TI 2021-074/I
Tinbergen Institute Discussion Paper

Evolution and the ultimatum game: Why do people reject unfair offers?

Revision: June 2022

Ashlihan Akdeniz[1,2]
Matthijs van Veelen[1,2]

1 University of Amsterdam

2 Tinbergen Research Institute

Evolution and the ultimatum game: Why do people reject unfair offers?

Aslıhan Akdeniz[1,2] & Matthijs van Veelen[1,2,*]
[1] University of Amsterdam, Amsterdam, The Netherlands, [2] Tinbergen Institute, The Netherlands, [*] Corresponding author email: c.m.vanveelen@uva.nl.

Abstract

In this paper we review, upgrade, and synthesize existing models from evolutionary game theory that aim at explaining behaviour in the ultimatum game, and we compare their predictions with the existing experimental evidence.

We find that the results in Gale et al. (1995) and Rand et al. (2013) are primarily driven by bias in the mutations. We make versions with local instead of global mutations for both. This minimizes the bias, and changes the results.

We also consider Quantal Response Equilibria in combination with the assumption that individuals are selfish after all. The Quantal Response Equilibrium is the noisy twin of the Nash equilibrium, and looking at this combination we explore an alternative explanation for what we observe in the lab, namely noise instead of deviations from selfishness.

Finally, we provide a refurbished version of the model of commitment in Nowak et al. (2000). The de-biased version of the model in Rand et al. (2013) becomes a special case of this more general model (with the possibility for commitment muted).

We find that the experimental evidence does not align with the models in Gale et al. (1995), Rand et al. (2013), or our de-biased versions of them, and that it also rejects the combination of selfishness and the Quantal Response Equilibrium. All of these models predict that the distribution of minimal acceptable offers should start with high frequencies at 0, end with low frequencies at 1, and have decreasing frequencies in between, which is not what is found in lab experiments.

Keywords: Ultimatum game, fairness, mutation-selection equilibrium, Quantal Response Equilibrium, commitment, experimental evidence

Preprint submitted to *14th June 2022*

1. Introduction

Humans are not just selfish. When deciding what to do, we do not only look at how our behaviour affects ourselves, but we also take into account the consequences of our actions for others. How we came to deviate from straightforward selfishness is one of the bigger questions in human evolution.

One of the classical games in which we see deviations from selfish money-maximizing behaviour is the ultimatum game (Güth et al., 1982). This game is played between a proposer and a responder. The proposer makes a proposal how to distribute a given amount of money between herself and the responder. The responder then accepts or rejects the proposal. If she rejects, neither player gets any money. For responders, the selfish money-maximizing choice would be to accept any proposal in which she gets a positive amount of money. That, however, is not what we find in lab experiments (Güth et al., 1982; Oosterbeek et al., 2004), where low offers regularly get rejected.

In this paper, we will review existing models from evolutionary game theory that aim at explaining this behaviour. We will also create new and improved versions of those models, describe their predictions in greater detail, and compare these predictions with the existing empirical evidence. This means that the paper will make a series of points, but we hope that the multitude of observations does not conceal the importance of each individual one.

1.1. Mutation-selection equilibria, bias, and the asymmetry argument

Two well-known models from the literature are Gale et al. (1995) and Rand et al. (2013). Both of these models describe mutation-selection equilibria. The ingredient that these models aim to capture is that not all suboptimal behaviours are equally costly. Rejecting a proposal in which the responder gets 1 euro and the proposer gets 9 costs the responder 1 euro. On the other hand, if offering 2 euros would have been accepted, then proposing 1 euro for the responder and 9 for oneself, and having this proposal rejected, costs the proposer 8 euros. In this example, one could therefore say that both the proposer and the responder made a mistake, and that the mistake made by the proposer is much more costly than the mistake made by the responder.

The concept of a mutation-selection equilibrium assumes that mutation creates a constant inflow of suboptimal strategies. In the ultimatum game, the asymmetry in how bad these mutations are for the fitness of their carriers then translates to an asymmetry in how long it takes for selection to eliminate them, and an asymmetry in how much they hurt the fitness of those they meet on their way out. Mutant proposers and mutant responders therefore ripple through the population differently, and that ends up having a nontrivial effect on what we should expect to see if mutation and selection balance in equilibrium.

A problem with both Gale et al. (1995) and Rand et al. (2013), however, is that both these models have global, and therefore biased mutations. As a result, the deviations from selfish, money maximizing behaviour that they find are primarily driven by the bias in the mutations, and not so much by the asymmetry in how costly different suboptimal behaviours are. Because mutation bias is not a good basis for an explanation, we redo both models, with mutations that are local instead of global. Switching from global to local mutations reduces the bias to a minimum, and changes the results significantly. In Section 2 we do this for the model in Rand et al. (2013), and in Section 3 we do this for the model in Gale et al. (1995).

The original papers, understandably, only focus on predictions regarding the average offer and the average threshold, below which responders start rejecting. We also look at other aspects of the prediction, such as, for instance, the relation between the average offer and the average threshold, the within population variance, and the variance across time, or across populations. In the model from Rand et al. (2013) low intensities of selection push offers and thresholds up, thereby getting them closer to levels found in experiments. In Section 2 we show that this comes at a cost, and that is that weakening selection also deteriorates the match between model predictions and empirical findings on these other dimensions.

1.2. Quantal Response Equilibria, learning dynamics, and the asymmetry argument (again)

Instead of looking at the ultimate level immediately, one can also look at a more proximate level first, and ask the question if the conclusion that humans deviate from selfish money-maximizing behaviour is justified. An alternative interpretation of the behaviour observed in the lab could be that individuals are in fact selfish, but that they are not perfectly informed, or otherwise not perfectly aware of what it is that they should do in order to earn as much money as they can. In order to do that, we calculate the Quantal Response Equilibria (McKelvey and Palfrey, 1995) for the ultimatum game (Yi, 2005), under the assumption of selfishness. This is a concept from classical game theory, that can also be seen as resulting from learning dynamics with noise. Also here, as with mutation-selection equilibria, the asymmetry argument is relevant, and also here, one can crank up the noise to get the average offer and the average threshold up quite a bit. We do however observe that other characteristics of the equilibrium distribution provide a poor match with the empirical evidence. In one of the two types of Quantal Response Equilibria, the distribution of MAO's is predicted to be downward sloping; the higher the MAO, the rarer they should be. This is not confirmed by the data, which suggests that there is more to human behaviour in the ultimatum game than everyone trying to get high monetary payoffs, but not knowing exactly what to do to get them.

This also carries over to explanations at the ultimate level. An equivalent argument about the mismatch between the shape of the distribution and the experimental evidence actually implies that every model that is only based on the asymmetry argument, and that does not include a pathway through which rejecting has actual fitness benefits, is to be rejected too. This therefore also applies to the models in Gale et al. (1995) and Rand et al. (2013) as well as our de-biased versions of them.

1.3. Commitment, and a unified model

Another well-known model for the evolution of behaviour in the ultimatum game is Nowak et al. (2000). In this model, rejecting itself is still bad for fitness, but accepting lower offers than others do, can also lead to getting lower offers than others – provided that proposers have a way of finding out how low they can go and still have their offer accepted. This means that the model specifies a pathway through which being the accepting type can actually be bad for fitness, and being the rejecting type can be good for fitness. That makes this model different from the models in Gale et al. (1995) and Rand et al. (2013), and their de-biased versions, in which there is never a fitness advantage to being the rejecting type.

We provide a version of the model from Nowak et al. (2000) which lifts some exogenously imposed restrictions on what strategies individuals can and cannot use. Our version is moreover a general model, in the sense that it contains our de-biased version of the model from Rand et al. (2013) as a special case. This helps illustrate the interaction between commitment (Frank, 1987; 1988; Akdeniz and van Veelen, 2021) and the asymmetry in costliness of mistakes.

2. Mutation-selection equilibria: Rand et al. (2013)

2.1. The simulation model in Rand et al. (2013)

Rand et al. (2013) consider a finite population model, in which 100 individuals play ultimatum games in both roles. Every individual has a strategy that specifies the offer they make in the role of proposer, as well as their minimal acceptable offer (MAO) in the role of responder. These offers and thresholds range from 0 to 1, and in the simulations we will be focusing on, they do so continuously. Each generation, every individual plays the ultimatum game with every other individual, once as a proposer and once as a responder. The resulting payoff is the average of the payoffs over all 99 pairings (in which a total of 198 games are played).

The population is updated according to a Moran process. One agent is picked at random to die, and individual $i \in \{1, ..., 100\}$ is picked with probability proportional to $exp\,(w\pi_i)$ to reproduce, where w is the intensity of selection, and π_i is the average payoff of individual i. Mutations happen at rate u at reproduction; with probability $1 - u$, the new individual inherits the strategy from the reproducing individual, and with probability u, the new individual carries a randomly selected strategy. The distribution from which the mutant is drawn is independent of the strategy before mutation; both the new offer and the new MAO are always drawn from a uniform distribution on $[0, 1]$. We will refer to this as global mutation.

The average trait value of the mutant is always $\frac{1}{2}$ – which is the value halfway the interval out of which the mutants are drawn – regardless of the trait value before mutation. Selection always works in favour of low values of the MAO's, and for low values of the MAO, it works in favour of low offers. That means that selection pulls these values towards the bottom of the interval $[0, 1]$. Therefore, when selection is at work, the average mutant has a higher offer and a higher MAO than the average offer and the average MAO in the population. In other words, mutation is biased, and mutants will result in increased offers and MAO's more often than they result in decreased offers and MAO's.

2.2. Our version

There are two inconsequential differences between their simulations and ours. The first is that we use a Wright-Fisher process instead of a Moran process. The Wright-Fisher process is computationally more efficient, but other than that, it perfectly reproduces the findings in Rand et al. (2013) for global mutations. The second inconsequential difference is that Rand et al. (2013) have co-occurring mutations; if an individual mutates, then both a new offer and a new MAO are drawn. Our version of the model has independent mutations. At any reproduction event, the offer mutates with probability u, and so does the MAO. That means that with probability u^2 mutations of the offer and of the MAO co-occur, and with probability $2u(1-u)$ only one of them mutates. Also this does not make much of a difference (see Supplementary Material 1.4 for details).

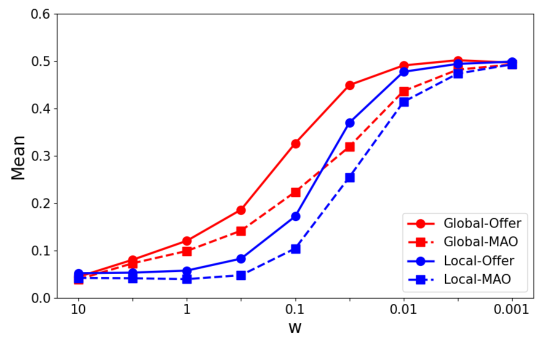
The important, and consequential difference is that in our version mutations are not global. Instead, mutations are changes with respect to the old trait value. That means that if a mutation of the offer happens, and the old offer is p, then the new offer is $p + \Delta p$, where Δp is drawn from a uniform distribution on $[-0.1, 0.1]$. There are two exceptions. The first is a result of the fact that we do not allow for offers below 0. Therefore, if $p + \Delta p < 0$, the new offer is 0. Similarly, we also do not allow for offers over 1, and therefore, if $p + \Delta p > 1$, the new offer is 1. This implies that mutations are unbiased for trait values in $[0.1, 0.9]$, and become a little biased if they drop below 0.1 or go over 0.9 (in which case the bias is still very small compared to the bias with global mutations in Rand et al., 2013). The same procedure applies to the MAO.

2.3. Global versus local mutation

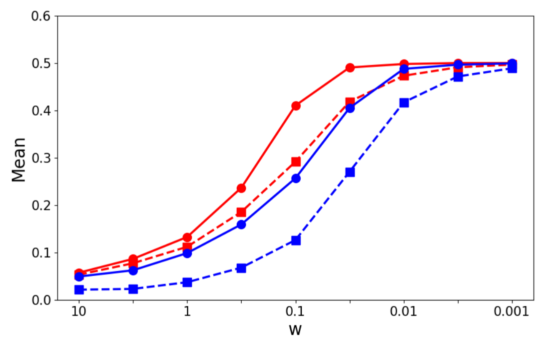
In Figure 1 we compare the results from Rand et al. (2013) with the results for our version. That makes this figure the counterpart of Figure 2 in Rand et al. (2013). For our figure, we did however choose to invert the horizontal axis. Their Figure 2 has low intensities of selection on the left and high intensities on the right. We do the opposite. The reason for that is that we want to make it clear that the benchmark, on the left, is the situation where responders accept all positive offers, and proposers offer nothing or close to nothing. The models investigate ways to arrive at dynamics that push the average offer and the average MAO up from 0, and we want it to be clear that reducing the intensity of selection does exactly that in both versions.

2.3.1. Lower offers, lower MAO's

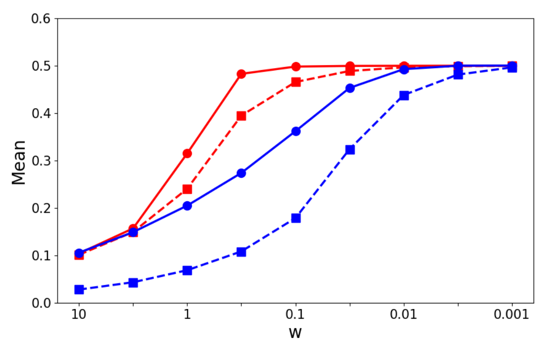
From the simulations, we learn that there are two important differences between global and local mutations. The first is that with the bias significantly reduced, the average offers and MAO's stay low for longer, and require further reduced intensities of selection to reach the same average offers and average MAO's. Here it is important to note that on the right hand side of the graph,



(A)



(B)



(C)

Figure 1: Global versus local mutations. In red the average offers and MAO's for the model in Rand et al. (2013), which has global, co-occurring mutations. In blue the same, but for local, independent mutations. Both the average offers and the average MAO's are higher with global, and therefore biased mutations, and lower with local, and therefore much less biased mutations. In order to get offers, or MAO's, up to average levels found in experiments, one would have to move to lower intensities of selection with local mutations than one would with global mutations. Section 2.4 explains why that is problematic. The mutation rate is 0.001 in panel A, 0.01 in panel B, and 0.1 in panel C.

with low intensities of selection, average offers and average MAO's end up at 0.5 in both versions. The reason why this eventually happens, and why that would happen for a range of reasonable modeling choices, is that 0.5 is halfway the parameter space, and therefore this must be the average over time in the limit of weak selection, where payoffs cease to matter. In Section 2.4 we will discuss in more detail why allowing for arbitrarily low intensity of selection, while focusing on how far offers and MAO's can be pushed up on average, limits the predictive power of the model on other criteria.

2.3.2. Gap vs. no gap

The second difference is there for all mutation rates in Figure 1, but it is the most visible for $u = 0.1$ (Fig. 1C). Here we see that on the left side of the graph, at high intensities of selection, there is no perceptible gap between the average offer and the average MAO for global mutation, while there is a very visible gap for local mutations. The latter is consistent with the asymmetry argument. Given that there is a consistent inflow of *local* mutations, proposers benefit from creating some space between their offer and the average MAO in the population; this way they reduce the risk that their offer is rejected by a responder with an above average MAO. Responders always get higher payoffs if they accept, and therefore they always benefit from moving their MAO down. The closer they get to an MAO of 0, however, the less of a difference a further decrease in their MAO makes. Also, if proposers increase their offers, that reduces the selection pressure against low MAO's. Both sides therefore want to create some distance, but since MAO's cannot drop below 0, that will result in mutations moving both averages up from 0, with a gap in between.

If we then start on the left hand side of the graph, and move a little to the right, then first the effect of reducing the intensity of selection is that this keeps mutants around for longer. With local mutations, this creates a wider distribution of offers and MAO's, which selects for strategies that on average keep more distance. This, in turn, leads to higher offers and MAO's due to the asymmetry in selection pressure. On the left end of the graph, we therefore see a widening gap, and an increase in offers and MAO's. Later on, when selection gets even weaker, and we get closer to the right end of the graph, everything just becomes noise. That causes both average offers and average MAO's to approach 0.5, which closes the gap.

With global mutations, on the other hand, there is hardly any gap at first. Here, the moving up of the offers and MAO's as selection gets weaker is the result of the bias in mutations balancing against ever weaker selection. The absence of a gap in the beginning therefore is understandable, because with *global* mutations, the equilibrium distribution of MAO's away from the mode is much more spread out. This implies that moving away from where most MAO's are does not make enough of a difference for the probability to have one's offer accepted, and that makes the reason to move away, that is there with local mutations, vanish.

Both differences – there being a gap versus there not being a gap at the left end of the graph, and the overall difference in average offers and average MAO's – indicate that with local mutations, the dynamics are mainly driven by the asymmetry in fitness effects, while the dynamics with global mutations are primarily driven by bias in the mutations. The latter is not a good basis for an explanation of deviations of selfishness.

2.4. Predictions for weak selection

In Section 2.3.1, we have seen that both with global and with local mutations, lowering the intensity of selection allows the average offer and the average MAO's to move away from 0, and towards $\frac{1}{2}$. There are however limitations to how observations about the averages for weak selection can be interpreted meaningfully. To see the reasons why, we will look a bit more closely at the dynamics in the absence of selection.

2.4.1. Reason 1: going against selection by shutting selection down

When the intensity of selection is 0, the dynamics in the model by Rand et al. (2013) are only driven by mutations. That implies that with global mutations, what we are seeing is the result of a sequence of random draws from a uniform distribution on $[0, 1]$, where no value of the draw is more likely to survive for longer and reproduce more than any other. Therefore, if we let the simulation run long enough, and we choose the intensity of selection to be 0, we will see the average offer and the average MAO converge to $\frac{1}{2}$ (which is the midpoint of the interval $[0, 1]$, and the expected value of the uniform distribution over it). By choosing a sufficiently weak intensity of selection, one can moreover get these averages anywhere between 0 – the limit for unfettered selection – and $\frac{1}{2}$ – the limit for unfettered mutation.

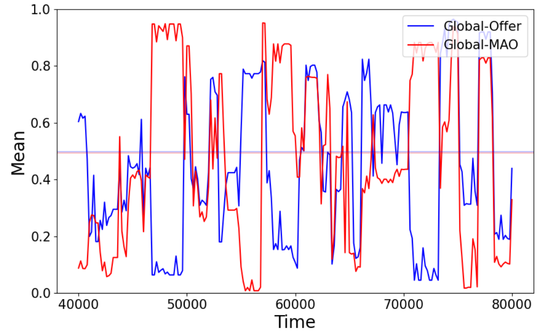
All of this implies that the fact that it is possible to get the average offer or the average MAO up to any value between 0 and 0.5, by choosing a sufficiently low intensity of selection, is not necessarily informative about selection – other than that selection always points towards lower offers and lower MAO's. Selection pulls both of them down, and if one reduces selection strength ever more, one can reduce by how much both are dragged down. The observation that one can find parameter choices for which the averages in the simulations match averages from experiments therefore is a somewhat arbitrary result of the fact that the average trait value in the strategy set is $\frac{1}{2}$, and not a reflection of what selection does to the strategies in this set. By definition of what happens in the limit of weak selection, and what happens in the limit of strong selection, deviations of which we try to explain, the model covers everything between offers and MAO's being 0, and the equal split.

A more general probabilistic symmetry argument, given in Supplementary Material 2.1, also applies to the version with local mutations. In this case, the results are driven much less by bias in the mutation, and much more by the asymmetry in costliness of mistakes, but the fact that also here *any* average offer below 0.5 can be reached by choosing a sufficiently low intensity of selection is an artefact of the fact that the neutral process, with mutation only, finds itself in the middle of the strategy space on average.

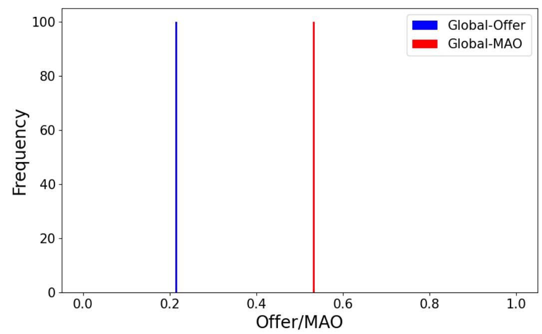
2.4.2. Reason 2: averages over the population and time versus averages over the population

There is also a second reason why not too much should be made of the fact that one get the average offer and the average MAO in the simulations to match average offers and average MAO's from lab experiments, if that requires choosing low intensities of selection. That reason has to do with the fact that the averages reported for the simulations are averages over populations and over time, and the averages in lab experiments are only averages over a given population. It is important to stress that these two are not the same. There are different ways in which the average in a population in a lab experiment can be the same as the average over the population and over time in the simulations, while other aspects of the simulations generate a remarkable mismatch with the empirical evidence.

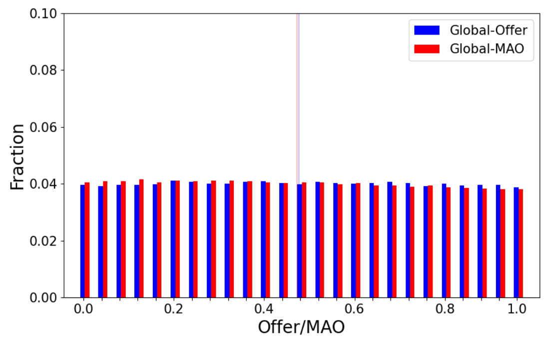
Figure 2A displays how the average offer and the average MAO within the population change over time in part of a run with relatively infrequent mutation, and weak selection. Mutations there are global and co-occurring, as they are in Rand et al. (2013). Figure 2B is a snapshot, which illustrates that most of the time, the population is at fixation, or close to it. The variance within the population therefore is almost always 0 or close to 0. Over time, however, the



(A)



(B)



(C)

Figure 2: Weak selection. The top panel shows how the average offer and the average MAO in the population change over time for part of a run with an intensity of selection of $w = 0.001$. Mutations are global and co-occurring, and the mutation rate is $u = 0.001$. In the neutral process, the average offer and the average MAO move completely independently. Here, with weak selection, they move almost completely independently (although the timing of changes to both coincides because of co-occurring mutations). The middle panel is a snapshot during the run. The bottom panel gives the average distribution over time, where we collected strategies within intervals of length 0.04. This average distribution is very close to the uniform distribution from which the mutants are drawn. The average over time of the average offers (MAO's) is a horizontal red (blue) line in panel A, and a vertical red (blue) line in panel C.

offers and MAO's are highly variable; Figure 2C indicates that they are quite literally all over the place.

When considering the results from lab experiments, we can assume that different populations are undergoing the same, or similar dynamics, and that implies that we may treat experiments in different populations as equivalent to different moments in time in the same simulation. If we do that, then the within population variance in experiments is much too large, and the between population variance in experiments is much too small to match the simulations, even if we can find model parameters for which the average over time of the average over the population in the simulations match the average for a sample from a population at a given moment in time.

2.4.3. Reason 3: lack of correlation between offers and MAO's

Another remarkable observation is that the offer and the MAO in these simulations are almost completely uncorrelated (this is also visible in Figure 2A). As a consequence, the average offer within the population is sometimes higher than the average MAO within the population, but almost equally often it is the other way around. Only when also averaged over time, is the average offer a bit above the MAO, but that masks that they move almost completely independently. Therefore, under weak selection, we should expect to find the average offer to be lower than the average MAO almost as often as the other way around. That is at odds with what is found in for instance cross-cultural experiments, where the offers in any given population are not independent of the income-maximizing offer in that population (Henrich et al., 2001; 2005; 2006). The lack of correlation between offers and MAO's in the simulations therefore is a remarkable mismatch with the empirical data.

Supplementary Material 2 shows that these mismatches are not confined to the combination of global and infrequent mutation. Whether mutations are global and co-occurring, as in Rand et al. (2013), or local and independent, as in our version, and whether mutations are frequent or infrequent, when selection is weak, averages over time from the simulations may coincide with averages from lab experiments, but predictions from the model that are not aggregated over time are not in line with the empirical evidence.

2.5. Mutation rates

The model in Rand et al. (2013) has two variables that can tilt the balance between mutation and selection; the intensity of selection, and the mutation rate. Decreasing the intensity of selection and increasing the mutation rate both make mutations overwhelm selection against rejecting positive offers. The reasons above point to limitations we encounter if we use the intensity of selection to push up average offers and MAO's for a given mutation rate. That still leaves the door open for increasing the mutation rate as a way to push offers and MAO's up.

A natural next question is therefore what a reasonable mutation rate is. For global mutations, it is important to realize that the "upward force" as a result of the bias scales up with the mutation rate. With global mutations, at a mutation rate of 1, individuals with high fitness still reproduce more than individuals with low fitness, but all selection is washed out completely by the bias. That means that, whether or not 1 is a realistic mutation rate, what is unrealistic for sure is that the force that is pushing the offers and MAO's up is just the bias in the mutations going in the other direction than selection.

For local mutations, on the other hand, there is only a little bit of bias around the edges (close to 0 and 1). That means that if dynamics take the offers and MAO's in the population up to intermediate levels, then even the moderate amount of bias that is there for trait values close to 0 disappears (instead of scaling up). The argument against high mutation rates with global mutations therefore does not apply with local mutations. One can moreover decide not to interpret the mutation rate too literally. There may be alternative genetic architectures that maintain the same variance within the population with much lower mutation rates. With sexual reproduction, for instance, no offspring is an exact copy of either of the parents. It is however important to realize that, for a given intensity of selection, with local mutations, it is *not* possible to push the average offer and MAO up to any level between 0 and 1. At 1, the highest possible mutation rate, these averages are not at $\frac{1}{2}$, but somewhere strictly (and, depending on the intensity of selection, possibly substantially) below $\frac{1}{2}$. All of this is discussed in more detail in Supplementary Material 1.3, where we fix intensities of selection, and let mutation rates vary.

2.6. WEIRD people

Another consideration that suggests we should allow for a margin of error when comparing averages from simulations and averages from experiments, is that those experiments tend to be done with WEIRD subjects, and the environment that makes us WEIRD is evolutionarily new. This is a point made by Henrich et al. (2010). One of the examples they point to is behaviour in the ultimatum game, and this is based on Henrich et al. (2006). In this study, they find that the income maximizing offers in two WEIRD populations (Emory students and rural Missouri) are relatively high compared to 13 non-WEIRD populations – and for the income maximizing offer to be high, there needs to be a relatively large share with a relatively large MAO. With WEIRD people having relatively high MAO's, experiments with WEIRD people therefore may set a bar that is a bit higher than necessary.

2.7. The shape of the distribution

Section 4 discusses a possible explanation of the data from lab experiments based on noise (instead of deviations from selfishness). This explanation is rejected by the empirical evidence, and this rejection is based on properties of the distribution other than the average offer or the average MAO. This mismatch between the empirical evidence and the predictions of the noise-based Quantal Response model also carries over to mutation-selection equilibria. It is helpful to first look at what one could consider to a be a somewhat more proximate explanation in order to understand what the prediction is, and why that would also follow from a mutation-selection model. Therefore, we will postpone this point to the end of Section 4. It may be good though to point to the fact that the mutation-selection equilibrium has another prediction in store, and to the fact that this one does not pertain to the average offer and the average MAO.

2.8. Summarizing

The results in Rand et al. (2013) are for a large part driven by bias in the mutations. If we un-bias the mutation process by replacing global mutations with local mutations, average offers and average MAO's in the simulations drop significantly. We can still get these averages up to levels found in experiments, but in order to do that, we have to choose really low intensities of selection. The fact that one can always do that, is, first of all, a somewhat gratuitous result of the fact that the intensity of selection can serve as a slider that can put us anywhere between the middle of the strategy space, and the point where selection alone would take us. Moreover, as we lower the intensity of selection, we may get the average offer and MAO (over time and over the population) closer to the averages (over the population) in experiments, but other characteristics of the prediction move away from what we observe – including the fact that for really low intensities of selection, the average offer and the average MAO are almost uncorrelated over time.

3. Mutation-selection equilibria: Gale et al. (1995)

Another paper that describes mutation-selection equilibria in the ultimatum game is Gale et al. (1995). While Rand et al. (2013) allow for an interpretation with genetic transmission as well as social learning, Gale et al. (1995) explicitly focus on the latter. There are also some technical differences. The model in Rand et al. (2013) has a finite population, for which they run stochastic simulations. Gale et al. (1995) on the other hand assume an infinitely large population, for which they calculate deterministic replicator dynamics. The strategy space in the main part of Rand et al. (2013) is continuous. The strategy space in Gale et al. (1995), on the other hand, is discrete; individuals can choose offers or MAO's only with certain, fixed increments. There are also some subtle differences concerning how mutation events and reproduction events relate.

These differences in modelling details come with differences in results. We will describe some of those differences here, and, in more detail, in the Supplementary Material. The similarities, however, are more important, and more prominent, than the differences. We will therefore first reproduce their main set of equations, and discuss what we see in equilibrium.

3.1. The model in Gale et al. (1995)

In Gale et al. (1995), the size of the pie is 40, but it is clear that one can choose any integer for size. We will therefore let n denote the amount to be divided. In Gale et al. (1995), proposers can offer $i = 1, ..., n$ to the responder; they can only offer integer numbers equal to or smaller than the pie size, but not including 0. Responders are characterized by an MAO, which is denoted by j, and which also ranges from 1 to n in steps of 1.

The differential equations that describe the dynamics are then given by

$$\dot{x}_i = (1-\delta)\left(\pi_{i,P} - \overline{\pi}_P\right) x_i + \delta\left(\frac{1}{n} - x_i\right)$$

for proposers, where x_i is the share of proposers that propose i, $\dot{x}_i$ is its time derivative, δ is the mutation rate, $\pi_{i,P}$ is the payoff of proposers that propose i, and $\overline{\pi}_P$ is the average payoff in the proposer population, and by

$$\dot{y}_j = (1-\delta)\left(\pi_{j,R} - \overline{\pi}_R\right) y_j + \delta\left(\frac{1}{n} - y_j\right)$$

for responders, where y_j is the share of responders with an MAO of j, $\dot{y}_j$ is its time derivative, $\pi_{j,R}$ is the payoff of responders with an MAO of j, and

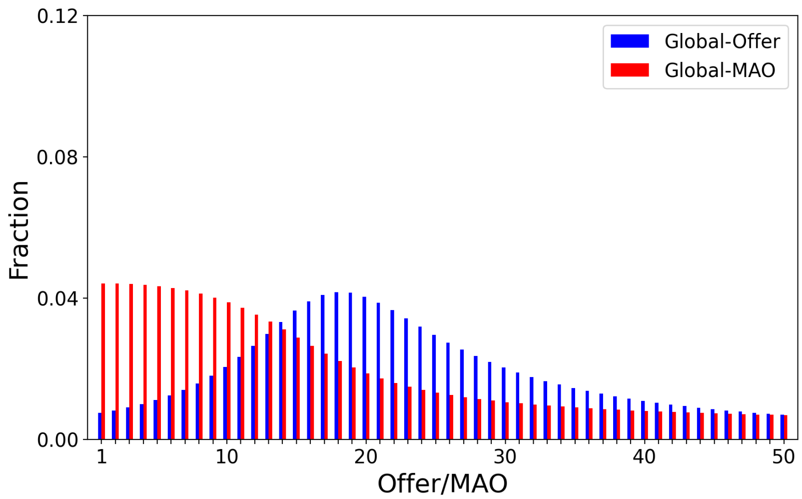


Figure 3: Mutation-selection equilibrium in Gale et al. (1995). The original model has global mutations, and this mutation-selection equilibrium has a mutation rate δ of 0.15. The thick tails of the distributions are a symptom of the bias in the mutation. With local mutations, the tails are much less thick (see Figure 4).

$\overline{\pi}_R$ is the average payoff in the responder population. The payoffs to different types of proposers depend on the composition of the responder population, and the payoffs to different types of reponders depend on the composition of the proposer population. In their paper, Gale et al. (1995) allow for the mutation rates to differ between the proposer and responder populations, but we will start with their default case, where they are the same.

We would like to keep the models of Gale et al. (1995) and Rand et al. (2013) as comparable as possible. Some of the simulation results from the model in Rand et al. (2013) are represented by frequencies of strategies in intervals of finite size; see for instance Figure 2B and C. In order to be as close as possible to that way of representing properties of simulation runs, we adjust the spacing of the strategies a little – which does not affect the equations above; the change only induces a minor change in how the payoffs are calculated. Instead of having proposer strategy i propose i, we choose for strategy i to propose the midpoint of the interval $[i-1, i]$, which is $i - \frac{1}{2}$. Similarly, we let responder strategy j have an MAO of $j - \frac{1}{2}$. That means we still have n strategies for both roles, but now we are not treating one end of the range from 0 to n differently; the smallest offer now is $\frac{1}{2}$ up from 0, and the largest is $\frac{1}{2}$ down from n, while before, 0 was excluded and n was included. This change is not consequential for what

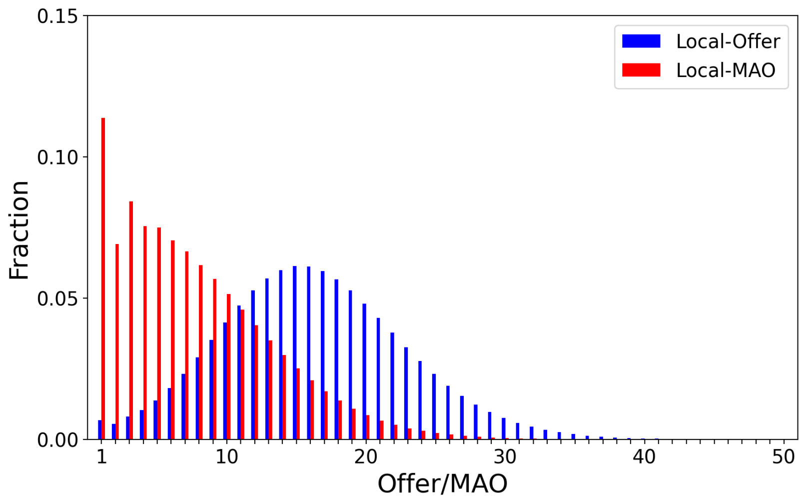


Figure 4: Mutation-selection equilibrium in Gale et al. (1995) with local mutations. This mutation-selection equilibrium has a mutation rate δ of 0.75. The tails are much thinner than with highly biased, global mutations. The spike at 1 and the dip at 2 are part of a dampening wave pattern caused by the remaining bias in mutations at the edges of the strategy space.

the mutation-selection equilibria look like.

Without mutations, at $\delta = 0$, almost all starting populations will converge to a population state where all responders have the lowest possible MAO and all proposers make the lowest possible offer. With mutations, that need not be the case. Mutations in Gale et al. (1995) are again global, as they introduce all MAO's and all offers at the same rate. This means that introducing mutations will by definition increase the average offer and the average MAO above 0 as a result of the bias. There are obviously also asymmetries in how fast suboptimal strategies are selected away, which creates the patterns in the offers and MAO's in Figure 3, but the main force behind the deviations from 0 with global mutations is the bias.

3.2. Our version

Because mutation bias is still not a good basis for an explanation, we also made a version of Gale et al. (1995) with local instead of global mutations. Local mutations work in a similar way as in our version of the model from Rand et al. (2013) with local mutations. A mutation induces a change in the offer, and this change can be up to a fixed number of steps to the right, or to the left, where all changes within that range are equally likely (with exceptions similar to those for our local mutations for Rand et al., 2013, if those changes would lead to offers or MAO's below 0 or over n). An example of a mutation-selection equilibrium with local mutations is given in Figure 4. Comparing the mutation-selection equilibria in Figures 3 and 4, we see that with local mutations, it takes much higher mutation rates to get to the same levels of average offers and MAO's, and that with local mutations, that happens without the thick tails that are symptomatic of the fact that with global mutations, higher mutation rates imply more upward push from the bias.

3.3. Finite versus infinite populations, and multiplicity of equilibria

In Gale et al. (1995), a mutation-selection equilibrium is a population state, characterized by a combination of frequencies of different strategies, for which the dynamics indicate no net change due to the combination of mutation and selection. The population state depicted in Figure 3 is such an equilibrium. These equilibria are moreover stable, in the sense that at least nearby population states move towards it, and sometimes there is even global convergence. What the authors seem to have overlooked, however, is that for one and the same combination of parameters, there can be multiple mutation-selection equilibria. In the Supplementary Material, we show that this is the case for low mutation rates. If the mutation rate is low enough, then there are multiple mutation-selection equilibria, at which almost all proposers make the same offer, and with a range of options for what that offer is. For higher mutation rates, there is just one, globally attracting, mutation-selection equilibrium.

The finite population dynamics in Rand et al. (2013) on the other hand are noisy, and not deterministic. The population will therefore keep moving around, and a mutation-selection equilibrium becomes a distribution over population states that reflects that some population states are visited (much) more often than others. By letting simulations run for a long time, we can figure out properties of this distribution of states, such as the average offer or the average MAO. This distribution is always unique, also if mutation rates are low enough for the infinite population version from Gale et al. (1995) to have multiple equilibria. The noise in Rand et al. (2013) would then make the population visit these different equilibria, and states close to them, over time.

In the Supplementary Material, we compare Gale et al. (1995) and Rand et al. (2013) by choosing versions of the latter with increasing population sizes.

We find that infinite population models are not a great approximation for finite population dynamics with small or even moderately sized populations.

3.4. Unequal mutation rates and Quantal Response

The setup in Gale et al. (1995) does allow for the possibility that mutation rates differ between proposers and responders. This is more reasonable for social learning than it would be for genetic transmission. With social learning, one could argue that if not much is at stake, there is less incentive to try to retain what you have learned. This kind of control over mutation rates make the agents more sophisticated than they are in the default version of the model, in which mutation rates are the same for both roles in the game. It also makes agents more sophisticated than they are in the model of Rand et al. (2013), where mutation rates are the same for offers and for MAO's.

The motivation that Gale et al. (1995) give for the unequal mutation rates is strikingly similar to the motivation given for the definition of a Quantal Response Equilibrium (McKelvey and Palfrey, 1995). A Quantal Response Equilibrium does not describe a mutation-selection equilibrium, so conceptually these are two different things, but both do have in common that they are ways in which the asymmetry in costliness of mistakes shapes how noise ripples through the population. In Quantal Response Equilibria, this noise in caused by perception error, or otherwise failures to maximize, and in mutation-selection equilibria the mutations are the sourse of the noise. The next section discusses Quantal Response Equilibria for the ultimatum game, and one important thing that we will see there, is that there is a whole set of models, including Quantal Response Equilibria and mutation-selection equilibria, that predict types of distributions that are not in line with the empirical evidence. We will make this general observation once we have also looked at Quantal Response Equilibria.

4. Quantal Response Equilibria

In this section, we will try to see if one can explain human behaviour in the ultimatum game without assuming that people deviate from selfishness. Instead, we assume that people are in fact selfish, but that they are also limited in their understanding of what it is that they need to do in order to maximize their fitness, or something that translates to fitness, like money. This imperfect understanding is formalized by the game-theoretic notion of a Quantal Response Equilibrium (QRE, McKelvey and Palfrey, 1995), which can be described as a statistical version of a Nash equilibrium, where suboptimal behaviour is not ruled out, but only assumed to be unlikely. The reason why this can be interesting for the question how behaviour in the ultimatum game has evolved, is that Quantal Response Equilibria can emerge as the result of a variety of learning dynamics. Like the mutation-selection equilibria discussed in Section 2, the QRE for the ultimatum game is shaped by the asymmetry in how costly mistakes are. After working our way through the details of the different types of QRE's, we will see that the empirical evidence actually rejects that humans play a QRE in which they try to maximize how much money they earn.

Looking at QRE's and comparing them to the empirical evidence is first of all interesting, because it helps rule out that people are selfish after all, and that their behaviour in the ultimatum game is just the result of not knowing exactly how to maximize their payoff. The deviations from selfishness we observe in experiments therefore cannot be explained away by people making mistakes. On top of this, the discrepancies between the empirical evidence on the one hand and the predictions of a combination of QRE and selfishness on the other also carry over to a larger class of evolutionary models at the ultimate level. Any dynamical model in which the reason why rejections are still present in equilibrium is that there is some source of noise that keeps introducing suboptimal behaviour, while selection keeps selecting against it, turns out to be inconsistent with the empirical evidence. That includes models that are in principle also open to an interpretation in which individuals evolve a preference for rejecting, as is the case for all mutation-selection equilibria discussed in the previous sections. This is an important, consequential observation, because it rules out a whole category of models that aim to explain the behaviour in the ultimatum game; all models that do not include a mechanism through which an actual fitness benefit is associated with rejecting proposals, do not explain behaviour in the ultimatum game. Before being able to articulate what the prediction is, and how that is refuted by the empirical evidence, it will be helpful to work through the technical details of the QRE, and first answer the more proximate question whether the behaviour can be reconciled with selfishness after all.

4.1. Quantal Response Equilibria and learning dynamics

The idea behind a Quantal Response Equilibrium (QRE, McKelvey and Palfrey, 1995, Goeree et al., 2016) is that players are imperfectly informed about the consequences of different behaviours. This concept does not assume anything about whether people are selfish or not; it can be combined with any type of preference, be it selfish, pro-social, anti-social, or inequity averse. Here, however, we will combine QRE with selfish preferences. The predictions therefore will be the result of a combination of selfish preferences and the Quantal Response model. Because of the fact that we do assume selfish preferences, words like "payoffs" will in fact coincide with money amounts when we use them below.

The defining property of a QRE is that strategies that result in high payoffs are played with higher probability than strategies that earn the agent lower payoffs. In the standard specification, the difference in probabilities is determined using a rationality parameter λ. The higher this rationality parameter, the larger the difference between these probabilities, and in the limit of $\lambda \to \infty$, only strategies that get the highest payoff are played. As a result, Quantal Response Equilibria become Nash equilibria in the limit of $\lambda \to \infty$. One reason why players might not be infinitely, or perfectly rational, is that increasing one's λ might not be free. At some point, getting better at recognizing which actions lead to high payoffs might not be worth the additional costs of boosting this capacity.

There are different ways to model individual behaviour that would imply dynamics that justify using the notion of a QRE. One such way is if individuals observe the payoffs in the population with a little bit of noise. This implies that their idea of what actions would get them the highest payoffs is mostly accurate, but due to the noise, they may sometimes think that the best they can do is choose an action that does not in fact come with the highest expected payoff. This is more likely to happen for actions that are close to optimal, for which it takes only a tiny shock to make it seem as if this is the optimal choice. In the resulting "*perturbed best response dynamics*" (Hofbauer and Sandholm, 2002; Sandholm, 2010; Alós-Ferrer and Netzer, 2010), individuals play what they think is the optimal thing to do against the current state of the population. If the noise follows a certain distribution, then this perturbed best response dynamic becomes the *logit response dynamics*, which can bring populations playing a game to a (logit) QRE.

A second way would be to assume instead that players adjust their behaviour locally, where they take their current strategy as their point of departure, and tend to adjust it in the direction in which payoffs increase. This is then combined with some amount of noise in how they adjust. The balance between those two factors implies that if payoffs increase steeply in one direction, individuals are most likely to adjust their behaviour in the right direction, and, in expectation, by a lot, whereas if payoff differences are small, then noise makes it more likely that they misdirect their adjustment. The resulting dynamics also converge to a QRE (Anderson et al., 2004). We do not focus on either of these two dynamic justifications; we just want to point to the fact that there is a variety of dynamic justifications for the concept of a QRE.

There are two versions of the QRE that we can apply to the ultimatum game; the *Agent*-QRE and the *Normal form*-QRE. There is a rationale behind those names, but it is not important for our purposes, and below we will just describe what they are for the ultimatum game.

4.2. Agent-QRE

If an offer in which the responder would get x is made, the responder chooses between, on the one hand, accepting, and getting x by doing so, and, on the other, rejecting and getting 0. In an *Agent*-QRE, that means that the responder is more likely to accept than to reject – unless $x = 0$ – and that this gap grows as x increases. For $x = 0$, there is no payoff difference, and therefore she accepts with 50% chance. In the logit specification of an *Agent*-QRE, the probability that the responder accepts depends on the offer x as follows:

$$\mathbb{P}\left(accept|x\right) = \frac{e^{\lambda \cdot x}}{e^{\lambda \cdot x} + e^{\lambda \cdot 0}} = \frac{e^{\lambda x}}{e^{\lambda x} + 1}$$

The formula itself is not too important, but for the comparison with the empirical evidence, it is important to observe that this would indeed predict that all positive proposals are more likely to be accepted than they are to be rejected, while the proposal $x = 0$ would have to be accepted half of the time. This is also illustrated by the red lines in Figure 5, that plot how the acceptance rates would depend on the proposal x for different rationality parameters λ.

Which offer would maximize the earnings for the proposer depends on what responders do. More precisely, what the best offer is, depends on the way in which the probability with which the responder accepts, changes with the offer that is made. For rationality parameters λ between 0 and 2, the probability with which the responder accepts is so insensitive to the proposal, that proposers are best off just proposing nothing for the responder and everything for themselves.[1] This proposal will then be accepted with 50% probability. Increasing the offer

[1]This can be found by taking the derivative of the expected earnings. These expected earnings are the amount the proposer gets if the offer is accepted (which is $1 - x$) times the probability with which it is accepted (which is $\frac{e^{\lambda x}}{e^{\lambda x}+1}$). The derivative of $(1 - x)\,\frac{e^{\lambda x}}{e^{\lambda x}+1}$ to x is negative for all $x \in (0, 1)$ for $0 \leq \lambda \leq 2$, while for all $\lambda > 2$, there is one – and only one – x within the interval $(0, 1)$ for which this derivative is 0.

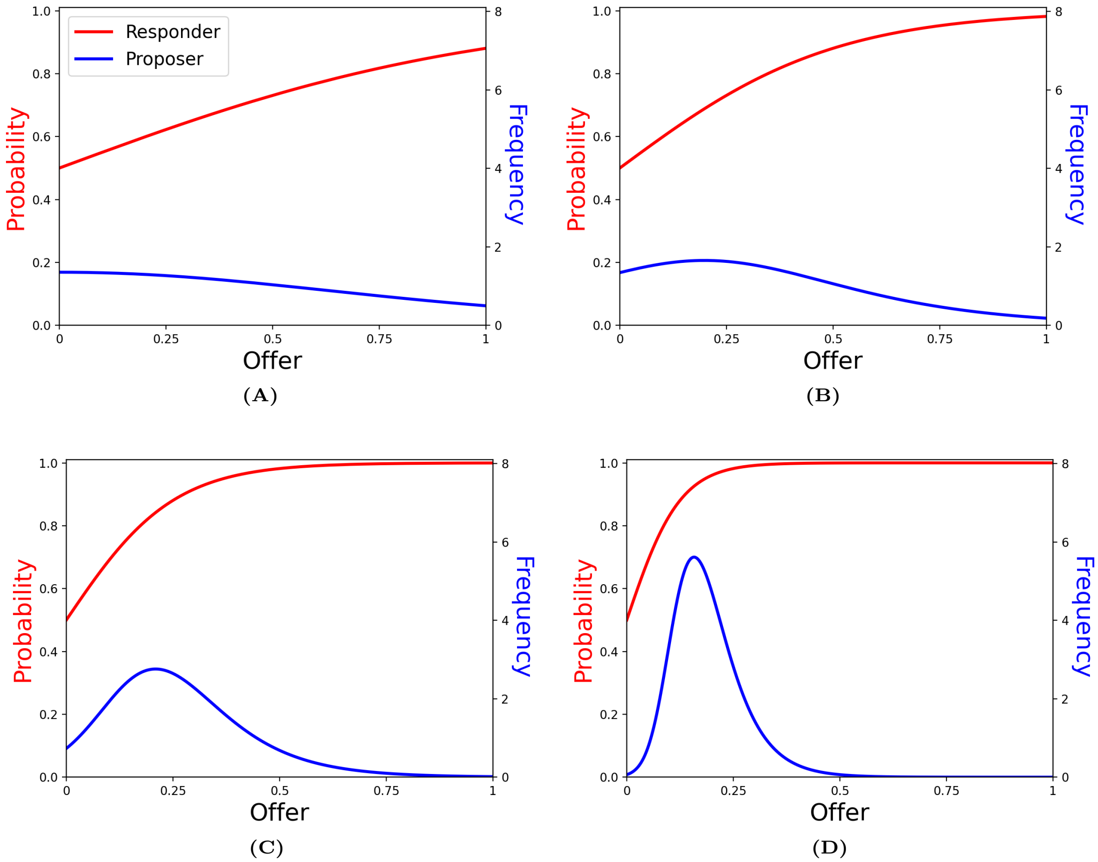


Figure 5: ***Agent*-QRE.** The red lines represent the probability with which the responder accepts an offer in which she gets x. The higher x, the larger the difference in payoff between accepting and not accepting, and therefore the higher the probability of acceptance. How strong the probability to accept responds to the payoff difference depends on the rationality parameter λ, which is $2, 4, 8$ and 16 in panels A, B, C and D, respectively. The blue lines represent the probability distribution over the offers made by the proposers in the QRE (just to be sure: this makes it is a different type of line than the red line is). The red line always starts at 0.5; the proposal in which the responder gets nothing is accepted with 50% chance.

does increase the probability with which the proposal is accepted a bit, but not enough to offset the reduction of the share of the pie when it is accepted. Therefore, for low λ's, the offer with the highest payoffs to the proposer, and therefore with the highest probability in the QRE, is 0 (see Figure 5A).

For λ's larger than 2, what the best response is first increases with λ. This can be seen in Figure 5B, where the peak of the blue graph has moved to the right. Later, for even higher λ's, changes in responder behaviour push the offer with the highest expected payoff back down again, which can be seen in Figure 5C and 5D, where the position of the peak moves back to the left, and gets ever closer to 0 as λ gets ever larger.

One can also see this from the formula for the density of offers made by the proposer in the logit specification for the *Agent*-QRE.

$$\frac{e^{\lambda(1-x)\frac{e^{\lambda x}}{e^{\lambda x}+1}}}{\int_0^1 e^{\lambda(1-y)\frac{e^{\lambda y}}{e^{\lambda y}+1}}dy}$$

The exponent in the numerator is λ times the expected payoff to the proposer of offering x. The density therefore peaks at the point where this expected payoff is maximized. High λ's moreover make for larger differences between the density for strategies with low expected payoffs and high expected payoffs. The peak therefore gets ever higher as λ increases, while the position of the peak, which is determined by the behaviour of responders, first moves to the right, and then back to the left.

4.3. Comparison empirics

For the comparison with the empirical evidence, we focus on responder behaviour, for which we pool data from different studies that use the direct-response method together (see Figure 6). The following studies are included: Andersen et al. (2011); Barmettler et al. (2012); Bornstein and Yaniv (1998); Cameron (1999); Carpenter et al. (2005a;b); Croson (1996); Forsythe et al. (1994); Lightner et al. (2017); Ruffle (1998); Slonim and Roth (1998). For each experiment, we extract the data for the standard ultimatum game, and discard the data for other treatments that vary certain aspects. Offers are calculated proportional to the total amount available in the ultimatum game in order to standardize the behavior across different experiments. Because it is not universally agreed upon whether stakes size matters, we also exclude the observations for the largest stakes in Andersen et al. (2011); Cameron (1999); Slonim and Roth (1998), while the Supplementary Material contains a version where we do include all stake sizes. The Supplementary Material also contains a version where we use data obtained with the strategy method to calculate rejection rates and compare them to the predictions of the *Agent*-QRE.

To test whether the predictions of the *Agent*-QRE fit the experimental evidence, we ran a logistic regression. With a logistic function, the probability that the offer is accepted is given by

$$\mathbb{P}\left(accept|x\right) = \frac{1}{1+e^{-(\beta_0+\beta_1 x)}}$$

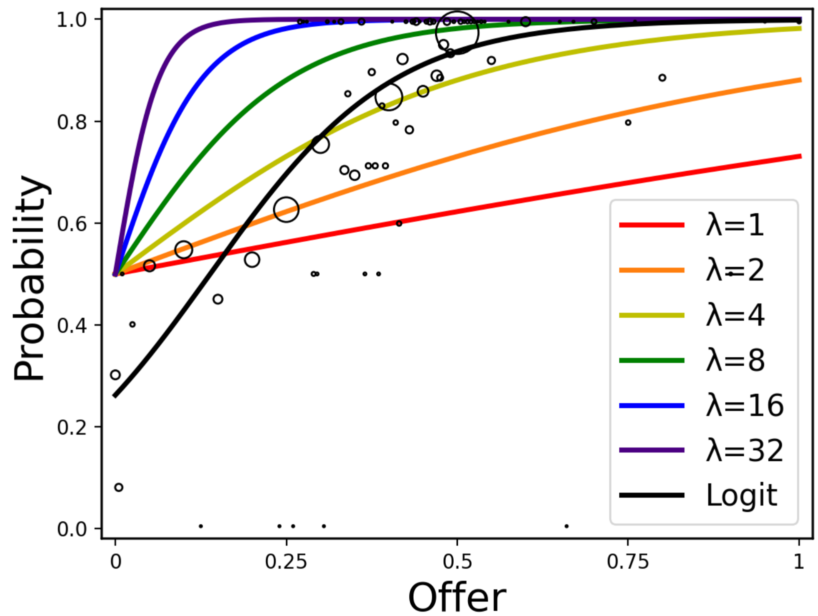


Figure 6: Acceptance/rejection rates in *agent*-QRE vs. empirical acceptance/rejection rates. The coloured lines are the acceptance rates in the *agent*-QRE for different λ's. The circles indicate acceptance rates for different proposals, pooling data from a number experiments together. Their size reflects the number of observations for that offer. The black line is the fitted acceptance rate as a function of the offer for a logit regression. Here, we use data obtained with the direct-response method, and we exclude some treatments with high stakes. In the Supplementary Material we include all stake sizes, and we compare the predictions of responder behaviour in the *Agent*-QRE with acceptance rates based on data obtained with the strategy method.

In the *Agent*-QRE, the acceptance probability of an offer in x is

$$\mathbb{P}\left(accept|x\right) = \frac{e^{\lambda x}}{e^{\lambda x}+1} = \frac{1}{1+e^{-\lambda x}}$$

This means that this probability in the *Agent*-QRE is nested in the logit specification, because the only difference between the two specifications is the intercept term β_0. The intercept is moreover the critical term, especially at the offers of 0, since the probability of accepting an offer of 0 *Agent*-QRE is

$$\mathbb{P}\left(accept|0\right) = \frac{1}{2}$$

whereas, for the logistic function including the intercept, it is

$$\mathbb{P}\left(accept|0\right) = \frac{1}{1+e^{-\beta_0}}$$

Depending on whether β_0 is statistically significantly different from 0 or not, we can therefore directly say something about the probability of accepting an offer of 0 being different from $\frac{1}{2}$.

We fit logistic functions with and without the intercept to test between the two specifications, and our results show a highly statistically significant coefficient on the intercept (p-value < 0.001). This indicates that the logistic regression including the intercept provides a significantly better fit than the *Agent*-QRE (see the Supplementary Material for details). The estimated coefficient on β_0 moreover has a negative sign, resulting in an estimated acceptance probability of an offer of 0 that is below 50%;

$$\mathbb{P}\left(accept|0\right) = \frac{1}{1+e^{-\beta_0}} < \frac{1}{2}$$

as $e^{-\beta_0} > 1$.

What this implies for the *Agent*-QRE is that the empirical acceptance rates are not consistent with the prediction of the QRE under the assumption that individuals are purely focused on maximizing their monetary payoff. Offers of 0 are accepted in significantly less than 50% of the cases, and there is an interval of low offers, for which subjects reject more often than they accept. The fact that there is such an interval is inconsistent with the idea that monetary payoffs are the only determinant of rejecting behaviour. Instead, it is consistent with subjects balancing the money they would get from accepting the offer against something else, which is best described as the joy of rejecting an unfair offer, or an aversion to accepting it.

4.4. Normal form-QRE

In the *Normal form*-QRE, we assume that proposers choose a proposal between 0 and 1, and responders choose an MAO between 0 and 1. That means that instead of considering responder strategies for each proposal separately, we consider strategies for the whole spectrum of possible offers. Moreover, the strategies we consider all have a natural, simple form; they have a threshold, above which they accept all offers, and below which they reject all offers. Many models reduce the strategy set this way, including the models in Gale et al. (1995) and Rand et al. (2013), as well as our version of the latter, all of which we discussed in Sections 2 and 3. The experimental evidence moreover suggests that this is not an unreasonable simplification; many people reject low offers and accept high offers, and switch from one to the other at some point in between.

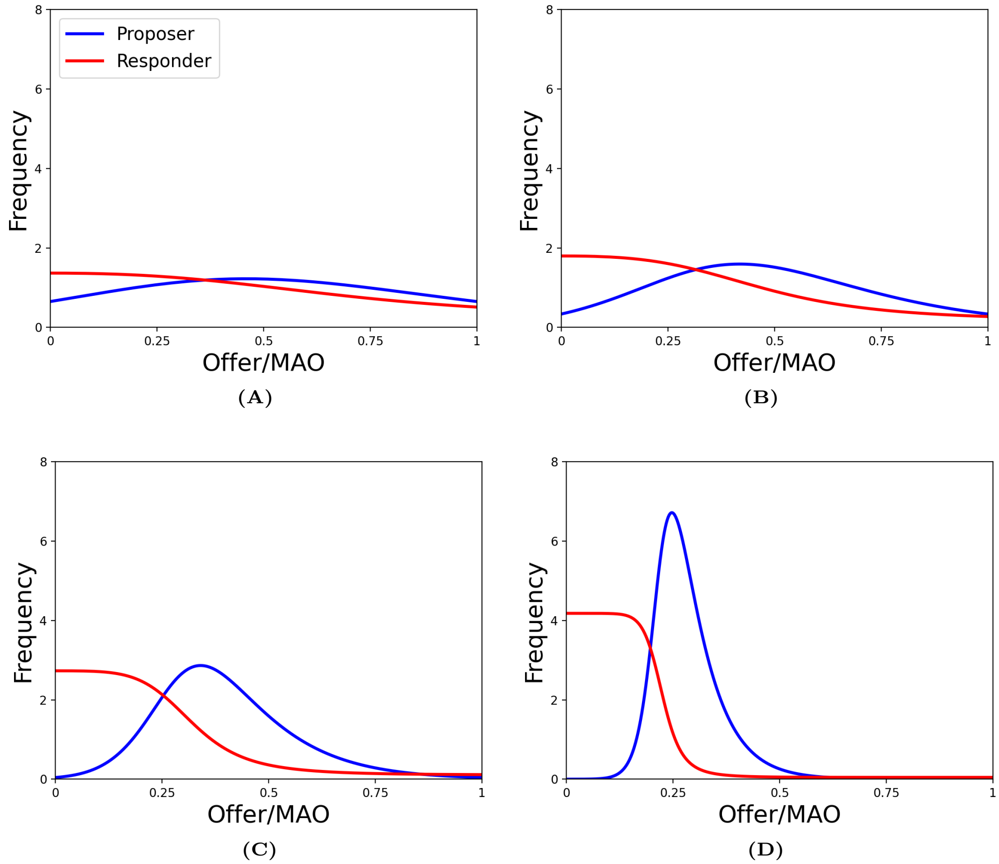


Figure 7: ***Normal form*-QRE.** The red lines represent probability distributions over MAO's for the responder, and the blue lines represent probability distributions over the offers made by the proposers, both in one and the same QRE. The rationality parameter, λ, is $2, 4, 8$ and 16 in panels A, B, C and D, respectively. The red graphs are all decreasing; in *Normal form*-QRE, lower MAO's occur with higher frequency than higher MAO's. The blue graphs have ever higher peaks, that start in the middle, and move ever more to the left.

In this setup, a QRE is a combination of distributions, one for the proposer and one for the responder. What the payoffs to different strategies for the proposer are, depends on the distribution of MAO's of responders, and vice versa. In equilibrium, strategies with higher expected payoffs are chosen with higher probabilities, and strategies with lower payoffs are chosen with lower probabilities, and this is true both for proposers and for responders.[2]

The *Agent*-QRE and the *Normal form*-QRE are not the same. In the *Normal form*-QRE, the proposer strategy that maximizes expected payoff starts at 0.5

[2]Just for completeness: if $f(x)$ is the distribution of proposals and $g(x)$ is the distribution of MAO's, then the following needs to be true for the combination of them to be a *Normal*

for $\lambda = 0$, and then only moves down. Therefore, if we look at Figure 7, we see that the peak starts in the middle, and always moves to the left (besides also becoming ever higher). This is different from how the distribution of proposer strategies changes with the increase of λ in the *Agent*-QRE, where the position of the peak starts at 0, first moves to the right, and then back to the left[3] (while the red lines are just not comparable, because they represent responder strategies in different ways).

These technicalities are not unimportant, but for the comparison with what subjects in labs do, what matters most is that the frequency with which players choose different MAO's decreases with the MAO; in the *Normal form*-QRE, an MAO of 0 is chosen the most, an MAO of 1 is chosen the least, and in between, if $0 < x < y < 1$, then x is chosen more often than y.

4.5. Comparison empirics

For the comparison with the empirical evidence, we focus on responder behaviour, for which we pool data from different studies that use the strategy method together (see Figure 8). The following studies are included: Bader et al. (2021); Bahry and Wilson (2006); Benndorf et al. (2017); Chew et al. (2013); Demiral and Mollerstrom (2020); Inaba et al. (2018); Keuschnigg et al. (2016); Peysakhovich et al. (2014). For each experiment, we extract the data for the standard ultimatum game, and discard the data for other treatments that vary certain aspects.

The majority of studies that use the strategy method restrict the subjects to strategies that can be characterized with an MAO. They ask their participants to submit a number, and if the offer they get is less than that number, it is rejected,

---

form-QRE for the ultimatum game:

$$f(x) = \frac{e^{\lambda \int_0^x (1-x) g(y) dy}}{\int_0^1 e^{\lambda \int_0^z (1-z) g(y) dy} dz}$$

$$g(y) = \frac{e^{\lambda \int_y^1 x f(x) dx}}{\int_0^1 e^{\lambda \int_z^1 x f(x) dx} dz}$$

[3]In the *Agent*-QRE, what responders do, changes with λ, but only directly, because a higher λ gives more weight to strategies with higher payoffs. Which responder strategies would result in what payoffs is not changing with λ, because in the *Agent*-QRE, these are calculated for any given proposal, which, if your partner just made it, is happening with probability 1. In the *Normal form*-QRE, the expected payoffs that different responder strategies generate do depend on what proposers do. The distribution of what responders do in the *Normal form*-QRE therefore depends on λ in an additional way, because λ also has an effect on what proposers do.

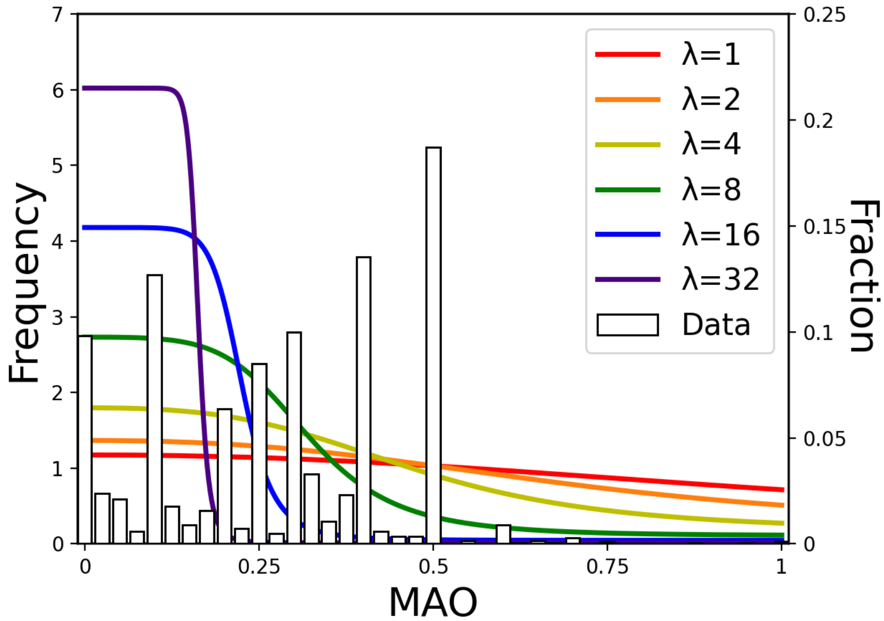


Figure 8: MAO's in *Normal form*-QRE vs. empirical MAO's. The coloured lines are the MAO's for different λ's. The bars indicate the frequencies of different MAO's in experiments. Because subjects gravitate towards round numbers, and because the increments subjects can choose also differ between experiments, we group the MAO's together as follows; the first bar is the frequency of MAO's of exactly 0, the second bar is the frequency of MAO's strictly between 0 and 0.05, the third bar is the frequency of MAO's of exactly 0.05, the fourth bar is the frequency of MAO's strictly between 0.05 and 0.1, and so on. Here, we use data obtained with the strategy method.

and if it is higher than or equal to it, it is accepted. There are however a few exceptions; Bahry and Wilson (2006) and Keuschnigg et al. (2016); Bader et al. (2021) ask participants to submit their accept/reject decisions for each possible offer separately. Participants in these studies therefore have the flexibility to switch between accepting and rejecting more than once, as opposed to the single switched point imposed by the method of submitting an MAO. For these studies we include participants who never switched, who switched only once (who start with rejecting and switch to accepting at a certain offer level), and those who switched twice (once from rejecting to accepting in the first half of the strategy space for offers, and another time from accepting to rejecting in the second half of the strategy space). We included this last group of subjects as well, as it seems that also rejecting hyper-fair offers is not a mistake, but aligns with an existing, consistent preference. In this case, we take the first switching point as their MAO. We do exclude other participants who do not fall into one of these categories. If the participant accepted all offers, we take their

MAO to be 0. MAO's are calculated proportional to the total amount available in the ultimatum game in order to standardize the behavior across different experiments (see the Supplementary Material for more details).

Figure 8 shows that the distribution of MAO's in experiments does not follow the pattern predicted by the *Normal form*-QRE. It is clear that the frequency of MAO's is not a decreasing function of the MAO, and as a simple indication of this, we can consider all MAO's below 0.25 on the one hand, and all MAO's above 0.25 up to, and including 0.5. The first interval, $[0, 0.25)$, contains fewer observations than the second one, $(0.25, 0.5]$, which is at odds with the distribution being a decreasing function.

4.6. Evolutionary dynamics in general

The evolutionary game theory models in the literature fit the setup of the *Normal form*-QRE perfectly. In Gale et al. (1995), Rand et al. (2013), and in our de-biased versions of both, there is a population of proposers that are characterized by their offers, and a population of responders that are characterized by their MAO's. Also there are similarities in the predicted distributions of offers and MAO's. We will therefore focus on how the mismatch for the *Normal form*-QRE carries over to evolutionary explanations at the ultimate level.

A good first observation is that the mutation-selection equilibria in Gale et al. (1995), Rand et al. (2013), as well as in our de-biased version, all have the property that the equilibrium distribution of MAO's goes from frequent to infrequent as the MAO's go from low to high. In other words, the density is the highest at 0, and then it decreases, until it is the lowest at an MAO of 1. That is a straightforward consequence of the fact that rejecting proposals is bad for fitness, and therefore having a lower MAO is always better than having a higher MAO.

The simple version of the original evolutionary question regarding human behaviour in the ultimatum game is: if rejecting is always bad for fitness, why do we observe rejections at all? At first sight, one might think that the mutation-selection models of Gale et al. (1995) and Rand et al. (2013) offer an escape from the iron logic that rejecting can only be bad for fitness, and should be selected against. Depending on parameter values, the average MAO in mutation-selection equilibrium can after all be sizable, and even if we de-bias the model, as we did in Section 2, the average MAO in equilibrium can still be pushed up to non-negligible amounts by choosing high mutation rates and, in Rand et al. (2013), low intensities of selection. In mutation-selection equilibria, or models with noise in general, we should however realize that the presence of across-the-board selection against rejecting does not imply that the average MAO in the population should be 0. The only thing that it implies, is that lower MAO's are favoured by selection over higher MAO's, and therefore higher MAO's should be observed less often than lower ones. The fact that the data do not align with that prediction, implies that no model with mutation-selection equilibria, in which selection works against rejections, can explain the rejecting of positive offers in humans. That remains true for all models in which rejecting is only bad for fitness. The observation that 0 is not the most common MAO in humans (far from it) therefore rejects all models that do not open up channels through which rejecting proposals can also bring fitness benefits.

4.7. Implications, great and small

The first implication of the comparisons of QRE's and observed behaviour in the lab is totally intuitive and unsurprising. People really deviate from selfishness, and what we observe in the lab is not some mirage caused by noise rippling through a population of selfish individuals asymmetrically.

The way this carries over to models that aim at giving ultimate explanations for human behaviour in the ultimatum game is less straightforward, and probably a bit more surprising, but therefore not less logically sound. Models in which all that happens is that some source of noise is added to the dynamics, without introducing a selective pressure that actually *favours* rejecting behaviour, are also at odds with the empirical evidence. These models do not predict that everyone in the population should have an MAO of 0, but they do predict that 0 should be the most common MAO (or, more generally, they predict that the higher the MAO, the less frequently it should be observed). That is clearly at odds with the empirical evidence.

5. Commitment

Nowak et al. (2000) propose a model for the evolution of behaviour in the ultimatum game in which the mechanism why rejections evolve is commitment (see also Frank, 1988, and Akdeniz and van Veelen, 2021). The rejecting itself is still bad for fitness, but their model opens a door through which being committed to rejections can be good for fitness. In their model, much the same as in other models, responders are characterized by a minimal acceptable offer (MAO), which is a threshold below which they reject proposals. Unlike other models, Nowak et al. (2000) allow proposers to sometimes observe the behaviour in past interactions of individual responders, and if they see that the responder accepted a proposal below what they would offer without observing, they lower their offer to what they know this responder apparently accepts. As a result, having a lower than average MAO, while leading to fewer costly rejections if unobserved, now has the disadvantage of also leading to worse offers, in case a player is observed to accept them.

In this section, we present a slightly upgraded version of the simulation model in Nowak et al. (2000). This illustrates a few core properties of this mechanism. The first is that, obviously, having a high MAO should sometimes lead to getting a higher offer for the mechanism of commitment to work. An individual's MAO must therefore be recognized from time to time, and proposers should sometimes do something with that information. On the other hand, the MAO does not always have to be recognized, and it does not have to be recognized perfectly, in order for commitment to work. A modest individual effect, by which those with higher MAO's on average get somewhat better offers, can still move the whole population to a state in which proposers serve their own interests by making sizable offers, even in cases where they do not have any information about the particular responder they are matched with.

Some of the differences between our version and the original have to do with restrictions on the admissible strategies that Nowak et al. (2000) impose. While these restrictions are not necessarily unreasonable, we felt that it is better to see if and when strategies evolve that satisfy them, rather than imposing them exogenously. Our version of Nowak et al. (2000) is also a generalization of the version of Rand et al. (2013) that we presented in Section 2. This allows us to explore the relative effects of asymmetry and commitment, and it allows us to illustrate the power of the combination of them. Also, it is aesthetically nice to have a unified model.

5.1. The simulation model in Nowak et al. (2000)

In the simulation model in Nowak et al. (2000), each individual is defined by a default offer p and a minimal acceptable offer q. In any given interaction, the proposer will offer whatever is smaller; her own p value, or the lowest amount that she knows was accepted by the responder during previous interactions. In addition, there is a small probability that the proposer will offer her value p minus some random number between 0 and 0.1. This makes sure that even if everyone in the population has the same p for a number of subsequent generations, there still will be observations for a range of lower proposals. Strategies are restricted to those with values for p and q that add up to a number that does not exceed 1; $p + q \leq 1$.

5.2. Our version

Our version first of all abstracts away from the way in which players find out about the MAO's of their partners. In Nowak et al. (2000), players sometimes observe past behaviour, from which they can make inferences about the MAO. We just assume that there is a fixed probability with which individuals know what the MAO of their partner is, and with the remaining probability they do not. The pathway could be reputation, but it can also be that people have other ways of recognizing individual differences in attitudes before playing.

Because Nowak et al. (2000) use reputation as a way for proposers to get information on the MAO of their partners, the mechanism behind the evolution of rejections here is sometimes classified as reputation (see for instance Debove et al., 2016, or Henrich et al., 2010). This is a defensible and understandable choice. What we would like to emphasize, though, is that in a population that is playing the ultimatum game, there are interesting incentives concerning communication. Proposers would like to be informed about the MAO of the responder they are matched with, so that they can maximize how much they can keep without getting their proposal rejected. Responders with high MAO's would like the proposer they are matched to know what their MAO is. Responders with below average MAO's however would like the proposer not to find out what their true MAO is. This partial alignment of the incentives for successful communication suggests that also without reputation, one could imagine some exchange of information to be established. Experimental evidence suggests that also in the absence of reputation, humans do indeed pick up on cues that help them predict rejecting behaviour with some success (van Leeuwen et al., 2018).

Therefore, what we want to stress is that one can also see commitment as the underlying mechanism for the evolution of rejecting behaviour; rejecting itself is still bad for fitness, but being committed to rejections is good, because it results in better offers (Frank, 1988; Akdeniz and van Veelen, 2021). This does require that others are able to identify, to some degree, who is committed. Reputation is one of the pathways to do that, but since it is not the only one, we abstract away from how it is that proposers tell different responders apart, and just include a parameter that represents the degree to which they can.

The second way in which our version is different, is that our proposers are characterized by two variables instead of one. One variable is their default offer, which they make if they are uninformed. The other is the maximum MAO they are willing to match as a proposer if they are informed about the MAO of the responder. Rather than assuming that proposers always match the MAO they observe, in our version, what they do with this information also evolves.

Also, in Nowak et al. (2000), proposers never propose more than their default proposal p, and only propose less than p if they know that will be accepted too. We allow for the possibility that proposers evolve to match the MAO of an opponent, also if it lies above their default offer p. The reason is that we think it is important to model the advantage it brings to be committed to an MAO that is above average, at least as much as it is important to allow being more accommodating than average to be exploited and selected against.

A fourth way in which our simulations are different, is that we do not assume that individuals sometimes lower their offer with a random amount, as in Nowak et al. (2000). This is not needed, because we abstracted away from the mechanism by which proposers are sometimes informed about the MAO of the responder they are matched with. We do have mutations on all traits, including the offer without observing, the same way as in our version with local mutations of Rand et al. (2013).

Finally, we do not impose the restriction that the default offer and the MAO should add up to a number that does not exceed 1. Our individuals can be endowed with any combination of those, as long as both are between 0 and 1. We do think that there are reasons why the offer and the MAO have not evolved to values larger than 0.5, but we prefer not to impose restrictions on the set of admissible strategies in order to rule out values above 0.5. Also, for understanding the working of the model, it will actually be instructive to allow values for p and q that both are larger than 0.5, even if we see reasons why these would not evolve (see also section 3.2.1 in Debove et al., 2016, where they point to the consequences of this restriction).

5.3. Results

Without observability, the model is the same as our version of Rand et al. (2013), but with local mutations. At $s = 0$, at the left end of Figure 9, the MAO and the offer without observing therefore are the same as they are for local mutations in Figure 1B at $w = 1$. The MO (the maximum offer they will make as a proposer to match the MAO of the responder, if observed) is a trait

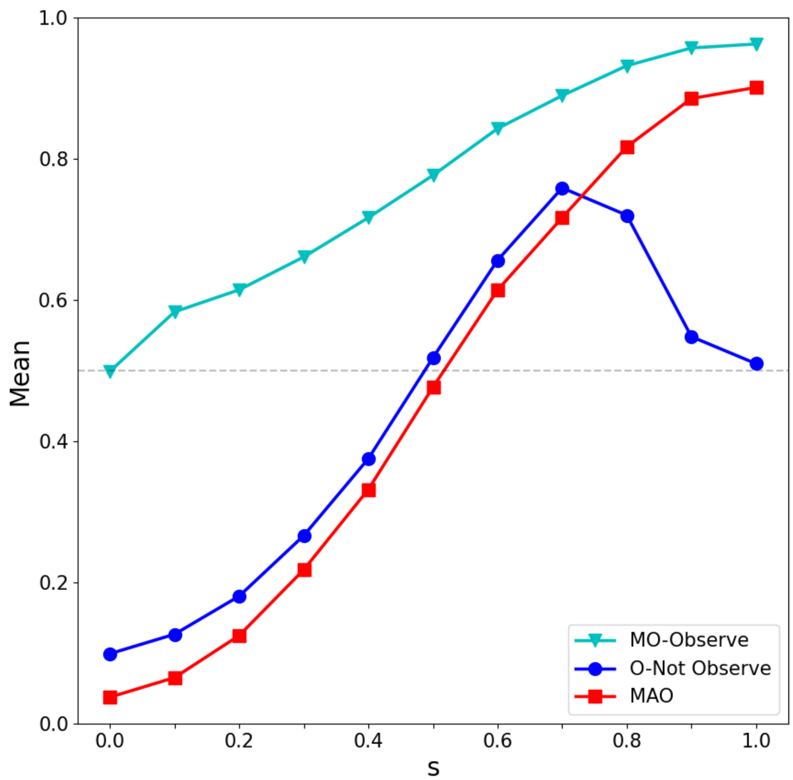
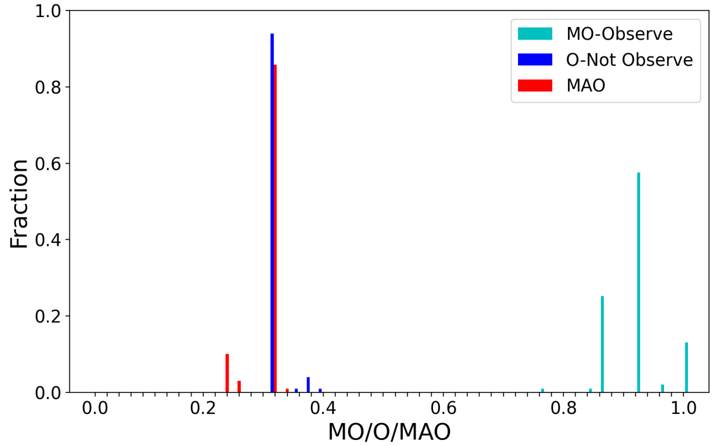


Figure 9: Local mutations and partial observability I. Individuals have three traits; the maximum offer (MO) they will make as a proposer to match the MAO of the responder, if observed; the offer (O) they make if they do not; and their MAO as a responder. The averages of these traits change with the probability with which proposers observe the MAO of the responder, which ranges from $s = 0$ (no observability) to $s = 1$ (full observability). Other parameter values are fixed at $u = 0.01$ and $w = 1$.

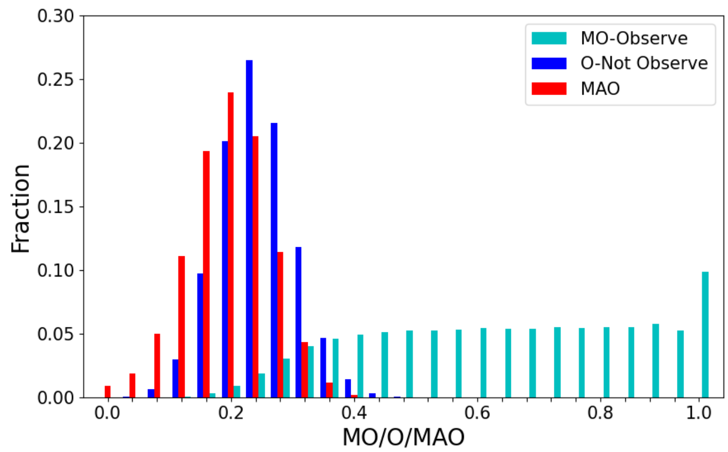
without fitness consequences if MAO's are never observed. That implies that at $s = 0$ it will drift neutrally within the interval $[0, 1]$, and will be 0.5 on average.

At $s = 1$ it is the offer without observing that becomes irrelevant, and will be 0.5 on average. Full observability moreover turns the tables on proposers and responders, because now the MAO of the responder is a given to proposers, who serve their own interest best by matching all positive MAO's. The situation at $s = 1$ therefore is the mirror image of the situation at $s = 0$, where the role of the offer without observing at $s = 0$ is played by the MAO at $s = 1$, and the role of the MAO at $s = 0$ is played by the MO at $s = 1$.

In between, we see that increasing the observability shifts the balance between



(A)



(B)

Figure 10: Local mutations and partial observability II. Panel A is a snapshot of the population, indicating the distribution of maximum offers (MO) they will make as a proposer to match the MAO of the responder; the offers (O) they make if they do not observe the responder's MAO; and their MAO as a responder. Panel B averages these across time, and thereby represents average distributions. Parameter values are fixed at $u = 0.01$, $w = 1$, and a probability of observing responder's MAO of $s = 0.3$. The spike at 1.0 and the dip just before 0.1 are part of a dampening wave pattern caused by the mutations being a little biased at the edges, where mutations beyond 1 are not possible.

the costs of being committed to rejecting low offers, which mainly occur when not observed, and the benefits, which only occur when observed. This pushes the average MAO up. The average offer without observing follows suit, because even though rejecting low offers evolves for when the MAO is observed, they are still a fact of life when not observed.

This is still a mutation-selection equilibrium, in which local mutations would flatten all distributions in the absence of selection, while selection can lift certain frequencies within the distributions up. When observability increases, selection for positive offers without observing becomes weaker. Up to $s = 0.7$, selection on the offers without observing the MAO keeps it above the average MAO within the population at all times. From $s = 0.8$ onward, with ever less selection countering the flattening of the distribution of offers without observing, the flattening sometimes wins, and sends the average offer without observing roaming below the average MAO, while selection at other times manages to temporarily stabilize the average offer without observing above the average MAO.

5.4. Model limitations

In this model, as in Nowak et al. (2000), we treat observability as an exogenous parameter. This is useful for illustrating how commitment works, but because of the partially aligned, partially misaligned interests between proposers and responders with respect to communicating the MAO's of responders, the observability is more likely to be endogenous, and subject to evolution itself.

One can also stack levels of observability on top of each other. On top of the probability with which proposers see the MAO of the responder, one could also introduce the probability with which the responder observes the MO of the proposer, and introduce the minimal MO she is willing to adjust her MAO to. While setting the first observability to 1 turns the tables to the benefit of responders, setting this second observability to 1 would turn the tables back to the benefit of proposers. We do not think this would be a particularly useful modelling exercise. We do however believe that there is a good reason why offers and MAO's would not exceed 0.5. The very nature of the ultimatum game makes it easier for proposers to commit than it is for responders. That implies that in a commitment tug-of-war between proposers and responders, we would expect that proposers will structurally find themselves on the shorter end of the stick.

6. Summary, discussion, reflection

In this paper, we have looked at a few prominent models from the evolutionary game theory literature that aim at explaining human behaviour in the ultimatum game. Gale et al. (1995) and Rand et al. (2013) both describe properties of mutation-selection equilibria, without a mechanism by which rejecting unfair proposals would get a selective advantage. We find that in both of them, the main driver of the results is bias in the mutations. This is not a good basis for an explanation. We made versions of both with local instead of global mutations. This minimizes the bias, and makes sure that the results are driven, not primarily by bias, but by the asymmetry in how costly mistakes are for proposers and responders. The reduction in bias makes average offers and average MAO's go down much more than the effect of the asymmetry makes them go back up again. The net change from global to local mutations therefore comes with significantly lower average offers and average MAO's. While the versions with local mutations capture the effect of the asymmetry in costliness of mistakes much better than the originals with global mutations, they still assume that rejecting is always bad, which is reflected by the fact that the mutation-selection equilibrium is characterized by higher MAO's always occurring less frequently than lower MAO's.

We also looked at Quantal Response Equilibria under the assumption that individuals are selfish. This noisy version of the Nash equilibrium, where people make mistakes in maximizing their payoff, and are more likely to make smaller mistakes than larger ones, comes in two versions. The first is characterized by probabilities of accepting that start at 50% for the offer of 0, and increases from there onward. The second is characterized by an equilibrium distribution in which an MAO of 0 has the highest density, an MAO of 1 the lowest, and the density always decreases as the MAO goes up everywhere in between. Both predictions are not confirmed by the experimental data, where we use data from existing experiments that use the direct response method to go with the first, and data from experiments that use the strategy method to go with the second prediction. The empirical evidence therefore rejects that the behaviour in the lab is the result of selfish people being imperfect at maximizing the amount of money they earn.

The second mismatch, between the observed shape of the distribution of MAO's and what Quantal Response would predict if people were selfish, also carries over to models at the ultimate level, as long as these models maintain the assumption that rejections can only be bad for fitness. This includes the mutation-selection equilibria in Gale et al. (1995), in Rand et al. (2013), and those in our de-biased versions of them.

The last model we looked at is Nowak et al. (2000). The mechanism at work there is commitment. If proposers have a way of finding out what the MAO of responders is, then having a higher MAO helps getting higher offers. The act of rejecting itself therefore is still bad for fitness, but being the rejecting type is good for fitness. We made an upgraded version of their model, that avoids making assumptions that rule out certain strategies a priori. Our version of Rand et al. (2013) also becomes a special case of our version of Nowak et al. (2000).

6.1. Inequity aversion and the mismatch hypothesis

The three papers we have focused on here are the best known evolutionary game theory models from the literature on this topic. They are however not the only ways in which one could try to explain human behaviour in the ultimatum game. One other possibility would be to assume that humans have inequity averse preferences (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000) that have evolved for playing other games, and that we inadvertendly bring to the ultimatum game too. That would imply that the rejecting behaviour in the ultimatum game is maladaptive.

Akdeniz and van Veelen (2021) show that there are two weak links in this argument. The first is that in most models for the evolution of deviations from selfishness, these "other games" are prisoner's dilemmas, and in those models, altruism evolves, or maybe spite, but not inequity aversion. The second is that this would imply that rejection rates should not depend on who makes the proposal – the person that the money is to be split with, or a computer – and that it should not depend on what the menu of possible proposals is that the proposer can choose from. Blount (1995) find that the first is not true, Falk et al. (2003) find that the second is violated (while an explanation with commitment would predict that whether or not we reject should depend on whether or not the other player is in fact responsible for an unfair proposal).

6.2. Other explanations

Akdeniz and van Veelen (2021) we also discuss why the fairness norm in the ultimatum game is not really group-beneficial – thereby ruling out a group selection argument – and they argue that repeating an ultimatum game would also not help explaining the behaviour that we find. Because those arguments are made elsewhere, we do not repeat them here. We also do not aim at making an exhaustive review of all existing models; there is already an excellent overview of the literature (Debove et al., 2016), and to complement that, we limited ourselves to discussing a few prominent ones in more depth.

6.3. Deviations from selfishness in general

There are many ways in which humans deviate from simply selfish money-maximizing behaviour. Studying deviations from selfishness in the ultimatum game therefore is part of a broader endeavour, that also tries to explain deviations from selfishness in other games. This gives another argument against asymmetry-based models – while mutation bias-based explanations are hardly ever a good option. For these other games, it is much more straightforward to see that asymmetry-based explanations could never work. Behaviour in the trust game can not be explained with models based on asymmetry in the costliness of mistakes; trustees not sending back money and trustors not trusting is very stable, also with mutations or noise. Also behaviour in the prisoners' dilemma or in the public good game, with or without punishment, cannot be explained on the basis of asymmetry. Here the simple reason is that these games are just not asymmetric. As an explanation for the human sense of fairness in general, therefore, asymmetry-based explanations would need to be combined with other mechanisms for deviations from selfishness in other games. That makes for instance commitment as a mechanism more parsimonious, because that gives an explanation of deviations from simple selfishness in a much wider variety of games (Frank, 1988; Akdeniz and van Veelen, 2021). That is not to say that asymmetries are irrelevant (they are not) but it makes it even more unlikely that asymmetry is the core driver of rejections in the ultimatum game.

Acknowledgements

We are very grateful to Donna Bahry, Rick Wilson; Volker Benndorf, Claudia Moellers, Hans-Theo Normann; Lisa Cameron; Jeffrey Carpenter, Eric Verhoogen, Stephen Burks; Soo Hong Chew, Richard Ebstein, Songfa Zhong; Rachel Croson; Elif Demiral, Johanna Mollerstrom; Robert Forsythe, Joel Horowitz, Gene Savin, Martin Sefton; Marc Keuschnigg, Felix Bader, Johannes Bracher, Bastian Baumeister, Roger Berger; Bradley Ruffle; Robert Slonim, Alvin Roth; and Damon Tomlin, for sharing their data with us, and to Steffen Andersen, Seda Ertaç, Uri Gneezy, Moshe Hoffman, John List; Franziska Barmettler, Ernst Fehr, Christian Zehnder; Gary Bornstein, Ilan Yaniv; Misato Inaba, Yumi Inoue, Satoshi Akutsu, Nobuyuki Takahashi, Toshio Yamagishi; Aaron Lightner, Pat

Barclay, Edward Hagen; Alexander Peysakhovich, Martin Nowak, and David Rand, for making their data publicly available.

Additional information

The code used in our simulation runs will be made publicly available on github.

References

Aslıhan Akdeniz and Matthijs van Veelen. The evolution of morality and the role of commitment. *Evolutionary Human Sciences*, pages 1–53, 2021.

Carlos Alós-Ferrer and Nick Netzer. The logit-response dynamics. *Games and Economic Behavior*, 68(2):413–427, 2010.

Steffen Andersen, Seda Ertaç, Uri Gneezy, Moshe Hoffman, and John A List. Stakes matter in ultimatum games. *American Economic Review*, 101(7):3427–39, 2011.

Simon P Anderson, Jacob K Goeree, and Charles A Holt. Noisy directional learning and the logit equilibrium. *The Scandinavian Journal of Economics*, 106(3):581–602, 2004.

Felix Bader, Bastian Baumeister, Roger Berger, and Marc Keuschnigg. On the transportability of laboratory results. *Sociological Methods & Research*, 50 (3):1452–1481, 2021.

Donna L Bahry and Rick K Wilson. Confusion or fairness in the field? Rejections in the ultimatum game under the strategy method. *Journal of Economic Behavior & Organization*, 60(1):37–54, 2006.

Franziska Barmettler, Ernst Fehr, and Christian Zehnder. Big experimenter is watching you! Anonymity and prosocial behavior in the laboratory. *Games and Economic Behavior*, 75(1):17–34, 2012.

Volker Benndorf, Claudia Moellers, and Hans-Theo Normann. Experienced vs. inexperienced participants in the lab: Do they behave differently? *Journal of the Economic Science Association*, 3(1):12–25, 2017.

Sally Blount. When social outcomes aren't fair: The effect of causal attributions on preferences. *Organizational Behavior and Human Decision Processes*, 63 (2):131–144, 1995.

Gary E Bolton and Axel Ockenfels. ERC: A theory of equity, reciprocity, and competition. *American Economic Review*, 90(1):166–193, 2000.

Gary Bornstein and Ilan Yaniv. Individual and group behavior in the ultimatum game: are groups more "rational" players? *Experimental Economics*, 1(1): 101–108, 1998.

Lisa A Cameron. Raising the stakes in the ultimatum game: Experimental evidence from indonesia. *Economic Inquiry*, 37(1):47–59, 1999.

Jeffrey Carpenter, Eric Verhoogen, and Stephen Burks. The effect of stakes in distribution experiments. *Economics Letters*, 86(3):393–398, 2005a.

Jeffrey P Carpenter, Stephen Burks, and Eric Verhoogen. Comparing students to workers: The effects of social framing on behavior in distribution games. In *Field experiments in economics*. Emerald Group Publishing Limited, 2005b.

Soo Hong Chew, Richard P Ebstein, and Songfa Zhong. Sex-hormone genes and gender difference in ultimatum game: Experimental evidence from china and israel. *Journal of Economic Behavior & Organization*, 90:28–42, 2013.

Rachel TA Croson. Information in ultimatum games: An experimental study. *Journal of Economic Behavior & Organization*, 30(2):197–212, 1996.

Stephane Debove, Nicolas Baumard, and Jean-Baptiste André. Models of the evolution of fairness in the ultimatum game: a review and classification. *Evolution and Human Behavior*, 37(3):245–254, 2016.

Elif E Demiral and Johanna Mollerstrom. The entitlement effect in the ultimatum game–does it even exist? *Journal of Economic Behavior & Organization*, 175:341–352, 2020.

Armin Falk, Ernst Fehr, and Urs Fischbacher. On the nature of fair behavior. *Economic Inquiry*, 41(1):20–26, 2003.

Ernst Fehr and Klaus M Schmidt. A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 114(3):817–868, 1999.

Robert Forsythe, Joel L Horowitz, Nathan E Savin, and Martin Sefton. Fairness in simple bargaining experiments. *Games and Economic behavior*, 6(3):347–369, 1994.

Robert H Frank. If *homo economicus* could choose his own utility function, would he want one with a conscience? *American Economic Review*, 77(4): 593–604, 1987.

Robert H Frank. *Passions Within Reason: The strategic role of the emotions.* WW Norton & Co, 1988.

John Gale, Kenneth G Binmore, and Larry Samuelson. Learning to be imperfect: The ultimatum game. *Games and Economic Behavior*, 8(1):56–90, 1995.

Jacob K Goeree, Charles A Holt, and Thomas R Palfrey. Quantal response equilibrium. In *Quantal Response Equilibrium*. Princeton University Press, 2016.

Werner Güth, Rolf Schmittberger, and Bernd Schwarze. An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization*, 3(4):367–388, 1982.

Joseph Henrich, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis, and Richard McElreath. In search of homo economicus: behavioral experiments in 15 small-scale societies. *American Economic Review*, 91 (2):73–78, 2001.

Joseph Henrich, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis, Richard McElreath, Michael Alvard, Abigail Barr, Jean Ensminger, et al. "Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences*, 28(6): 795–815, 2005.

Joseph Henrich, Richard McElreath, Abigail Barr, Jean Ensminger, Clark Barrett, Alexander Bolyanatz, Juan Camilo Cardenas, Michael Gurven, Edwins Gwako, Natalie Henrich, et al. Costly punishment across human societies. *Science*, 312(5781):1767–1770, 2006.

Joseph Henrich, Steven J Heine, and Ara Norenzayan. The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2-3):61–83, 2010.

Josef Hofbauer and William H Sandholm. On the global convergence of stochastic fictitious play. *Econometrica*, 70(6):2265–2294, 2002.

Misato Inaba, Yumi Inoue, Satoshi Akutsu, Nobuyuki Takahashi, and Toshio Yamagishi. Preference and strategy in proposer's prosocial giving in the ultimatum game. *PloS One*, 13(3):e0193877, 2018.

Marc Keuschnigg, Felix Bader, and Johannes Bracher. Using crowdsourced online experiments to study context-dependency of behavior. *Social Science Research*, 59:68–82, 2016.

Aaron D Lightner, Pat Barclay, and Edward H Hagen. Radical framing effects in the ultimatum game: the impact of explicit culturally transmitted frames on economic decision-making. *Royal Society Open Science*, 4(12):170543, 2017.

Richard D McKelvey and Thomas R Palfrey. Quantal response equilibria for normal form games. *Games and Economic Behavior*, 10(1):6–38, 1995.

Martin A Nowak, Karen M Page, and Karl Sigmund. Fairness versus reason in the ultimatum game. *Science*, 289(5485):1773–1775, 2000.

Hessel Oosterbeek, Randolph Sloof, and Gijs Van De Kuilen. Cultural differences in ultimatum game experiments: Evidence from a meta-analysis. *Experimental Economics*, 7(2):171–188, 2004.

Alexander Peysakhovich, Martin A Nowak, and David G Rand. Humans display a 'cooperative phenotype' that is domain general and temporally stable. *Nature Communications*, 5(1):1–8, 2014.

David G Rand, Corina E Tarnita, Hisashi Ohtsuki, and Martin A Nowak. Evolution of fairness in the one-shot anonymous ultimatum game. *Proceedings of the National Academy of Sciences*, 110(7):2581–2586, 2013.

Bradley J Ruffle. More is better, but fair is fair: Tipping in dictator and ultimatum games. *Games and Economic Behavior*, 23(2):247–265, 1998.

William H Sandholm. *Population games and evolutionary dynamics*. MIT press, 2010.

Robert Slonim and Alvin E Roth. Learning in high stakes ultimatum games: An experiment in the slovak republic. *Econometrica*, pages 569–596, 1998.

Boris van Leeuwen, Charles N Noussair, Theo Offerman, Sigrid Suetens, Matthijs van Veelen, and Jeroen van de Ven. Predictably angry–Facial cues provide a credible signal of destructive behavior. *Management Science*, 64(7): 3364–3364, 2018.

Kang-Oh Yi. Quantal-response equilibrium models of the ultimatum bargaining game. *Games and Economic Behavior*, 51(2):324–348, 2005.

Supplementary Material for:

Evolution and the ultimatum game: Why do people reject unfair offers?

Abstract

In the supplementary material, we discuss a few things in more detail. Section S1 compares Rand et al. (2013) with our version with local mutations. Section S2 describes why choosing arbitrarily weak selection is problematic. Section S3 describes the model in Gale et al. (1995) as well as our version with local mutations, and illustrates the possibility of multiple equilibria for the former. Section S4 illustrates the link and the differences between the models in Gale et al. (1995) and Rand et al. (2013). Section S5 discusses some details of Quantal Response Equilibria, and how their predictions are compared to experimental data.

S1 Finite population models

S1.1 The model in Rand et al. (2013)

The model in Rand et al. (2013) has a finite population, in which 100 individuals play ultimatum games in both roles. Every individual has a strategy that specifies the offer they make in the role of proposer, and their MAO in the role of responder. These offers and thresholds range from 0 to 1. Each generation, every individual plays the ultimatum game with every other individual, once as a proposer and once as a responder. The resulting payoff is the average of the payoffs over all 99 pairings.

The population is updated according to a Moran process. One agent is picked at random to die, and individual $i \in \{1, ..., 100\}$ is picked with probability proportional to $exp\,(w\pi_i)$ to reproduce, where w is the intensity of selection, and π_i is the average payoff of individual i. Mutations happen at rate u at reproduction; with probability $1 - u$, the new individual inherits the strategy from the reproducing individual, and with probability u, the new individual carries a randomly selected strategy. If a mutation happens, both the new offer and the new MAO are drawn from a uniform distribution on $[0, 1]$.

S1.2 Our versions

There is one general, inconsequential difference between their simulations and ours, and that is that we use a Wright-Fisher process instead of a Moran process. The Wright-Fisher process is computationally more efficient, but other than that, it perfectly reproduces the findings in Rand et al. (2013) for global mutations. The more important difference is that in our version, mutations are local. We consider two local alternatives for the mutation process.

S1.2.1 Local, co-occurring mutations

In the first one, mutations on both dimensions (offer and MAO) are co-occurring, as they are in Rand et al. (2013). That means that if a mutation happens, then both a new offer and a new MAO are drawn. The only difference with Rand et al. (2013) is that they are drawn from a local distribution, instead of a global one. If the old offer is p, then the new offer is $p + \Delta p$, where Δp is drawn from a uniform distribution on $[-0.1, 0.1]$. There are two exceptions. If $p + \Delta p < 0$, the new offer is 0. Similarly, if $p + \Delta p > 1$, the new offer is 1. The same procedure applies to the MAO.

S1.2.2 Local, independent mutations

In the second version, mutations in the offer or the MAO happen independently. At any reproduction event, the offer mutates with probability u, and so does the MAO. That means that with probability u^2 mutations of the offer and of the MAO co-occur, with probability $2u\,(1-u)$ only one of them mutates, and with probability $(1-u)^2$ neither of the two mutates. Mutations still happen locally, as described above.

The differences between these two versions are relatively small (see Section S1.4). Because the second version is more elegant, this is the one that we use here and in the main text.

S1.3 Global versus local mutation

The first question that Rand et al. (2013) answer for their model, and that we answer for ours, is: which combinations of the intensity of selection and the mutation rate put the average offer and the average MAO in the range of the averages in empirical findings. There are two ways to rephrase that question, or to visualize the answer. The first is: for a given mutation rate, how low would the intensity of selection have to be in order to get the offers and MAO's up to levels found in experiments. The second is: for a given intensity of selection, how high would the mutation rate have to be in order to get offers and MAO's up to the levels found in experiments.

For the figures in the main text, we took the first approach: we fixed a mutation rate, and considered a variety of intensities of selection. This way these figures indicate how far we would have to reduce the intensity of selection in order to push average offers and average MAO's up to values in the range found in experiments. Here, we complement that with the second approach. In Figure S1, below, the intensities of selection are fixed, and we look at the average offers and MAO's for a variety of mutation rates. For reasons explained in the main text, and in Section S2 of the supplementary material, we would like to stay away from the limit of weak selection, and therefore we choose the three larger intensities of selection that feature in Figure 1 in the main text.

In this representation, with fixed intensities of selection and increasing mutation rates, the simulations suggest the same as Figure 1 in the main text does, and maybe even more strongly so. If we compare the version from Rand et al.

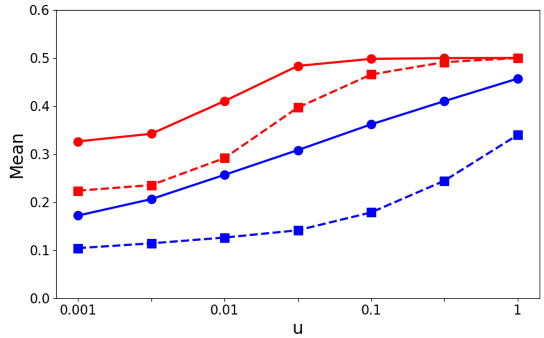
(2013), with global, biased mutations, to our version with local, and therefore much less biased mutations, then the average offers and the average MAO's are significantly lower in the latter. We also see that with local mutations, the average offer and the average MAO do not always reach the averages from experiments, even at the maximum mutation rate, where everybody always mutates.

With global mutations, both average offers and average MAO's inevitably get to 0.5 as mutation rates increase. The reason is that when mutations are global, then at $u = 1$, when both the offer and the MAO mutate at every reproduction event, it becomes irrelevant who is reproducing. The parents therefore stop passing on any (genetic) information; every new individual is a mutant, and all mutants are drawn from the same distribution, regardless of what the parents are. Therefore, at $u = 1$, on both dimensions, the population at any point in time just becomes a collection of independent random draws from $[0, 1]$.

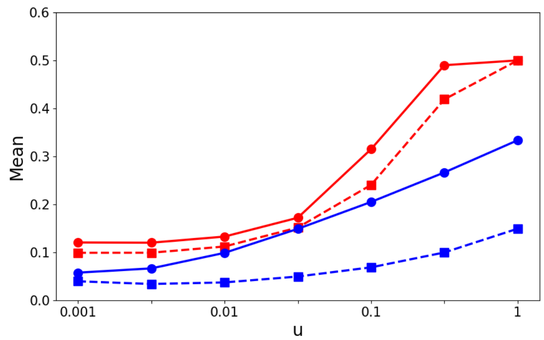
With local mutations, average offers and average MAO's do not necessarily get to 0.5 as the mutation rate increases. In this case, parents still pass on genetic information, because even if everyone mutates, these mutations are drawn from a distribution that is centered around the trait value of the parent. The trait value of the parent matters for payoffs, and therefore for the expected number of offspring, and that makes it possible for the average offer and the average MAO to stay below 0.5, even if the mutation rate is 1.

This illustrates that one can also push the average offers and the average MAO's up by increasing the mutation rate. It also illustrates that there are limits to how far one can push them up, and for moderate to high intensities of selection, even a mutation rates of 1 does not push them up high enough.

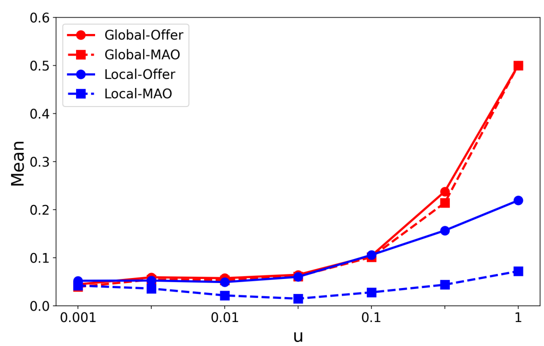
It is possible to model the genetics underlying the behaviour differently. If we for instance assume that there is a number of different loci that all can increase or decrease the offer or the MAO by a little bit, and we assume sexual reproduction, then it is possible that also with lower mutation rates, one can sustain similar levels of variation in the population, and thereby push the offers and MAO's up by the same amount. It should however be noted that this would naturally make mutations local, and therefore with global mutations, where the effect of the bias scales up with the mutation rate, there is less space to think of reasons why high mutation rates make sense.



(A)



(B)



(C)

Figure S1: Global versus local mutations. In red the average offers and MAO's for the model in Rand et al. (2013), which has global, co-occurring mutations. In blue the same, but for local, independent mutations. The intensity of selection is 0.1 in panel A, 1 in panel B, and 10 in panel C. For $w = 0.1$ one can still get to the averages observed in experiments, but with local mutations it requires very high mutation rates. For $w = 1$ and $w = 10$, even a mutation rate of $u = 1$ is not high enough for local mutations.

S1.4 Co-occurring versus independent mutations

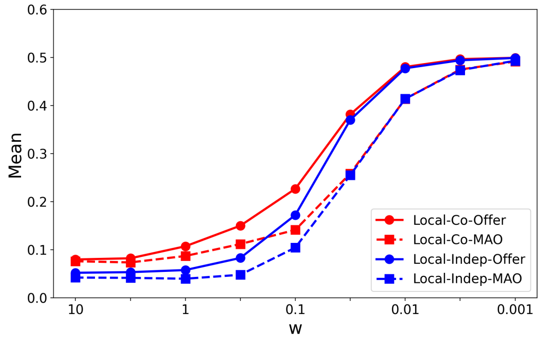
Figure S2 shows that only for a combination of strong selection and a low mutation rate is there a modest difference between local, co-occurring mutations and local, independent mutations. For higher mutation rates and/or weaker selection, this difference disappears. Because there is no real reason why mutations would co-occur, we chose to use the version with independent mutations. The comparison here is done to make sure that the lion share of the difference between simulations with the model from Rand et al. (2013) and simulations with ours is due to replacing global, and therefore biased mutations with local, and therefore much less biased ones, and not to switching from co-occurring to independent mutations.

S2 Weak selection

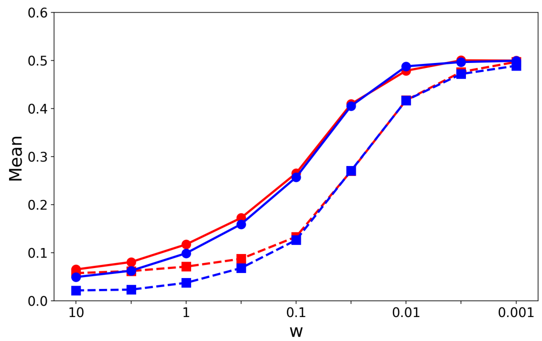
In the main text, we have seen that for a fixed mutation rate, we can always push average offers and the average MAO's up, from 0, to any point between 0 and 0.5, by reducing the intensity of selection. We have also seen that there are limitations to how much the intensity of selection can be reduced, and still produce a meaningful prediction. Here we will make that argument a bit more precisely and elaborately.

S2.1 Probabilistic symmetry

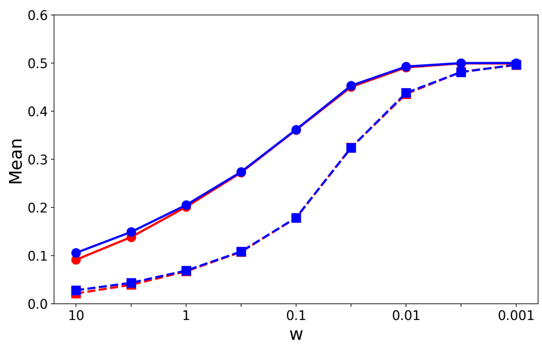
When the intensity of selection is 0, the dynamics in the model by Rand et al. (2013) become symmetric, in the sense that any transition from one population state to the other is equally likely as its mirror image. More precisely, if p_i denotes the offer of player i in the role of proposer, and q_i is the MAO of player i in the role of responder, then a population state is characterized by vectors $\mathbf{p} = [p_1, ..., p_N]$ and $\mathbf{q} = [q_1, ..., q_N]$, where N is the population size. Symmetry means that a transition from population state $(\mathbf{p}, \mathbf{q})$ to population state $(\mathbf{p}', \mathbf{q}')$ is equally likely as its mirror image, going from population state $(\mathbf{1} - \mathbf{p}, \mathbf{1} - \mathbf{q})$ to population state $(\mathbf{1} - \mathbf{p}', \mathbf{1} - \mathbf{q}')$, where $\mathbf{1}$ is a vector of 1's. This symmetry implies that if we average the population states over time, we will find a symmetric distribution. The average offer over this distribution therefore will be 0.5, and the average MAO will also be 0.5, and both of these are a consequence of the fact that 0.5 is halfway the strategy set that the population



(A)



(B)



(C)

Figure S2: Co-occurring versus independent mutations. In red the average offers and MAO's with local, co-occurring mutations, and in blue the same, but for local, independent mutations. The mutation rate is 0.001 in panel A, 0.01 in panel B, and 0.1 in panel C.

moves around in – with probabilistic symmetry.

All of this implies that the fact that it is possible to get the average offer or the average MAO up to any value between 0 and 0.5 by choosing a sufficiently low intensity of selection is not necessarily something that reflects anything to do with selection. Selection pulls both of them down, and if one reduces selection ever more, then one can reduce how much both are dragged down. The fact that one can get them to average at values arbitrarily close to 0.5 by almost eliminating selection, however, is a somewhat arbitrary result of the shape of the strategy set, and not of what selection does to the strategies in it.

In the main text we illustrate this by looking at a simulation run for global and infrequent mutation, and, of course, weak selection. Here in the supplementary material we will also consider global and frequent, local and infrequent, and local and frequent mutation.

S2.2 Weak selection, global mutation, low mutation rate

The left hand side of Fig. S3 depicts a few aspects of a run with global mutation, a low intensity of selection ($w = 0.001$), and a low mutation rate ($u = 0.001$). Panel A shows how the average offer and the average MAO change over time for a part of a simulation run. Panel C gives a snapshot of the distribution at some moment in time, and here we find both traits to be at fixation, as is expected to be the case for most of the time with such a low mutation rate. Panel E averages these distributions (like the one given in panel C) across time, producing the average distribution over time. As is to be expected, this is quite close to the uniform distribution on $[0, 1]$, which is the distribution that all mutants come from.

The fact that the average offer and the average MAO move around quite a bit over the course of a run limits the predictive power of the model for this combination of low intensity of selection and low mutation rate. Any average that we find in experiments would be close to the average in the simulations at some points in time, but it would be far away from the averages that the simulations produce at many other points in time. Also, at most points in time, there is not much variation; the variation in panel E is generated by the variability across time, not by the variation at any moment in time. The prediction of this model therefore is that we should observe a close to monomorphic population, where the probability with which we would observe a certain average is the result of a draw from the uniform distribution. The fact that the MAO of everyone in the population is regularly also above the offers of everyone in the population (almost 50% of the time) also implies that if we really believe in

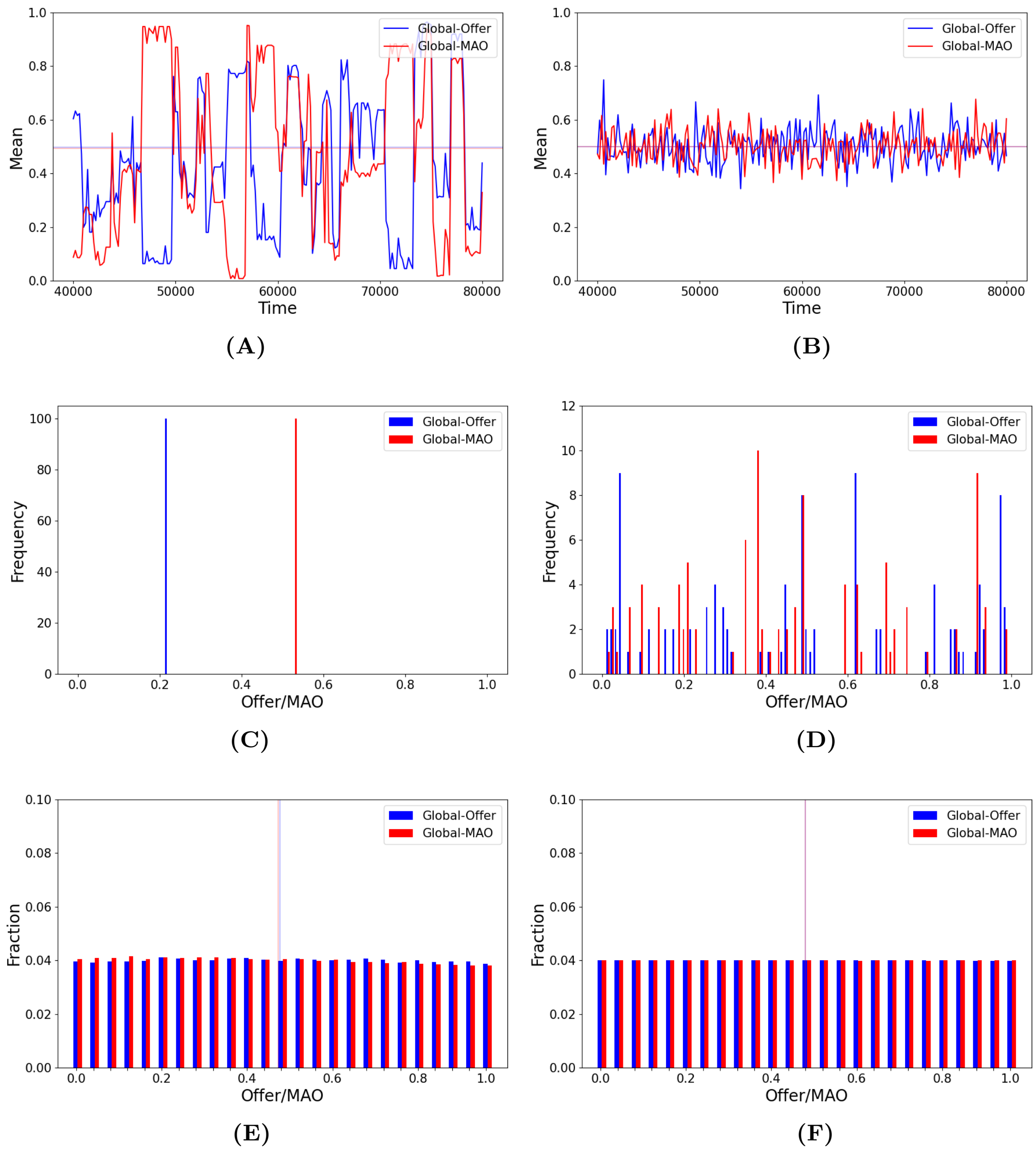


Figure S3: Global, co-occurring mutations, w = 0.001, and u = 0.001 (left) and u = 0.1 (right). The top panels give the average offer and MAO over time for a part of the run. The middle panels give the distribution of strategies at some random moment in the simulation run. The bottom panels give the average distribution over time, where we bundled strategies within intervals of length 0.04 together. The average offer of the average distribution and the average MAO of the average distribution are horizontal lines in panel A and B, and vertical lines in panel E and F.

weak selection, we should also conclude that if we now find the average MAO to be below the average offer, then this is just a coincidence, and it could also have been the other way around. That would make it very unlikely that between different populations they would correlate, and that the first always turns out to be below the second (Henrich and Boyd, 2001; Henrich et al., 2001; 2006).

S2.3 Weak selection, global mutation, high mutation rate

The right hand side of Fig. S3 depicts the same aspects for a run, also with global mutation, and also with a low intensity of selection ($w = 0.001$), but with a high mutation rate ($u = 0.1$). Here, the averages in the population do not move around as much, and the shape of the distribution at any point in time is relatively close to the distribution of the inflow of mutants, which is a uniform distribution over $[0, 1]$. Given the low intensity of selection, this makes sense. With much less variability over time, this produces a much sharper prediction: the distribution should be close to uniform on $[0, 1]$ at all times. This does not match the empirical evidence either, because the distributions that we find in experiments typically are not that close to uniform. Moreover, as before, the average offer and the average MAO move close to independently, and this does not predict the average offer to be above the average MAO.

S2.4 Weak selection, local mutation, low mutation rate

With weak selection, local mutations, and low mutation rates, the populations are typically also close to monomorphic, as they are with global mutations in combination with low mutation rates. Over time, they also move around quite a bit, and, similar to global mutations, not in synchrony. What is different is that changes in the averages come in much smaller steps, as a result of the mutations being local, and therefore the averages move around much slower (see Fig. S4A, where time runs 10 times faster than in Fig. S3A). The distribution over time is not the same as the "distribution that all mutants come from", as it is with global mutations, because with local mutations, there is no such thing as a constant mutant distribution. Because the average is a random walk, restricted to $[0, 1]$, the distribution still ends up looking like a uniform distribution, with some deviations at and close to the boundaries (see Fig. S4E). These boundaries are a bit stickier than other monomorphic states; trait values 0 and 1 collect more incoming mutations, that otherwise would have gone below 0 or over 1, and once the population is temporarily absorbed in one of the boundaries, and

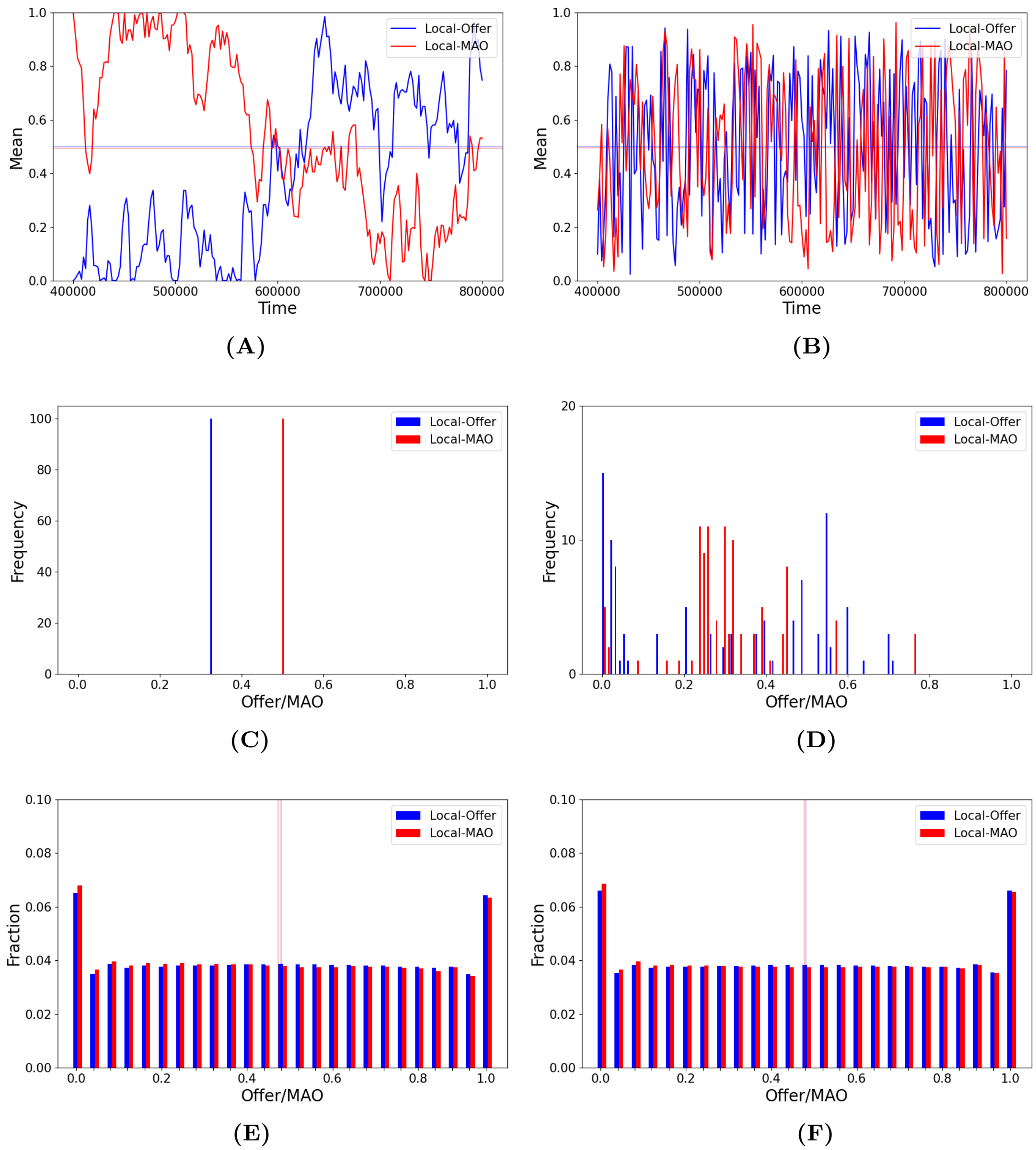


Figure S4: Local, independent mutations, w = 0.001, and u = 0.001 (left) and u = 0.1 (right). The top panels give the average offer and MAO over time for a part of the run. The middle panels give the distribution of strategies at some random moment in the simulation run. The bottom panels give the average distribution over time, where we bundled strategies within intervals of length 0.04 together. The average offer of the average distribution and the average MAO of the average distribution are horizontal lines in panel A and B, and vertical lines in panel E and F.

everyone has trait value 0, or everyone has trait value 1, leaving requires a mutant with the right sign. The population therefore spends some extra time at these extreme points.

S2.5 Weak selection, local mutation, high mutation rate

With weak selection and high mutation rates, the mutations being local rather than global allows random effects to persist for longer, because mutations are not biased towards the middle anymore – except for trait values at or close to 0 or 1. Deviations from the average over time therefore have much more amplitude than they do with global mutations (notice that also here, time is running 10 times faster in Fig. S4B than it is in Fig S3B). Compared to low mutation rates (Fig. S4D vs. Fig. S4C), the distribution at any given point in time is much less concentrated, and over time, the population moves around faster (Fig. S4B vs. Fig. S4A), but otherwise, also here the average (over time) of the averages (over the population) is a uniform distribution, with some deviations at the edges (see Fig. S4F).

S2.6 Variance over time

In order to have an indication of how stable or unstable the distributions are over time, we can calculate the variance in average offers, or the variance in average MAO's, over time. If $\overline{p}^t$ is the average offer in the population at time t, and $\overline{\overline{p}} = \frac{1}{T}\sum_{t=1}^{T}\overline{p}^t$ is the average over time of these averages over the population, then

$$\frac{1}{T}\sum_{t=1}^{T}\left(\overline{p}^t - \overline{\overline{p}}\right)^2$$

is the variance across time. Simulations with a low variance have more predictive power than simulations with a high variance.

As a benchmark of something that has no predictive power, one could use the variance that would go with randomly drawing a new average offer or MAO from a uniform distribution on $[0, 1]$ every period. In that case, the variance is

$$\int_0^1 (x - 0.5)^2 = \frac{1}{3}\left[(x - 0.5)^3\right]_0^1 = \frac{1}{12} \approx 0.083$$

Calculating these variances for the simulations paints a picture that is in line with what one would expect from Fig. S3 and Fig. S4; whether mutations are

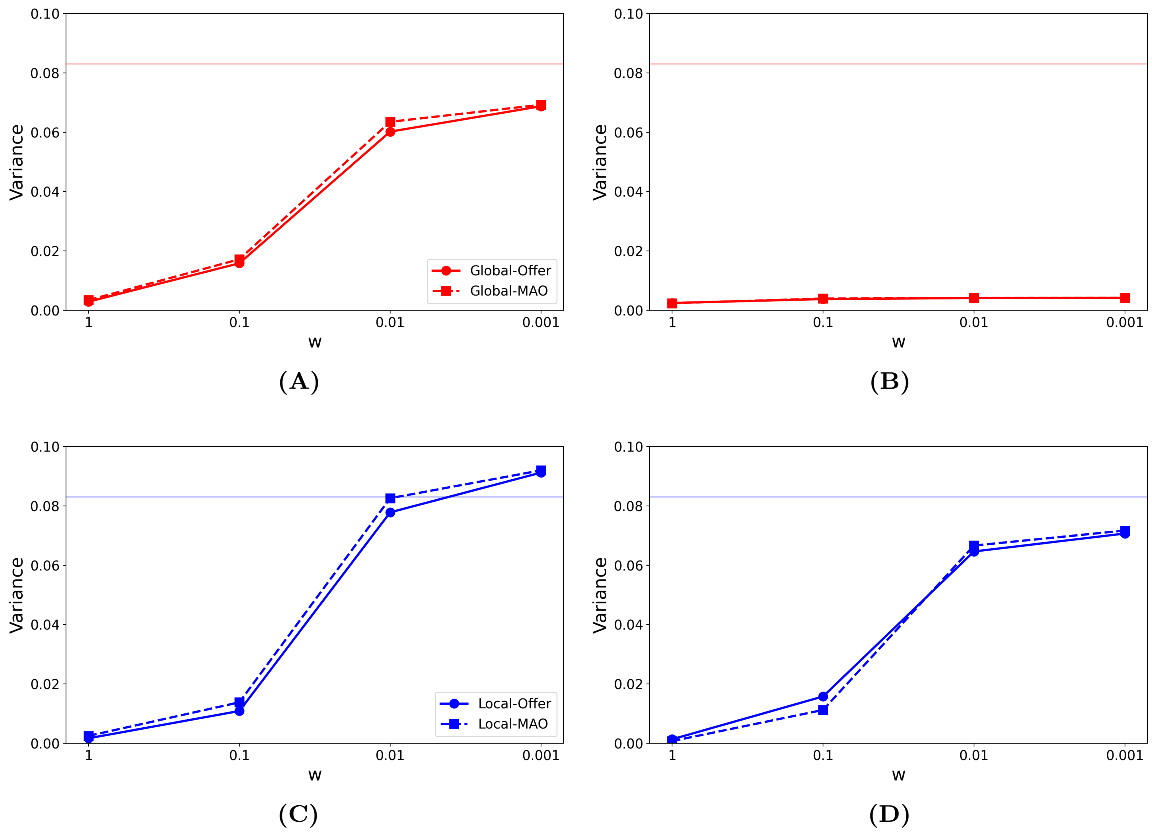


Figure S5: Variances over time for weak selection. With global, and rare mutations ($u = 0.001$), the variance over time gets very high for $w = 0.01$ and $w = 0.001$ (A). With global, and frequent mutations ($u = 0.1$), the average is very stable, and the variance stays low (B). With local, and rare mutations, the variance over time gets high again, even a bit higher than the variance one would get from independent draws from the uniform distribution – indicated in the figures by straight horizontal lines (C). With local, and frequent mutations, the variance also gets very high for weak selection (D).

global and rare; local and rare; or local and frequent, variances get very high when selection gets weak (see Fig. S5). Because the edges of the interval $[0, 1]$ are temporarily absorbing for local mutations, the variances there become even higher than $\frac{1}{12}$ when mutations are rare. The variance only remains low for mutations that are both global and frequent. In this case, the distribution at any point in time will be close to the distribution that the mutants come from.

This implies that with weak selection, the simulations are literally all over the place. For global and infrequent mutations, they are extremely variable over time; and for global and frequent mutations, they are extremely variable at any moment in time. We should, however, not have a model with global, and therefore biased mutations anyway – as we have seen that it is this bias that drives the results – and if we choose local, and therefore much less biased mutations, the averages are, again, all over the place for weak selection, this time regardless of the mutation rate. That implies that the trick to push average offers and average MAO's up by lowering the intensity of selection goes at the expense of predictive power; any population average that one would find at some point in time would literally be equally likely under the model.

S3 Infinite population models

S3.1 The model in Gale et al. (1995)

First we repeat the equations for the model in Gale et al. (1995). The amount to be divided is denoted by n. The share of proposers that propose i is denoted by x_i, for $i = 1, ..., n$, and the share of responders with an MAO of j is denoted by y_j, for $j = 1, ..., n$. The mutation-selection dynamics are given by

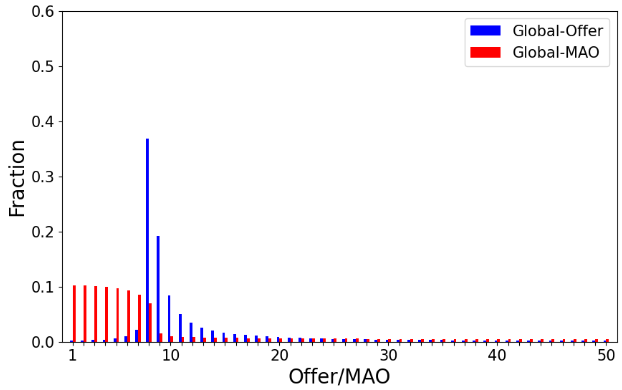
$$\dot{x}_i = (1-\delta)\left(\pi_{i,P} - \overline{\pi}_P\right) x_i + \delta\left(\frac{1}{n} - x_i\right)$$

for proposers using strategies $i = 1, ..., n$, where $\dot{x}_i$ the time derivative of x_i, δ is the mutation rate, $\pi_{i,P}$ is the payoff of proposers that propose i, and $\overline{\pi}_P$ is the average payoff in the proposer population, and by

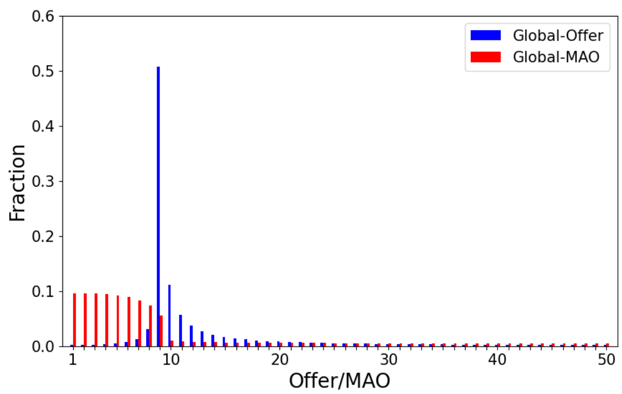
$$\dot{y}_j = (1-\delta)\left(\pi_{j,R} - \overline{\pi}_R\right) y_j + \delta\left(\frac{1}{n} - y_j\right)$$

for responders that use strategies $j = 1, ..., n$, where $\dot{y}_j$ is the time derivative of y_j, $\pi_{j,R}$ is the payoff of responders with an MAO of j, and $\overline{\pi}_R$ is the average payoff in the responder population.

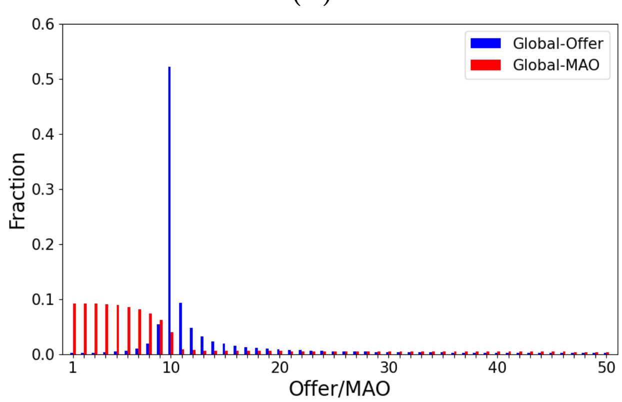
The first term on the right hand side reflects selection, the second term reflects mutation. Global mutation means that all strategies have the same inflow due to mutation (it is $\frac{\delta}{n}$ for all strategies) and an outflow that is proportional to the current shares (it is δx_i for proposers and δy_j for responders). Gale et al. (1995) allow for the mutation rate δ to differ between proposers and responders, but we will first assume that they are the same.



(A)



(B)



(C)

Figure S6: Multiple equilibria for Gale et al. (1995). With $\delta = 0.05$ and global mutation, there are 3 mutation-selection equilibria. The most frequent strategy for proposers in those equilibria ranges from $i = 8$ (A) to 10 (C), making the predominant offer in those equilibria range from $7\frac{1}{2}$ (A) to $9\frac{1}{2}$ (C).

S3.2 Our small changes to the version with global mutations

The offers and MAO's in the original model run from 1 to n, and exclude 0. When we compare this model, which has a discrete strategy space, with the model from Rand et al. (2013), that has a continuous strategy space, it can be nice to make strategies in the former comparable to strategies within an interval in the latter. We therefore shifted all proposals to the left by $\frac{1}{2}$; instead of having proposer strategy i propose i, we choose for strategy i to propose the midpoint of the interval $[i-1, i]$, which is $i - \frac{1}{2}$. Similarly, we let responder strategy j have an MAO of $j - \frac{1}{2}$. This only affects the equations above indirectly, in the sense that the payoff calculations now involve slightly shifted offers and MAO's. Below, we will sometimes still just refer to those strategies as strategy i or j, because that is shorter, but sometimes we will explicitly refer to the offer, and then we write $i - \frac{1}{2}$, or to the MAO, in which case we write $j - \frac{1}{2}$.

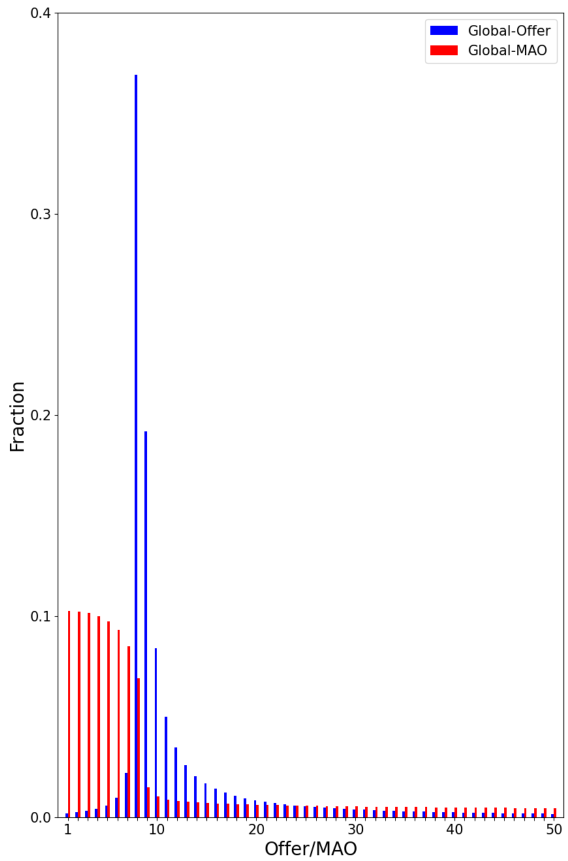
We also normalize the payoffs, so that the maximum payoff is 1 and the minimum payoff is 0. With normalization, one can see n, not as an indicator of the pie size, but as an indicator of how finely one unit can be subdivided. This helps comparing the results to simulations from Rand et al. (2013), which have a fixed amount of 1 to be divided in the ultimatum game.

S3.3 Multiple equilibria

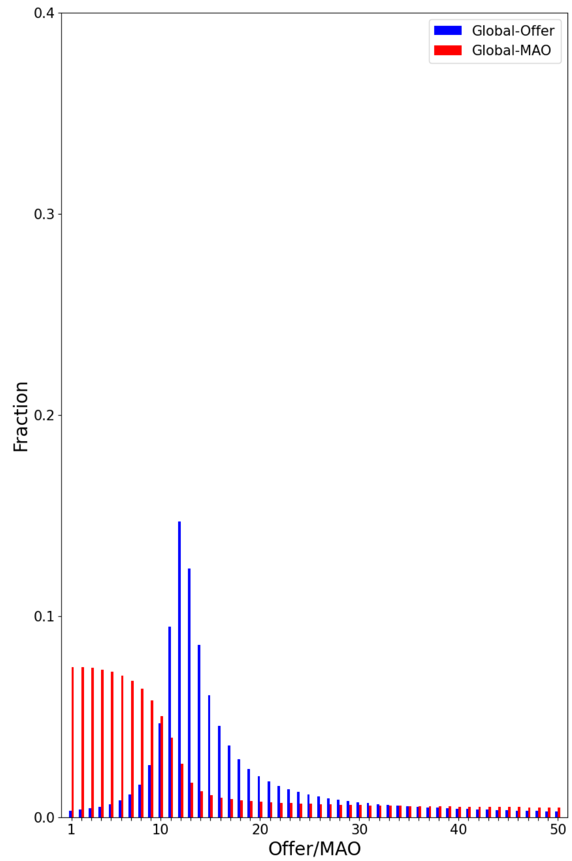
Figure S6 shows a variety of equilibria for the same combination of n and δ. All of those equilibria are similar, in that most proposers are making the same offer, with fewer proposers making higher offers, and even fewer making lower offers. Most responders have MAO's somewhere between the smallest possible MAO and the offer that most proposers make, and very few have larger MAO's. The different equilibria are characterized by what the most frequent offer is; for $n = 50$ and $\delta = 0.05$ − the parameters from Figure S6 − there are equilibria where the most frequent MAO is $7\frac{1}{2}$ (A), $8\frac{1}{2}$ (B), or $9\frac{1}{2}$ (C).[1]

The first parameter combinations in Figure S7 have multiple mutation-selection equilibria. The other three have unique, globally attracting mutation-selection equilibria. With the limited computing power of 1995, Gale et al.

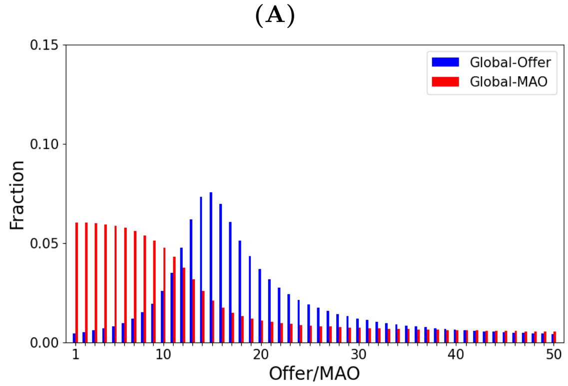
[1] For replicator dynamics with a continuous strategy space, one would expect a spectrum of equilibria, with positive point mass at some offer for proposers. These would be stable to perturbations that are small in the variational distance, but not to perturbations that are small in the Prohorov metric (see Van Veelen and Spreij, 2009).



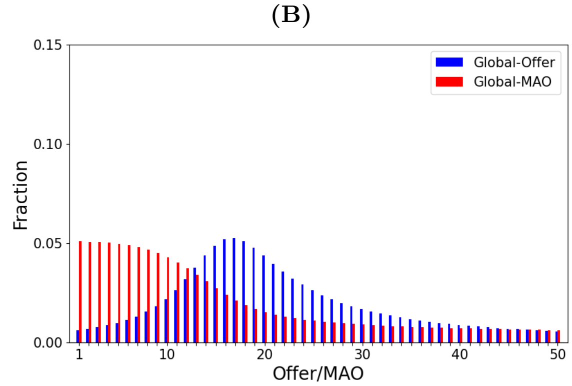
(A)



(B)



(C)



(D)

Figure S7: Mutation-selection equilibria in Gale et al. (1995) with global mutations. For $\delta = 0.05$ there are multiple equilibria (see Figure S6). We picked the first one for panel A. For $\delta = 0.075$ (B), $\delta = 0.1$ (C), and $\delta = 0.125$ (D), there is a unique, globally attracting mutation-selection equilibrium. The fat tails are a symptom of the bias in the mutations.

may have missed the possibility that the population might converge to different states depending on the starting point.

S3.4 Our version with local mutations

Also here, global mutations are biased, which is not a good basis for an explanation. Therefore, as we did with the model in Rand et al. (2013), we also made a version with local, and therefore much less biased mutations. In the local version, if an individual mutates that currently plays strategy i, then the new strategy becomes any strategy from $i-k$ to $i+k$, all with equal probability – provided that these changes do not make the offer or MAO drop below 0 or go over 1. The latter is guaranteed not to occur if $k < i \leq n-k$. If $i \leq k$, the mutant becomes any strategy from 2 to $i+k$ with probability $\frac{1}{2k+1}$, and strategy 1 with the remaining probability, and if $> n-k$, then the mutant becomes any strategy from $i-k$ to $n-1$ with probability $\frac{1}{2k+1}$, and strategy n with the remaining probability.

For $n = 50$ and mutations that take a mutant a maximum of $k = 2$ steps to the right or to the left, that means that the equations for the replicator dynamics for proposers become

$$
\begin{aligned}
\dot{x}_1 &= (1-\delta)\left(\pi_{1,P} - \overline{\pi}_P\right) x_1 + \delta\left(-\frac{2}{5}x_1 + \frac{2}{5}x_2 + \frac{1}{5}x_3\right) \\
\dot{x}_2 &= (1-\delta)\left(\pi_{2,P} - \overline{\pi}_P\right) x_2 + \delta\left(\frac{1}{5}x_1 - \frac{4}{5}x_2 + \frac{1}{5}x_3 + \frac{1}{5}x_4\right)
\end{aligned}
$$

for the first two, then

$$
\dot{x}_i = (1-\delta)\left(\pi_{i,P} - \overline{\pi}_P\right) x_i + \delta\left(\frac{1}{5}x_{i-2} + \frac{1}{5}x_{i-1} - \frac{4}{5}x_i + \frac{1}{5}x_{i+1} + \frac{1}{5}x_{i+2}\right)
$$

for strategies 3 to 48, and

$$
\begin{aligned}
\dot{x}_{49} &= (1-\delta)\left(\pi_{49,P} - \overline{\pi}_P\right) x_{49} + \delta\left(\frac{1}{5}x_{47} + \frac{1}{5}x_{48} - \frac{4}{5}x_{49} + \frac{1}{5}x_{50}\right) \\
\dot{x}_{50} &= (1-\delta)\left(\pi_{50,P} - \overline{\pi}_P\right) x_{50} + \delta\left(\frac{1}{5}x_{48} + \frac{2}{5}x_{49} - \frac{2}{5}x_{50}\right)
\end{aligned}
$$

for the last two. For responders, this is

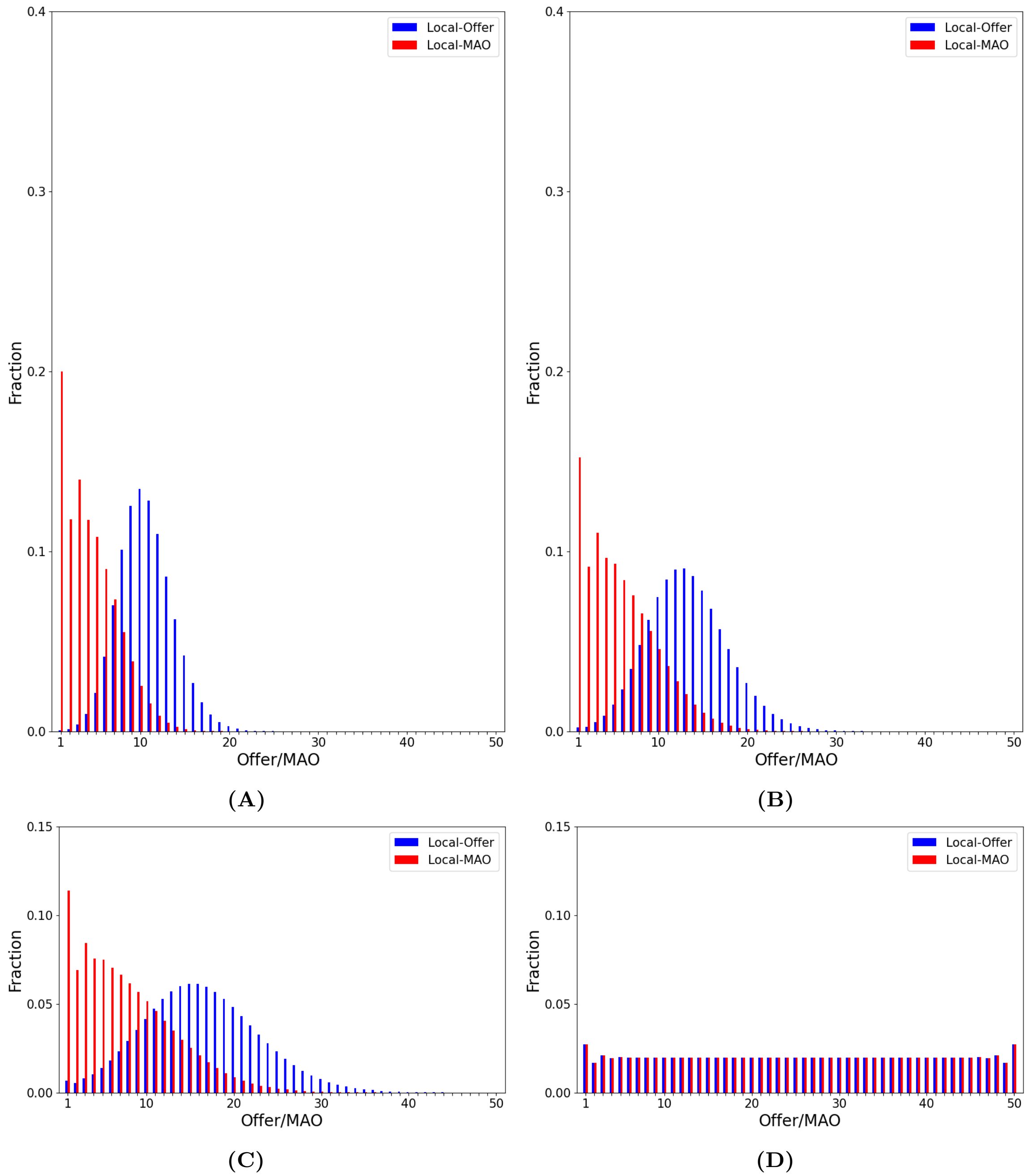


Figure S8: Mutation-selection equilibria in Gale et al. (1995) with local mutations. Mutation rates are $\delta = 0.25$ (A), $\delta = 0.5$ (B), $\delta = 0.75$ (C), and $\delta = 1$ (D). The wave pattern at the boundaries of the strategy space is caused by the remaining bias in mutations at the boundaries, where mutations to strategies below 0 or above 50 are ruled out.

$$\dot{y}_1 = (1-\delta)\left(\pi_{1,R} - \overline{\pi}_R\right) y_1 + \delta\left(-\frac{2}{5}y_1 + \frac{2}{5}y_2 + \frac{1}{5}y_3\right)$$
$$\dot{y}_2 = (1-\delta)\left(\pi_{2,R} - \overline{\pi}_R\right) y_2 + \delta\left(\frac{1}{5}y_1 - \frac{4}{5}y_2 + \frac{1}{5}y_3 + \frac{1}{5}y_4\right)$$

for the first two, then

$$\dot{y}_j = (1-\delta)\left(\pi_{j,R} - \overline{\pi}_R\right) y_j + \delta\left(\frac{1}{5}y_{j-2} + \frac{1}{5}y_{j-1} - \frac{4}{5}y_j + \frac{1}{5}y_{j+1} + \frac{1}{5}y_{j+2}\right)$$

for strategies 3 to 48, and

$$\dot{y}_{49} = (1-\delta)\left(\pi_{49,R} - \overline{\pi}_R\right) y_{49} + \delta\left(\frac{1}{5}y_{47} + \frac{1}{5}y_{48} - \frac{4}{5}y_{49} + \frac{1}{5}x_{50}\right)$$
$$\dot{y}_{50} = (1-\delta)\left(\pi_{50,R} - \overline{\pi}_R\right) y_{50} + \delta\left(\frac{1}{5}y_{48} + \frac{2}{5}y_{49} - \frac{2}{5}y_{50}\right)$$

for the last two.

Figure S8 shows mutation-selection equilibria for global mutation and a variety of mutation rates. Many observations made when comparing global and local mutations for Rand et al. (2013) can also be made here. The most obvious one is that when comparing Figures S7 and S8, we see that also here, all else equal, average offers and MAO's are lower with local than with global mutation (note that mutation rates in Figure S8 are higher than in Figure S7).

Because mutants have to remain within the strategy space, we assumed that mutations to strategies below 0 are replaced with mutations to 0, and mutations to strategies above n are replaced by mutations to n. That means that 0 and n have extra incoming mutations, while, in our case, with a maximum change in strategy of 2 due to mutation, strategies 1 and $n-1$ only have a reduced amount of incoming mutations, since strategies below 0 or above n that could mutate to 1 and $n-1$, respectively, do not exist. In equilibrium, this creates a spike at 1, a valley at 2, and those also ripple through the frequencies towards the middle. In panel (D) we see the same at the top end of the strategy space.

Finally, a difference between Rand et al. (2013) with local mutations and Gale et al. (1995) with local mutations, is that we have seen that even at a mutation rate of 1, the average offer and the average MAO are not $\frac{1}{2}$ in Rand et al. (2013) with local mutations. In Figure S8 we see that for the version of Gale et al. (1995) with local mutations, this is not true, and both averages are in fact $\frac{1}{2}$. This is caused by the difference in how reproduction events and mutations relate in both models. In Rand et al. (2013), mutations happen at reproduction. That means that at a mutation rate of 1, reproductions still happen, but at every one of those, a mutation occurs. With local mutations, that means that the trait value of the offspring is still correlated with the trait value of the parents (which is not true for global mutations). In Gale et al. (1995), mutations happen not at reproduction. Instead, the mutation rate reflects how many mutation events occur relative to the number of reproduction events. That means that here, at a mutation rate of 1, there are only mutations, and reproduction is just not happening.

S4 Link between Rand et al. (2013) and Gale et al. (1995)

For infinitely large population dynamics, if a population is in equilibrium, it does not move. In finite population dynamics, also in equilibrium, the population moves around, but visits some states (much) more often than others. With finite population dynamics, the equilibrium therefore is a distribution over population states. As the population size increases, the noise decreases, and the variation in population states across time goes down. In the limit of infinitely large populations, the dynamics become deterministic.

In order to investigate the link between the finite population model in Rand et al. (2013) and the infinite population model in Gale et al. (1995), we ran simulations with increasing population size for Rand et al. (2013). When we compare snapshots from the population with the average distribution (over time), we find that the difference between these two does indeed decrease with population size – which is an indication that the population does indeed move around less. For global mutation and a population size of 100, the difference between the snapshot and the average distribution (which averages these snaphots over time) is very large, and also for a population size of $1,000$ it is still considerably different, and only at a population of $10,000$, they come close. The characteristics of the average distribution therefore are not perfectly representative for the average characteristics of the distribution at any given point in time, although much more at $10,000$ than the other population sizes. The variance within the population at any moment in time is also smaller than the variance in the average distribution. This difference also goes down, but is quite substantial at 100 and $1,000$. One could therefore say that the infinite population model in Gale et al. (1995) is only a good approximation of the finite population model in Rand et al. (2013) for quite large finite populations. The discrepancies are however smaller for our versions of the two with local mutations (not depicted).

S5 Quantal Response Equilibria

S5.1 Predictions and data

S5.1.1 Data

Author, Year	# Obs. all stakes	# Obs. low/medium	DR vs. STR
Andersen et al. (2011)	458	325	DR
Bader et al. (2021)	485	–	STR
Bahry and Wilson (2006)	288	–	STR
Barmettler et al. (2012)	100	–	DR
Benndorf et al. (2017)	98	–	STR
Bornstein and Yaniv (1998)	20	–	DR
Cameron (1999)	202	165	DR
Carpenter et al. (2005a;b)	107	–	DR
Chew et al. (2013)	207	–	STR
Croson (1996)	56	–	DR
Demiral and Mollerstrom (2020)	283	–	STR
Forsythe et al. (1994)	67	–	DR
Inaba et al. (2018)	121	–	STR
Keuschnigg et al. (2016)	487	–	STR
Lightner et al. (2017)	42	–	DR
Peysakhovich et al. (2014)	576	–	STR
Ruffle (1998)	44	–	DR
Slonim and Roth (1998)	820	570	DR

Table S1: Alphabetical list of empirical papers whose data we use. The second column shows the number of observations after eliminating the treatments using the non-standard versions of the ultimatum game, for studies with a variety of stake sizes, the third column shows the number of observations after excluding also the treatments with the largest stake sizes in studies testing the stake size effects, and the last column shows whether the experimental design uses the direct-response (DR) or the strategy (STR) method.

We use the data from the papers listed in Table S1 in our main text and

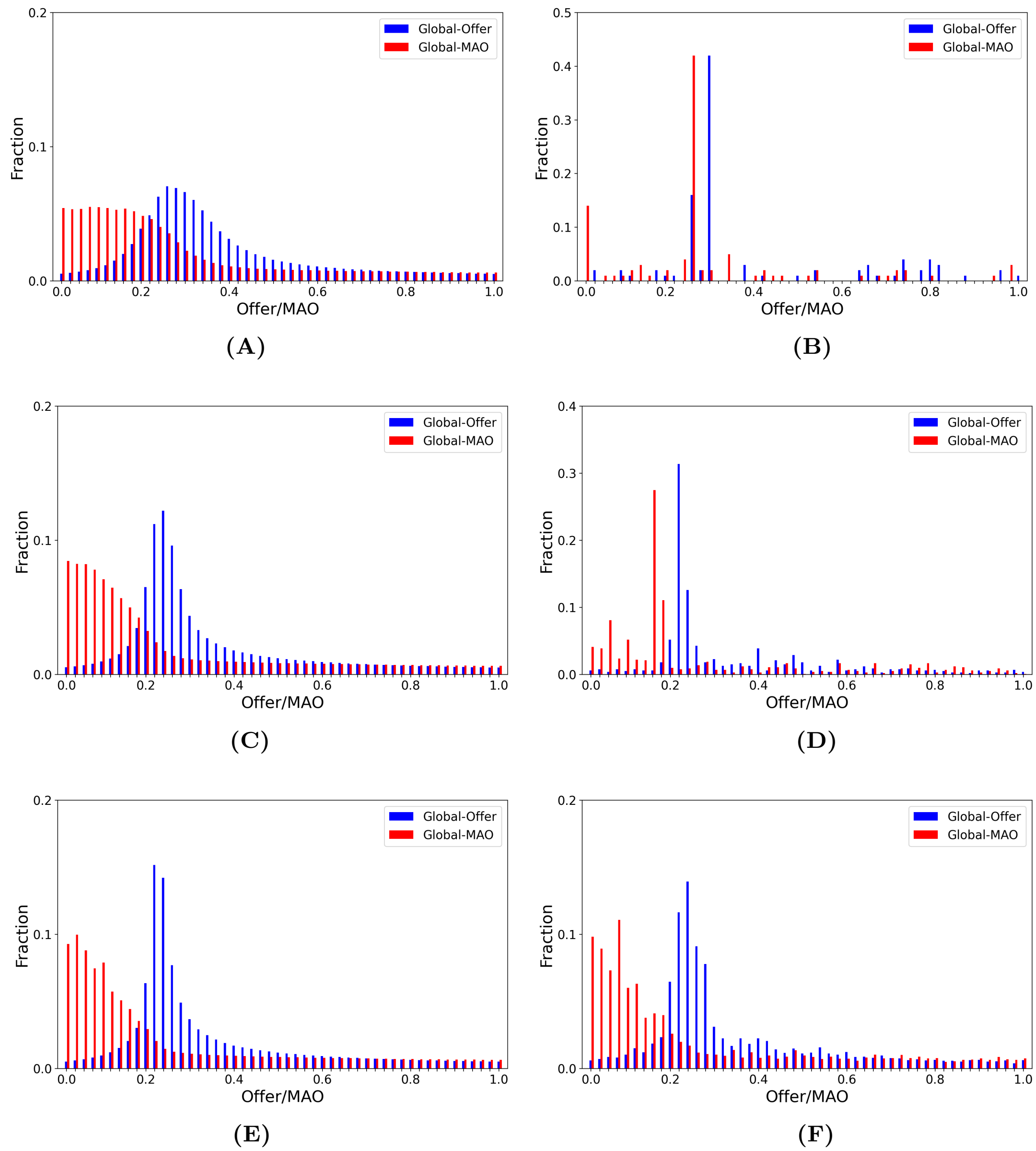


Figure S9: Rand et al. (2013) for different population sizes. The mutation rate is 0.125, the intensity of selection is 1, and the population size is 100 (top panels), $1,000$ (middle panels), and $10,000$ (bottom panels). Panels A, C, and E show the average distributions of offers and MAO's, where the average is taken over time. Panels B, D, and F show snapshots of offers and MAO's. The scaling on the horizontal axes in B and D is different from the other panels. Because running additional generations becomes rather expensive at a population size of 10,000, the distribution in E is a bit noisier than in A and C.

	(1)	(2)
	P(*accept*)	P(*accept*)
Offer	4.778***	7.463***
	(p<0.001)	(p<0.001)
Intercept		-1.035***
		(p<0.001)
Observations	1496	1496
Log-likelihood	-600.78098	-580.5621
AIC	1203.562	1165.124
BIC	1208.873	1175.745
Pearson's χ^2	190.56	302.69
LR-test	40.44 (p<0.001)	

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table S2: Logistic regression results with and without the intercept.

in the supplementary material. In addition to these, we use the data from the meta-analysis of Tomlin (2015) in Figure S10.[2] For both the direct-response method and the strategy method we calculate the outcome variable proportional to the total amount available in the ultimatum game to make observations from different studies comparable as much as allowed by the grid used in the experiments. All of the studies use real monetary stakes.

In studies that use the direct-response method we consider how often an offer is accepted out of the number of times that offer is made, as an estimate of its acceptance probability. In studies that use the strategy method we analyze the MAO's reported by participants, or their accept/reject decisions for each possible offer level, as an estimate of the acceptance probability for every possible offer.

S5.1.2 Logistic regression with and without the intercept

Since the *Agent*-QRE is equivalent to the logit specification without the intercept, we ran logit regressions with and without the intercept to test between the two specifications. Table S2 presents the results. As can be seen from Column (2), the coefficient on the intercept is highly statistically significant. In line with this, the AIC, BIC, Pearson's χ^2 criteria, and the likelihood-ratio (LR) test all

[2]Since the dataset in Tomlin (2015) does not include information on stake sizes, we excluded those observations in our analysis for the main text.

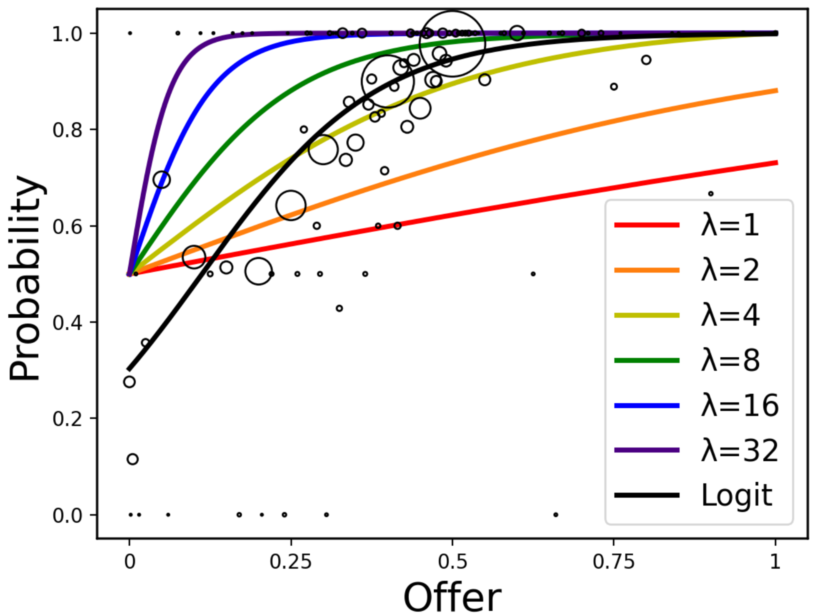


Figure S10: Acceptance/rejection rates in *agent*-QRE vs. empirical acceptance/rejection rates. The coloured lines are the acceptance rates in the *agent*-QRE for different λ's. The circles indicate acceptance rates for different proposals, pooling data from a number experiments together. Their size reflects the number of observations for that offer. The black line is the fitted acceptance rate as a function of the offer for a logit regression. Here, we use data obtained with the direct-response method, including all stake sizes.

indicate that the introduction of the intercept term in Column (2) improves the model fit significantly.

S5.1.3 With and without large stakes

In the main text, we compared the predictions of the *Agent*-QRE with rejection rates we calculated by pooling data from experiments that use the direct-response method. Because it is not universally agreed upon whether stake size matters, we excluded the observations for the largest stakes in the papers by Andersen et al. (2011); Cameron (1999); Slonim and Roth (1998).[3] If we include all stake sizes, the pattern is similar (see Figure S10).

[3]In experiments that use three levels of stake sizes we exclude the largest one; in experiments that use four levels of stake sizes we exclude the largest two. As Carpenter et al. (2005a) uses stake sizes of only \$10 and \$100, we include both stakes in our analysis in the main text.

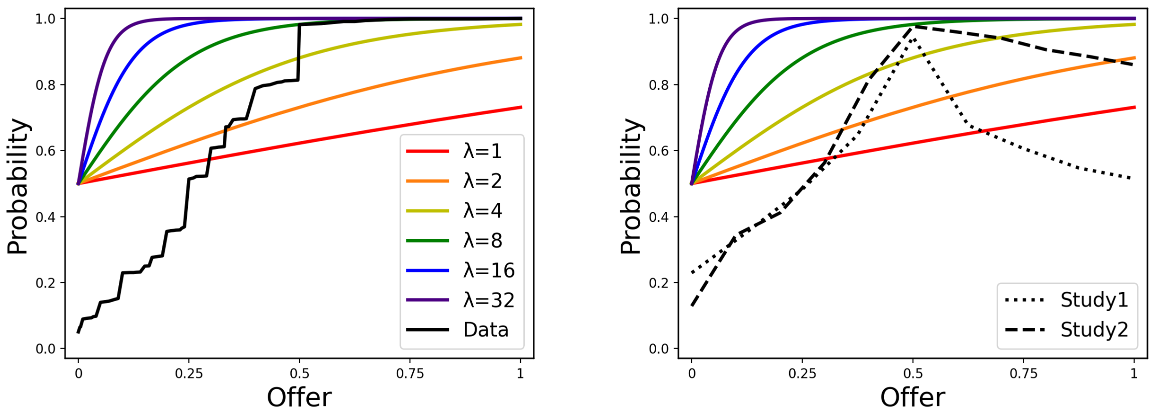
S5.1.4 Strategy method data for *Agent*-QRE

In the main text, we compared the predictions of the *Agent*-QRE with rejection rates that we calculate by pooling data from experiments that use the direct response method. This is the more natural thing to do, but one can also construct rejection rates from experiments that use the strategy method. The majority of studies that use the strategy method restrict subjects to strategies that can be characterized with an MAO; they exclude strategies for which there exist two offers, where the higher one is rejected, and the lower one is accepted. This restriction by construction implies that we will find that rejection rates are never decreasing in the offer. The prediction that rejection rates increase therefore cannot be tested with the experiments that use the strategy method in this way. The prediction that all acceptance rates should be above 50% can be tested, and would be rejected with data from experiments that use the strategy method – as it is with experiments that use the direct-response method (see the left panel in Figure S11).

Three of the studies in our sample (Bader et al., 2021; Bahry and Wilson, 2006; Keuschnigg et al., 2016) do not restrict subjects to submitting an MAO. They instead ask their participants to submit their accept/reject decisions for each possible offer level, which allows them to freely switch between accepting and rejecting. For these studies we include the participants who never switch, who switch only once (those that start with rejecting and switch to accepting at a certain offer level), and those who switch twice (from rejecting to accepting in the first half of the strategy space of offers, and from accepting back to rejecting in the second half of the strategy space). We exclude participants who do not fall into one of these categories. The number of observations given in Table S1 represents the number of observations after excluding these participants. Using the data from these studies we can test both the prediction that rejection rates monotonically increase in the offer and the prediction that all acceptance rates should be above 50%; and both would be rejected (see the right panel in Figure S11).

S5.2 *Agent*-QRE and learning models

In the *Agent*-QRE, a higher offer is accepted with higher probability than a lower offer. The underlying assumption is that the noise in the perception of what the payoff-maximizing thing to do is, is the same for all proposals. Combined with the fact that the payoff difference between accepting and rejecting gets larger



(A) **(B)**

Figure S11: Acceptance/rejection rates in *agent*-QRE vs. empirical acceptance/rejection rates under the strategy method. The coloured lines are the acceptance rates in the *agent*-QRE for different λ's. The black line in panel A indicates acceptance rates for different proposals, pooling data from a number experiments together that use the strategy method and ask subjects to submit an MAO. The two dotted lines in panel B indicate acceptance rates for different proposals, for the three experiments that use the strategy method and ask subjects to submit their accept/reject decision for each possible offer within their grid. Study1 is Bahry and Wilson (2006), and Study2 combines the data from Keuschnigg et al. (2016) and Bader et al. (2021) as they have an identical design.

when proposals increase, and therefore selection against rejecting also becomes stronger, this leads to the probability of accepting the offer being larger for larger offers and smaller for smaller offers.

The assumption of constant noise can be realistic, but one can also imagine that there are models for which this does not hold. For instance, one can also assume that the noise is higher for proposals that are made less frequently, and, depending on how the increased noise balances against the increased payoff difference, the rejection rate of a higher offer, that is made less frequently, could also end up being higher than that of a lower offer, that is made more frequently.

Also more in general, for models of selection that fit the setup of an *Agent*-QRE, where responses to different offers evolve separately, the property that higher offers get lower rejection rates may not universally hold. If we think of a model where strategies for different proposals do indeed evolve independently, then one could imagine that there is more selection happening for proposals that are made more frequently. If the mutation rate for responses to different offers is the same, one can imagine population states for which high offers are made so infrequently, that selection against rejections there is weak, and the rejection rate ends up being higher than that of a lower offer, that is made more frequently, and where there is more selection undoing the effect of mutations.

It is good to keep in mind, though, that this is only a detail, and perhaps a reason to prefer the *Normal form*-QRE over the *Agent*-QRE, but not an escape from the fact that the data would reject these models too. While the property of the *Agent*-QRE (higher x are always accepted with higher probability) might not carry over to all learning models or mutation-selection models that treat strategies for all offers separately, the property that acceptance rates should all be larger than 50% for all positive offers does. It is this property that is clearly violated by the data.

References

Steffen Andersen, Seda Ertaç, Uri Gneezy, Moshe Hoffman, and John A List. Stakes matter in ultimatum games. *American Economic Review*, 101(7):3427–39, 2011.

Felix Bader, Bastian Baumeister, Roger Berger, and Marc Keuschnigg. On the transportability of laboratory results. *Sociological Methods & Research*, 50 (3):1452–1481, 2021.

Donna L Bahry and Rick K Wilson. Confusion or fairness in the field? rejections in the ultimatum game under the strategy method. *Journal of Economic Behavior & Organization*, 60(1):37–54, 2006.

Franziska Barmettler, Ernst Fehr, and Christian Zehnder. Big experimenter is watching you! anonymity and prosocial behavior in the laboratory. *Games and Economic Behavior*, 75(1):17–34, 2012.

Volker Benndorf, Claudia Moellers, and Hans-Theo Normann. Experienced vs. inexperienced participants in the lab: Do they behave differently? *Journal of the Economic Science Association*, 3(1):12–25, 2017.

Gary Bornstein and Ilan Yaniv. Individual and group behavior in the ultimatum game: are groups more "rational" players? *Experimental Economics*, 1(1): 101–108, 1998.

Lisa A Cameron. Raising the stakes in the ultimatum game: Experimental evidence from indonesia. *Economic Inquiry*, 37(1):47–59, 1999.

Jeffrey Carpenter, Eric Verhoogen, and Stephen Burks. The effect of stakes in distribution experiments. *Economics Letters*, 86(3):393–398, 2005a.

Jeffrey P Carpenter, Stephen Burks, and Eric Verhoogen. Comparing students to workers: The effects of social framing on behavior in distribution games. In *Field experiments in economics*. Emerald Group Publishing Limited, 2005b.

Soo Hong Chew, Richard P Ebstein, and Songfa Zhong. Sex-hormone genes and gender difference in ultimatum game: Experimental evidence from china and israel. *Journal of Economic Behavior & Organization*, 90:28–42, 2013.

Rachel TA Croson. Information in ultimatum games: An experimental study. *Journal of Economic Behavior & Organization*, 30(2):197–212, 1996.

Elif E Demiral and Johanna Mollerstrom. The entitlement effect in the ultimatum game–does it even exist? *Journal of Economic Behavior & Organization*, 175:341–352, 2020.

Robert Forsythe, Joel L Horowitz, Nathan E Savin, and Martin Sefton. Fairness in simple bargaining experiments. *Games and Economic behavior*, 6(3):347–369, 1994.

John Gale, Kenneth G Binmore, and Larry Samuelson. Learning to be imperfect: The ultimatum game. *Games and Economic Behavior*, 8(1):56–90, 1995.

Joseph Henrich and Robert Boyd. Why people punish defectors: Weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas. *Journal of Theoretical Biology*, 208(1):79–89, 2001.

Joseph Henrich, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis, and Richard McElreath. In search of homo economicus: behavioral experiments in 15 small-scale societies. *American Economic Review*, 91 (2):73–78, 2001.

Joseph Henrich, Richard McElreath, Abigail Barr, Jean Ensminger, Clark Barrett, Alexander Bolyanatz, Juan Camilo Cardenas, Michael Gurven, Edwins Gwako, Natalie Henrich, et al. Costly punishment across human societies. *Science*, 312(5781):1767–1770, 2006.

Misato Inaba, Yumi Inoue, Satoshi Akutsu, Nobuyuki Takahashi, and Toshio Yamagishi. Preference and strategy in proposer's prosocial giving in the ultimatum game. *PloS One*, 13(3):e0193877, 2018.

Marc Keuschnigg, Felix Bader, and Johannes Bracher. Using crowdsourced online experiments to study context-dependency of behavior. *Social Science Research*, 59:68–82, 2016.

Aaron D Lightner, Pat Barclay, and Edward H Hagen. Radical framing effects in the ultimatum game: the impact of explicit culturally transmitted frames on economic decision-making. *Royal Society Open Science*, 4(12):170543, 2017.

Alexander Peysakhovich, Martin A Nowak, and David G Rand. Humans display a 'cooperative phenotype' that is domain general and temporally stable. *Nature Communications*, 5(1):1–8, 2014.

David G Rand, Corina E Tarnita, Hisashi Ohtsuki, and Martin A Nowak. Evolution of fairness in the one-shot anonymous ultimatum game. *Proceedings of the National Academy of Sciences*, 110(7):2581–2586, 2013.

Bradley J Ruffle. More is better, but fair is fair: Tipping in dictator and ultimatum games. *Games and Economic Behavior*, 23(2):247–265, 1998.

Robert Slonim and Alvin E Roth. Learning in high stakes ultimatum games: An experiment in the slovak republic. *Econometrica*, pages 569–596, 1998.

Damon Tomlin. Rational constraints and the evolution of fairness in the ultimatum game. *PloS One*, 10(7):e0134636, 2015.

Matthijs Van Veelen and Peter Spreij. Evolution in games with a continuous action space. *Economic Theory*, 39(3):355–376, 2009.