequity than they are to advantageous inequity (9), and research with children finds that disadvantageous inequity develops earlier than advantageous inequity (11–14). These results suggest that the two forms of fairness are most likely cognitively distinct. Some have argued that proposer behavior can be entirely explained by strategic motivations: given that many responders reject low offers, it may be payoff maximizing to offer even splits (15, 16). Others, however, contend that fairness concerns play at least some part in the high offers of proposers (17, 18); a comprehensive review (1) concludes that high proposer offers are likely the result of a combination of strategic and fairness-based motivations.

Fairness presents a proximate psychological motivation for the observed behavior. What, however, is the ultimate evolutionary explanation for why we should have come to possess such fairness preferences? To explore the origins of fairness, we study an evolutionary process in which strategies with higher payoffs tend to become more common in the population (19–22). This process could describe genetic evolution, or cultural evolution through social learning, both of which have been linked to play in the UG (4, 18, 23). In the context of genetic evolution, agents reproduce and die, and mutations introduce variation into the gene pool. In the context of cultural evolution, individuals sometimes change strategy and copy the strategy of another, which has been linked to play in the UG (4, 18, 23). In the context of genetic evolution, agents reproduce and die, and mutations introduce variation into the gene pool. In the context of cultural evolution, individuals sometimes change strategy and copy the strategy of another, with higher payoff strategies being more likely to be imitated. Here mutation represents either experimentation, in which individuals try new random strategies, or confusion regarding the strategies used by other players (leading the imitator to adopt a strategy different from that of the imitatee) (24). This type of process based on reproduction is distinct from strategy update rules that use prospective reasoning, such as best response dynamics (25).

Author contributions: D.G.R., C.E.T., H.O., and M.A.N. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

D.G.R. and C.E.T. contributed equally to this work.

Author to whom correspondence should be addressed. E-mail: martin_nowak@harvard.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1214167110/-/DCSupplemental.

www.pnas.org/cgi/doi/10.1073/pnas.1214167110
(for a comparison of evolutionary dynamics with prospective reasoning, see ref. 26).

Using this evolutionary framework, we can explore the conditions under which natural selection leads to fair behavior. The classical approach to evolutionary game theory (27–29) is deterministic. Such models assume that higher payoff strategies always become more common, whereas lower payoff strategies always die out. Under deterministic game dynamics, evolution favors self-interest, and in the UG selection leads to the rational self-interested strategy where agents offer and demand nothing (30). To explain fairness using deterministic dynamics, it is therefore necessary to invoke some additional evolutionary mechanism.

One approach involves reputation formation (31–36). It has been shown that fairness can be favored by natural selection if agents can recognize their partners’ strategies (37, 38) or have reputations that carry from game to game (30). Here it pays to reject a stingy offer today so that others will make you higher offers in the future. Without a sufficiently high expectation of future interactions and a sufficiently strong reputation system, however, fairness collapses. An alternate approach studies one-shot anonymous games but assumes an asymmetric mutation structure, such that proposers experiment with new strategies less often than responders: the greater variation in responder behavior structurally, such that proposers experiment with new strategies less frequently, favoring self-interest, and in the UG selection leads to the rational self-interested strategy.

We consider the stochastic evolution of strategies in a population of finite size $N$. Each player $i$ plays the UG with each of the $N−1$ other players, and receives an average payoff $p_i$. Player $i$’s effective payoff (or fecundity) is then defined as $\exp[p_i w]$, where $w$ is the “intensity of selection.” An intuition behind this effective payoff function is that the higher the intensity of selection, the more likely agents with higher payoffs are to be imitated (to reproduce). At the extreme of $w \to \infty$, only those who obtain the highest payoff are imitated (strong selection).

We specify an agent who obtains the highest payoff are imitated (to reproduce). At the extreme of $w \to \infty$, only those who obtain the highest payoff are imitated (strong selection).

The dynamics depends significantly on the mutation rate $u$. In the low mutation limit $u = 0$, a novel mutant will either die out or completely take over the population before a new mutant arises (43, 48–51). Thus, the population transitions between homogeneous states, in which all agents in the population play the same strategy at any given time. Here, strategies which can protect themselves from invasion do best. Conversely, in the high mutation limit $u \to 1$, all strategies are present at approximately equal abundances at the same time (24, 52). Thus, success is determined not by resisting invasion, but by performing best when playing against all strategies with equal probability (i.e., playing a randomly selected opponent). Intermediate mutation rates result in intermediate outcomes between these two dynamical extremes (for technical details, see Methods and Supporting Information).

**Results**

We begin with agent-based simulations using population size $N = 100$, and vary the intensity of selection $w$ as well as the mutation rate $u$. For each set of simulation parameters we determine the steady-state frequency distribution over the $[0,1]$ uniform distribution in $[0,1]$, and later show that the results are qualitatively unchanged when instead we use local mutation.

The dynamics depends significantly on the mutation rate $u$. In the low mutation limit $u \to 0$, a novel mutant will either die out or completely take over the population before a new mutant arises (43, 48–51). Thus, the population transitions between homogeneous states, in which all agents in the population play the same strategy at any given time. Here, strategies which can protect themselves from invasion do best. Conversely, in the high mutation limit $u \to 1$, all strategies are present at approximately equal abundances at the same time (24, 52). Thus, success is determined not by resisting invasion, but by performing best when playing against all strategies with equal probability (i.e., playing a randomly selected opponent). Intermediate mutation rates result in intermediate outcomes between these two dynamical extremes (for technical details, see Methods and Supporting Information).

We now turn from the most common strategy to consider the average (mean) strategy. Fig. 2 shows how the time-averaged values of $p$ and $q$ vary systematically with changes in selection strength $w$ and mutation rate $u$. When selection is very weak, the dynamics is dominated by neutral drift, and mean $p$ and $q$ are both $\sim 0.5$. As the selection intensity increases, both $p$ and $q$ decrease, approaching the rational self-interested strategy $p = q = 0$. Critically, however, $q$ decreases faster than $p$. Thus, in Fig. 2 we observe both aspects of fair behavior, $p > q$ and $q > 0$, across a wide range of parameter values. As with the modal strategy considered in Fig. 1, we see in Fig. 2 that although the selection strength and mutation rate quantitatively affect the mean $p$ and $q$,...
the qualitative result is general (outside of the weak and strong selection extremes): The mean proposal $p$ is greater than the demand $q$, and the average demand $q$ is substantially greater than 0. These findings are very different from the results of classical evolutionary game theory using either the replicator equation studying the mini UG (30) or adaptive dynamics studying the full UG (53), where the population converges to the rational self-interested strategy $p = q = 0$ unless other mechanisms are present.

Our results are robust to the manner in which mutants are selected. Replacing the uniform mutation described above with a local mutation scheme in which mutants are some random perturbation from the parent strategy gives qualitatively similar results. See SI Local Mutation Kernel and Figs. S2 and S3 for details.

A pessimistic interpretation of the results in Fig. 2 is as follows. Perhaps selection always favors the rational self-interested strategy $p = q = 0$, and the fact that the mean $p$ and $q$ transition from 0 to 0.5 as $w$ decreases is the trivial result of increasing neutral drift driving the mean away from the optimal (selfish) strategy. On the contrary, however, we clearly see that this is not the case. Instead, the frequency distributions in Fig. 1 are centered around fair strategies with large offers and demands, as long as selection is not too strong. Put differently, the strategy most favored by natural selection is the strategy that is most common under mutation-selection balance; thus, the fact that $p = q = 0$ is not the most common strategy when selection is not so strong shows that we truly are observing natural selection favoring fairness.

Thus far we have shown that evolution in finite populations can qualitatively reproduce both the negative and positive aspects of fair behavior demonstrated in experiments. Now we ask whether there can also be quantitative agreement between our model and the range of behaviors observed in the experimental data. On the negative side of fairness (disadvantageous inequity aversion), whereas *Homo economicus* would accept any nonzero offer, the mean demand $q$ is substantially greater than zero across experiments, tending to lie in the range $0.2 < q < 0.35$. On the positive side of fairness (advantageous inequity aversion), subjects also offer more than is demanded: across experiments, the average offer $p$ is substantially higher than the average demand, typically in the range $0.3 < p < 0.5$ (see Fig. S4 for mean $p$ and $q$ values from numerous experiments).

We now compare these experimental data with the average values of $p$ and $q$ from our agent-based simulations, across a range of selection strengths and mutation rates. We see that the evolutionary outcomes for a number of parameter combinations are quantitatively consistent with the experimental data, having mean offers $0.3 < p < 0.5$ and mean demands $0.2 < q < 0.35$. The parameter regions which lead to this agreement are highlighted in yellow in Fig. 2. We see that increasing the mutation rate leads to a corresponding increase in the selection strength needed to reproduce the experimental behavior. This
balancing is required to conserve the level of randomness in the system, which is increased by higher mutation and decreased by stronger selection (the opposite is true for the relationship between selection strength and population size, as shown in Fig. S5). We see that with the correct level of randomness, our evolutionary simulations can quantitatively reproduce the range of average behavior observed in experiments. This agreement stands in contrast with classical economic approaches as well as deterministic evolutionary dynamics, and demonstrates the potential power of finite population evolutionary analysis for understanding human behavior.

In addition to average behavior, it is also of interest to consider how the distribution of individual-level behaviors shown in Fig. 1 compares with experimental data (see Fig. S6 for histograms of \( p \) and \( q \) separately, rather than the joint \([p,q]\) distribution shown in Fig. 1). We begin with proposer behavior. Our model produces a unimodal distribution of \( p \) values that drops off sharply when \( p \) rises above 0.5. This result is generally consistent with the findings of behavioral experiments, with the exception of the model having substantially more variation in offers than is typically seen in experiments (i.e., a wider distribution), and including a low but nonzero probability of a rejection for offers above 0.5 (whereas virtually no subjects offer more than 0.5 in most experiments) (1). Turning to responder behavior, our model again produces a fairly broad unimodal distribution with relatively little probability weight above 0.5. It is harder to compare these results with experimental data as few studies provide distributions of minimum acceptable offers, and the few that do are not consistent with each other: both unimodal distributions (2) and bimodal distributions with modes at 0 and 0.5 (3) have been observed. Further exploration of individual-level behavior, both theoretically and experimentally, is an important direction for future study.

**Discussion**

To gain an intuition for the evolutionary success of fairness in our agent based simulations, we turn to mathematical calculations. In the weak selection limit, where the average abundance of all strategies is approximately equal (steady-state \( p \) and \( q \) are uniformly distributed on the unit square), it is possible to analytically determine which strategy is most common. We find that the most frequent strategy depends strongly on the rate at which mutations arise in the population (see SI Intuition on the Role of Mutation and Table S1).

In the high mutation limit, all strategies are present in the population simultaneously with approximately equal frequency. Hence, the optimum strategy is the one that maximizes its expected absolute payoff against a randomly chosen opposing strategy. As has been shown previously (30), it is intuitive that this strategy is \((1/2, 0)\). The offer of \( p = 1/2 \) maximizes the proposer’s expected payoff of \( p(1-p) \) when playing against a randomly chosen opponent; and the demand \( q = 0 \) maximizes expected payoff as responder because any nonzero demand results in lost profit. Thus, in the high mutation limit natural selection favors the first aspect of fair behavior (advantageous inequity aversion), with proposal \( p \) greater than demand \( q \).

In the low mutation limit, the other hand, the population dynamics is very different. A new mutant will either die out or take over the resident population before another mutant arises. Thus, although all strategies are still present at equal frequency in the steady-state distribution when in the limit of weak selection, at most two strategies are ever present in the population at the same time. Therefore, what matters in the low mutation limit is resisting invasion by a single (randomly chosen) other strategy: it is not expected absolute payoff that determines success, but rather expected relative payoff in pairwise competition with a random opponent (52).

What strategy then maximizes expected relative payoff? One can see that \( p = q \) is a logical first requirement for success in the low mutation limit: any strategy that offers more than it demands \((p > q)\) loses in pairwise competition with mutant strategies which offer \( p_m \) such that \( p > p_m > q \). The most common strategy is the mutant which is less generous than the resident, but still has its offers accepted; and any strategy that demands more than it offers \((p < q)\) always rejects its own offers and is outcompeted by all mutant strategies with \( p_m > q_m \). Thus, we focus our attention on the self-consistent set of strategies with \( p = q \). Consider the interaction between a relatively fair strategy \( S_F \) with \( p = q = x \), and a relatively unfair strategy \( S_U \) with \( p = q = x-e \). Both strategies receive the full payoff of 1 when playing against themselves; but, when \( S_U \) meets \( S_F \), surprisingly the more fair strategy receives the higher payoff (providing it offers less than half). When \( S_U \) is the proposer, her offer is accepted, and she earns \( 1-x \) whereas \( S_U \) earns \( x \); when \( S_U \) is the proposer, her lower offer is rejected, and neither player earns anything. Thus, fairer strategies always earn more than less fair strategies when they interact pairwise as long as \( x < 0.5 \). However, when considering expected relative payoff against a random opponent, there is a tradeoff: the more you offer (up to 0.5), the more strategies you outperform, but the smaller your margin of success is in each pairing. This creates two opposing forces resulting from increasing your offer: the decreasing marginal payoff versus the increasing number of strategies you out-perform. These two forces balance out at some intermediate, optimal value of \( x \). We find that this balance is achieved at \( p = q = 1/3 \), and that this result continues to hold when lifting the restriction \( p = q \). Thus, when mutations are rare, it pays to reject nonzero offers; in this case, the most common strategy has the second experimentally observed aspect of fairness (disadvantageous inequity aversion), with \( q > 0 \).

At intermediate mutation rates, the evolutionary dynamic has characteristics of both the fully heterogeneous and fully homogeneous extremes. Because of the somewhat heterogeneous nature of the population, \( p > q \) is favored by selection; and due to the somewhat homogeneous nature of the population, \( q > 0 \) is also favored. We therefore find that for \( Nu > 1 \), the most frequent strategy has \( p = (1 + Nu)/4 + 2 Nu \) and \( q = 1/(2 + Nu) \). As shown in Fig. 3, we see evolution favoring both qualitative attributes of experimentally observed human behavior that were so challenging for classical game theory’s “economic man”: nonzero rejection thresholds, \( q > 0 \), and proposer’s generosity beyond what is necessary to avoid rejection, \( p > q \). The same results hold if instead of one population in which any individual can be either a proposer or a responder in any given game, we consider two separate interacting populations, one of proposers and one of responders. See SI Single–Population Formulation for...
the weak selection limit analysis and SI Two–Population Formulation for the treatment of separate populations.

To summarize, we find that (i) weakening selection increases the favored demand \( q \) (and therefore also the favored offer \( p \)), and (ii) increasing the mutation rate increases the favored offer \( p \) but not demand \( q \). We now use these theoretical findings to generate two testable predictions, and evaluate these predictions by running a behavioral experiment.

The first prediction stems from the result regarding weakening selection. In our model, weaker selection means that agents have a harder time assessing which others have the highest payoffs when choosing whom to imitate. Therefore, we would predict that people who developed their strategies in settings where it was more difficult to assess the successfulness of others would make both larger offers and larger demands.

The second prediction comes from the result regarding increasing mutation. In our model, higher mutation means that agents are more likely to change their strategy at random. Therefore, we would predict that people who developed their strategies in settings where the behavior of others is less consistent would make higher offers, but not higher demands.

To evaluate these two predictions, we conduct an experiment using the online labor market Amazon Mechanical Turk (54–61). We recruit \( n = 140 \) subjects from around the world to play a one-shot anonymous UG. In addition, these subjects are asked “Among those you interact with in daily life, how clear is it which people are more or less successful?” as a measure of the intensity of selection under which they developed their strategy, and “How accurate do you think first impressions are when judging other people?” as a measure of the consistency of others (i.e., the inverse of the mutation/experimentation rate) under which they developed their strategy.

The results validate both of our theoretical predictions. Fig. 4A shows that subjects who report less clarity about the successfulness of others offer more and demand more. Fig. 4B shows that subjects who report less consistency of others offer more but do not demand more. These results are confirmed by statistical analysis using linear regression with robust SEs, including appropriate controls. See Methods for further details.

We have shown that in finite populations, where dynamics are stochastic and evolutionary trajectories are influenced by chance, natural selection favors fairness in the one-shot anonymous UG. Furthermore, we have validated the predictions of our analysis using behavioral experiments. Previous analyses of the UG have focused on situations where selection is strong and higher payoff strategies always produce more offspring (30, 39, 40). In these settings where evolution is deterministic, fairness needs additional mechanisms to evolve. When the role of random chance is included, however, the results are very different. Without any reputation systems (30, 37), asymmetries between proposers and responders (39), or a priori assumptions about other-regarding preferences (9), the self-interested process of evolution can lead to behaviors which defy classic rational self-interest models (26, 45). These stochastic evolutionary simulations for finite populations can quantitatively reproduce the range of human behaviors observed in the laboratory. [This is as true for cross-cultural results on UG play from small-scale societies as it is for play by Western undergraduates, with the exception of the few societies in which offers above 50% are consistently rejected (23).] Thus, we do not necessarily need to invoke additional mechanisms to provide an evolutionary account of the origins of fairness, as long as the system is not too deterministic.

The emotions, intuitions, and preferences which guide us toward generosity and righteous anger in the one-shot anonymous UG (62–65) may be the proximate biological (4) and/or cultural (23) implementations of behaviors which are advantageous in the presence of weak selection and imperfect learning (for further discussion of the role of intuition in economic games, see ref. 66). Thus, stochastic evolutionary dynamics may offer an explanation for why we have come to have such a preference for fairness. When populations are finite and selection is not too strong, evolution can be fickle: Fitter strategies sometimes die out, and less-fit strategies sometimes triumph. However, in this unfair world, myopic self-interest is vanquished whereas fairness triumphs.

Methods

Agent-Based Simulations. Our main results are produced using agent-based simulations. In our simulations, agents interact in a well-mixed population of constant size 100. Each agent has a strategy vector \((p, q)\) specifying that agent’s behavior when acting as proposer (offers \(p\)) and responder (demands \(q\)). Each generation, every agent plays the UG with every other agent, once in the proposer role and once in the responder role, and the resulting payoff \(x_i\) is the average of the payoffs over all 99 pairings.

Then one agent is picked proportional to \(\exp(w_i)\) to reproduce, where \(w_i\) is the intensity of selection; and one agent is picked at random to die. With probability \(1–u\), the dead agent’s strategy is replaced with the reproducing agent’s strategy; with probability \(u\), a mutation occurs and instead the dead agent’s strategy is replaced with a randomly selected strategy. Thus, \(u\) is the mutation rate.

Each agent’s strategy is initialized randomly at the beginning of the simulation, and the strategies of all agents are recorded over \(10^5\) generations.

For details of the weak selection analytical calculations, see SI Single-Population Formulation.

Behavioral Experiments. To assess our theoretical predictions, we conduct a behavioral experiment. We recruit \(n = 140\) subjects using the online labor market Amazon Mechanical Turk (AMT; for an overview of running experiments on AMT, as well as a discussion of the value of combining behavioral experiments and theoretical models, see ref. 55). Commensurate with standard wages on AMT, subjects receive a $0.20 baseline payment for participating, and then play a UG in which $0.40 is at stake. For evidence that these low stakes do not compromise the validity of behavioral data, see ref. 57. Subjects read a set of instructions explaining the game, and are told they will be randomly assigned to be either the proposer or the responder. They are then asked to calculate the payoff received by the proposer and responder in two different scenarios to ensure that they understand the payoff structure. Only subjects who answer correctly are allowed to participate.

After clearing the comprehension questions, subjects indicate the minimum offer they would accept if they are assigned to be the responder. Then they indicate the amount they would offer if they are assigned to be the proposer. Finally, they complete a demographic questionnaire that includes the questions “Among those you interact with in daily life, how clear is it which people are more or less successful?” and “How accurate do you think
first impressions are when judging other people?,” each reported using a 5-
point Likert scale (1 = Very unclear to 5 = Very clear for the first question; 1 = Very inaccurate to 5 = Very accurate for the second).

Once all subjects have been recruited, they are randomly paired and assigned roles, the resulting payoffs are calculated, and each subject is paid accordingly using the AMT payment system. No deception is used. The practice of having subjects specify a strategy which dictates a decision in each possible outcome and then having actual payoffs determined by ex post matching, referred to as the “strategy method,” is a common technique in experimental economics (and is used by all of the experimental papers whose data we visualize in Fig. S4). This is particularly true for eliciting responder behavior in the LG, as low proposer offers are rare and thus it is difficult to determine how subjects would respond to receiving a low offer.

To analyze the results, we use linear regression with robust SEs. We find a significant negative effect of clarity of the successfulness of others on LG offers (coefficient = 0.061, P = 0.004) as well as a significant negative effect of the consistency of others on LG offer (coefficient = -1.241, P = 0.031). Similarly, we find a significant negative effect of clarity of the successfulness of others on LG demand (coefficient = -0.097, P = 0.029) but no significant effect of the consistency of others on LG demand (coefficient = 0.294, P = 0.568). These results are qualitatively unchanged when including controls for age, sex, income, education and US residency (clarity of success predicting offer: coefficient = -1.142, P = 0.032; clarity of success predicting demand: coefficient = -1.102, P = 0.044; consistency predicting demand: coefficient = 0.100, P = 0.854). This experiment was approved by the Harvard University Committee on the Use of Human Subjects in Research, Application F174682.

Supporting Information

Rand et al. 10.1073/pnas.1214167110

SI Text

We study the evolutionary dynamics of the Ultimatum Game (UG) analytically in two ways. The first approach, described in SI Single-Population Formulation, is a “single-population” formulation: we consider a population of size $N$ in which each individual can be both a proposer and a responder with equal probability. The second approach, described in SI Two-Population Formulation, is a “two-population” formulation: we consider two populations of size $N$, one of which is a population of proposers and the other, a population of responders, and we study their coevolution. Both approaches will yield the same result; an intuitive explanation of this result is then provided in SI Intuition on the Role of Mutation.

We then turn to additional agent based simulation results. Fig. S1 shows that the results in Fig. 1 in the main text are robust to different mutation rates. SI Local Mutation Kernel and Figs. S2 and S3 describe our simulations using a local mutation kernel (rather than the global mutation kernel used in the main text), and show that the results in Fig. 2 in the main text are robust to this alternate mutation structure. Fig. S4 shows the average values of $p$ and $q$ from numerous behavioral experiments as well as from agent-based simulations using particular parameter sets highlighted in Fig. 2 in the main text. Fig. S5 shows that the results in Fig. 2 in the main text are robust to different population sizes. Fig. S6 shows the distribution of individual-level offers and demands for a representative set of parameters.

SI Single-Population Formulation

Here, we consider a population of $N$ players who play the role of proposers and responders with equal probability. We specify an agent’s strategy as the pair $S = (p, q)$, $0 \leq p \leq 1$, and $0 \leq q \leq 1$, where $p$ is the amount offered when acting as proposer, and $q$ is the minimum amount demanded when acting as responder, or the rejection threshold. Hence, the strategy space for the UG is the unit square. Let $A(S_1, S_2)$ be the expected payoff that strategy $S_1 = (p_1, q_1)$ gets from strategy $S_2 = (p_2, q_2)$. Because we assume that in the interaction between a player using strategy $S_1$ and a player using strategy $S_2$ each player can be in the role of the proposer with equal probability, $A(S_1, S_2)$ is given (up to a 1/2 factor which we henceforth omit) by the function

$A(S_1, S_2) = \begin{cases} 
1 - p_1 + p_2 & \text{if } p_1 \geq q_2 \text{ and } p_2 \geq q_1 \\
1 - p_1 & \text{if } p_1 \geq q_2 \text{ and } p_2 < q_1 \\
p_2 & \text{if } p_1 < q_2 \text{ and } p_2 \geq q_1 \\
0 & \text{if } p_1 < q_2 \text{ and } p_2 < q_1 
\end{cases}$

[S1]

Every individual in the population plays the UG with every other individual and they all get payoffs according to the function above. The (relative) fecundity (or effective payoff) of a player with an average payoff $\pi$ is given by $\exp[\alpha \pi]$, where the parameter $\alpha > 0$ represents the intensity of selection. Individuals reproduce proportional to their fecundity. In each time step a random player dies and another player (including the dying player himself) is picked proportional to fecundity to replace the dead. Reproduction is subject to mutation: the offspring inherits the strategy of the parent with probability $1 - u$ and with probability $u$ it adopts a strategy uniformly at random. We say that strategy $S$ is favored on average in the mutation-selection equilibrium, if its abundance exceeds the mean.

Let us first assume that our strategies do not cover the entire unit square, but in fact are only of the form $S = ((i/m, j/m)$ with $1 \leq i, j \leq m$ being integers. This discretizes the problem, making it possible to invoke previous results. To then go back to the continuous strategy space we simply take the limit $m \to \infty$.

Having turned our continuous problem into a discrete one, we are now interested in the stationary abundance of these discrete strategies. For this problem, we can use the result in ref. 1 to conclude that, for large population size $N$, strategy $S$ is favored by selection if $L_S + NuHS > 0$, where

$$L_S = \frac{1}{m^2} \sum_{i=1}^{m} \sum_{j=1}^{m} \{A(S, S) + A(S, S') - A(S', S) - A(S', S')\}$$

[S2]

$$H_S = \frac{1}{m^2} \sum_{i=1}^{m} \sum_{j=1}^{m} \sum_{i'=1}^{m} \sum_{j'=1}^{m} \{A(S, S') - A(S', S')\}$$

Here, $S' = ((i'/m, j'/m)$ and $S'' = ((i''/m, j''/m)$. Moreover, ref. 1 showed that the higher the quantity $L_S + NuHS > 0$, the more the strategy $S$ is favored by selection. Consequently, to determine which strategy is most favored by selection, one simply has to maximize $L_S + NuHS > 0$.

Taking the limit $m \to \infty$ as in Tarnita et al. (2), the sums in S2 converge to the integrals

$$\bar{L}_S = \int_0^1 \int_0^1 \{A(S, S) + A(S, S') - A(S', S) - A(S', S')\} \, dp \, dq$$

$$\bar{H}_S = \int_0^1 \int_0^1 \int_0^1 \int_0^1 \{A(S, S') - A(S', S')\} \, dp \, dq \, dp' \, dq'$$

[S3]

where $S = (p', q')$ and $S' = (p'', q'')$. Moreover, it follows that the condition for strategy $S$ to be favored by selection is $\bar{L}_S + Nu\bar{H}_S > 0$ and that the most favored strategy is determined by maximizing $\bar{L}_S + Nu\bar{H}_S > 0$. Depending on whether $p \geq q$ or $p < q$, the payoff function $A(S, S)$ takes two different values and hence we find the condition for strategy $S$ to be favored by selection to be

$$\bar{L}_S + Nu\bar{H}_S = \left[ I(p \geq q) + p^2(-Nu - 2) + q^2 \left( \frac{Nu}{2} - 1 \right) \right] + p(1 + Nu) + q - \frac{1}{2} > 0,$$

[S4]

where $I(condition)$ is 1 if condition is true and is 0 if condition is false.

Maximizing S4, we conclude that the optimum strategy (most abundant in the stationary distribution and hence, by our measure, most favored by selection) is achieved when $p \geq q$ and is given by...
Note that for low mutation, the optimum strategy is \( (1/3, 1/3) \). Hence, the most successful strategy is one that offers 33\% and also rejects any offer lower than 33\%. As mutation increases, the proposal increases and the rejection threshold decreases. For high mutation, the most frequent strategy is \((1/2, 0)\); thus the proposal is 50\% and the rejection threshold is 0.

**SI Two-Population Formulation**

Next we will derive the same results as above but using a different approach. Instead of considering a population where each individual can be both proposer and responder, we consider two distinct populations—the population of proposers and that of responders—and explore their evolutionary game dynamics. This means that, unlike in SI Single-Population Mutation, where the strategy of an individual was given by \( S = [p, q] \in [0, 1] \times [0, 1] \), here the strategy of an individual is given by one number. Thus, an individual from the population of proposers will have strategy \( S_{\text{prop}} = p \in [0, 1] \), which represents the offer he makes and individual from the population of responders will have strategy \( S_{\text{resp}} = q \in [0, 1] \), which represents his rejection threshold. When two such players meet, their payoffs are given by

\[
A_{\text{prop}}(p, q) = \begin{cases} 
1 - p & \text{if } p \geq q \\
0 & \text{if } p < q
\end{cases} \tag{S5}
\]

\[
A_{\text{resp}}(p, q) = \begin{cases} 
p & \text{if } p \geq q \\
0 & \text{if } p < q
\end{cases}
\]

where \( A_{\text{prop}} \) and \( A_{\text{resp}} \), respectively, represent payoffs of the proposer and the responder.

Suppose that there are \( N \) players in each population (i.e., \( N \) proposers and \( N \) responders). Each proposer plays the UG described above with every responder in the responders’ population and obtains an average game payoff. Similarly, each responder plays the game with every proposer and obtains an average game payoff. We assume that selection occurs in each population according to payoffs in the UG. More specifically, a random player in either population dies (in cultural evolution terms, he attempts to change his strategy) and another player in the same population (it can be the dying player himself) replaces the dead with its offspring with probability proportional to one’s fecundity. We assume that one’s (relative) fecundity is given by \( \exp[w\pi] \), where \( \pi \) represents one’s average payoff in the game and \( w > 0 \) represents the intensity of selection. Reproduction (imitation) occurs with mistakes—with probability \( u \), the offspring is susceptible to mutation and randomly adopts a strategy uniformly at random, independently of its parent’s strategy.

For simplicity of our analysis, we will first discretize the problem, as before. Thus, we first consider that the possible proposals have the form \( S_{\text{prop}} = i/m \) and the possible rejection thresholds have the form \( S_{\text{resp}} = j/m \), where \( m \geq 1 \) is an integer and \( 1 \leq i, j \leq m \). In this case, assuming weak selection, \( w \to 0 \).

Ohtsuki et al. (3) have obtained the result that the combination of proposer’s and responder’s strategies that is most abundant in the stationary distribution (and hence, by our definition, is most favored by selection) is the one that maximizes

\[
L(i/m, j/m) + 2(N - 1)u \cdot H(i/m, j/m),
\]

where

\[
L(i/m, j/m) = \sum_{i=1}^{m} A_{\text{prop}}(i/m, j/m) - A_{\text{prop}}(i/m, j/m)
\]

\[
+ A_{\text{prop}}(i/m, j/m) - A_{\text{prop}}(i/m, j/m)
\]

\[
+ 1/m^2 \sum_{i, j=1}^{m} A_{\text{resp}}(i/m, m) - A_{\text{resp}}(i/m, m)
\]

\[
+ A_{\text{resp}}(i/m, m) - A_{\text{resp}}(i/m, m)
\]

\[
H(i/m, j/m) = \sum_{i, j=1}^{m} A_{\text{prop}}(i/m, j/m) - A_{\text{prop}}(i/m, j/m)
\]

\[
+ A_{\text{resp}}(i/m, j/m) - A_{\text{resp}}(i/m, j/m)
\]

where \( I(\text{condition}) \) is 1 if condition is true and 0 if condition is false. Now let \( p = i/m \) and \( q = j/m \). Substituting \( i = pm \) and \( j = qm \)

\[
L(p, q) = I(p \geq q) - 2p^2 - q^2 + p + q - \frac{1}{2}
\]

\[
H(p, q) = -p^2 - \frac{1}{2}q^2 + p
\]

A direct calculation now shows that, for large \( N \), the most abundant pair of strategies \((p_{\text{opt}}, q_{\text{opt}})\) obtained by maximizing \( L(p, q) + 2(N - 1)uH(p, q) \) is given by

\[
(p_{\text{opt}}, q_{\text{opt}}) = \begin{cases} 
\left( \frac{1}{2} \right) & \text{if } 0 < 2Nu \leq 1 \\
\left( \frac{1}{2} + \frac{2N}{m} \right) & \text{if } 1 < 2Nu
\end{cases}
\]

This result is exactly the same as the one for the single-population model in SI Single-Population Mutation except that \( Nu \) is replaced with \( 2Nu \). This makes sense, because the total population size in the two-population formulation is \( N = 2N \), whereas it is \( N \) in the one-population formulation.

**SI Intuition on the Role of Mutation**

As we have seen, the magnitude of the rescaled mutation rate \( Nu \) is a crucial determinant of the strategy that is favored the most by natural selection (= the one that is most frequently observed in the mutation-selection equilibrium of our stochastic evolutionary dynamics). Here, we try to explain the reason for that.

In our formulation, a mutant strategy, \( S = (p, q) \), is randomly chosen from our strategy space, which is the unit square \([0, 1] \times [0, 1] \). This assumption means that a mutant almost surely takes an initially unoccupied strategy and that those that are unoccupied are chosen with equal probability. However, when the population size is large, the empirical frequency of observing a particular strategy becomes close to its theoretical distribution. This implies that if the number of mutants is large, no strategy is expected to persist in the population. For a particular strategy, \( S = (p, q) \), the probability of observing it is obtained by sampling from the distribution of strategies of the population. This can be done by integrating the contribution of all strategies to their probability of observation (which is obtained by maximizing \( L(p, q) + 2(N - 1)uH(p, q) \) for \( Nu > 1 \)).
ulation of size $N$. Because our model corresponds to Moran infinitely many alleles model in population genetics, many results are already known. The expected number of different strategies coexisting in the population is given by the following exact formula (4):

$$
\sum_{k=0}^{N-1} \frac{N \nu}{N u + k(1-u)}
$$

[S10]

Table S1 shows some values of this expression for various $N$ and $u$. From S10 (but also from Table S1) we see that, in the low mutation limit $N \nu \rightarrow 0$, the number of different strategies in the population is close to 1, suggesting that the population is almost always monomorphic. When a new and rare mutant appears in the population, the number of different strategies becomes 2 (resident and mutant), and we expect that mutants will either die out or take over the resident population before another new mutant arises. Thus, at most two strategies are involved in a takeover at any moment in time. Therefore, in the low mutation, a strategy is selected if it can resist invasion by a single randomly chosen opposing strategy. At intermediate mutation rates, the evolutionary dynamic has characteristics of both the fully heterogeneous and fully homogeneous extremes. How these conclusions play a role in determining the winning strategies is explained at length in the main text.

As $N \nu$ increases, the number of different strategies present in the population also increases. In the high mutation limit, all strategies are present in the population simultaneously with approximately equal frequency. Hence, the optimum strategy is the one that maximizes its expected absolute payoff against a randomly chosen opposing strategy. At intermediate mutation rates, the evolutionary dynamic has characteristics of both the fully heterogeneous and fully homogeneous extremes. How these conclusions play a role in determining the winning strategies is explained at length in the main text.

SI Local Mutation Kernel

In the main text analysis and simulations, a mutant’s $p$ and $q$ values are randomly picked from the uniform distribution $[0, 1]$. Thus, mutations are “global,” in the sense that a mutant’s new strategy has no relation to the previous strategy. An alternative scheme, however, uses local mutation, where the mutant strategy is some perturbation off of the parent strategy. To investigate the effects of local mutation, we use the following mutation kernel.

For a parent with $p = p^*$, the mutant $p$ is picked from a $\beta$-distribution defined by the probability density function

$$
f(p) = \frac{\beta^{a-1}(1-p)^{\beta-1}}{\int_0^1 u^{a-1}(1-u)^{\beta-1}du}
$$

[S11]

where $\alpha = (p^* \gamma - 2p^* + 1)/(1 - p^*)$ and $\beta = \gamma$ if $x < 0.5$ or $\alpha = \gamma$ and $\beta = (1 - p^*) \gamma - 2(1 - p^*) + 1)/p^*$ if $x \geq 0.5$, and $\gamma$ is a parameter determining how similar the mutant tends to be to the parent. A $\beta$-distribution is used as this distribution is bounded on the interval $[0, 1]$ and is unimodal if $\alpha, \beta > 1$. The particular values of $\alpha$ and $\beta$ are chosen such that the modal value of the distribution is $p^*$.

Fig. S2 shows sample probability density functions for different values of $p^*$, using $\gamma = 50$. The same distribution is used to independently generate $q$ values. Fig. S3 shows the results of repeating the simulations shown in Fig. 2 in the main text, but now using this local mutation kernel $\gamma = 50$. We see qualitative agreement: Across a wide range of $w$ and $u$ values, we observe average $q > 0$ and $p > q$. Thus, our results are robust to the use of a local mutation kernel.

Fig. S1. For $u = 10^{-3}$, $u = 10^{-2}$, and $u = 10^{-1}$, respectively, and for $N = 100$, we show frequencies of $|p, q|$ pairs over $10^8$ generations, binned in increments of 0.1. The most common bin is indicated with a black $\times$. 

Rand et al. www.pnas.org/cgi/content/short/1214167110
Fig. S2. Local mutation kernel probability density functions for parent $p^*$, using $\gamma = 50$. 
Fig. S3. Agent-based simulation results using local mutation with $\gamma = 50$. Shown are time-averaged values of $p$ and $q$ over $10^8$ generations, using population $N = 100$ and (A) $u = 10^{-3}$, (B) $u = 10^{-2}$, and (C) $u = 10^{-1}$.
Fig. S4. Mean values of offers by player 1, p (blue), and minimum amount demanded by player 2, q (red), are shown from data (A) and evolutionary agent-based simulations (B). Observations are sorted by q value. (A) Data as presented in refs. 1–7. (B) All simulations with 0.3 < p < 0.5 and 0.2 < q < 0.35 are shown, from the set of mutation rate $u = [10^{-3}, 10^{-1}]$ and selection strength $w$ in log-scaled increments of 0.1. Simulations use population size $N = 100$, and show time-averaged values of $p$ and $q$ over $10^8$ generations. Fig. 2 in main text shows a systemic exploration of the relationship between $u$, $w$, and the average strategy.

Fig. S5. Agent-based simulation results using global mutation (as in the main text) but varying the population size $N$. Shown are time-averaged values of $p$ and $q$ over $10^8$ generations, using mutation rate $u = 0.01$ and (A) $N = 50$, (B) $N = 200$, and (C) $N = 1000$. 

Rand et al. www.pnas.org/cgi/content/short/1214167110
Fig. S6. Distribution of offers $p$ (A) and demands $q$ (B) for $N=100$, $u=0.01$, and $w=10^{-0.8}$. This is an alternative method of presenting the data shown in Fig. 1 in main text: here, $p$ and $q$ distributions are shown separately, rather than showing the joint $[p, q]$ distribution as is done in Fig. 1 in main text.

Table S1. Expected number of different strategies coexisting in the population

<table>
<thead>
<tr>
<th>$Nu$</th>
<th>$N = 100$</th>
<th>$N = 1,000$</th>
<th>$N = 10,000$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>26.0417</td>
<td>47.0201</td>
<td>69.6545</td>
</tr>
<tr>
<td>1</td>
<td>5.22322</td>
<td>7.49132</td>
<td>9.78842</td>
</tr>
<tr>
<td>0.1</td>
<td>1.50298</td>
<td>1.73318</td>
<td>1.96342</td>
</tr>
<tr>
<td>0.01</td>
<td>1.05162</td>
<td>1.07468</td>
<td>1.09771</td>
</tr>
</tbody>
</table>