EGALITARIANISM AND EVOLUTION

CHRISTIAN TORSELL

ABSTRACT. An agent is *inequity averse* if their preferences reflect a distaste for material inequality. Inequity aversion is of both scientific and philosophical interest. In behavioral economics, it is central to popular explanations of "anomalous" behavior in resource-division games. In social and political philosophy, it is of interest because it embodies an egalitarian approach to distributive bargaining. We use the *indirect evolutionary approach*, an extension of standard evolutionary game theory, to consider whether inequity aversion can survive competition with selfishness under natural dynamics like social learning and Darwinian evolution. In particular, we study evolution in populations playing a variant of the Nash demand game, a popular model of bargaining. Unlike related extant work, we assume that individuals play the base game repeatedly within their lifetimes, learning as they do. The central result is that, under a broad range of assumptions about the evolutionary dynamics and the initial population state, inequity aversion goes to fixation in populations comprising both inequity averse and selfish agents. Incorporating learning turns out to be crucial to the evolutionary result. The model thus supports two conclusions. First, egalitarian attitudes can survive and proliferate in competition with selfish ones. Second, in applying the indirect evolutionary approach, integrating intragenerational learning can influence the evolutionary dynamics in important ways.

1. INTRODUCTION

Fairness is a topic of perennial philosophical interest. It has also attracted significant attention in economics, much of which has been directed toward experiments on resource-division games. In these experiments, subjects play games in which their choices determine how much of some material good—typically money—each will receive. Although there is significant cultural variation in the behavior observed in such games (Henrich et al. 1991), one cross-culturally robust finding is that subjects' choices routinely seem to violate the assumption that players choose so as to maximize their own material gain. In some games, subjects consistently turn down free money (see Güth and Kocher 2014). In others, they give away money they could instead keep without consequence (Engel 2011). In still others, they show a strong bias favoring equal splits that cannot be explained by appealing to standard assumptions about rationality and selfishness alone (Skyrms 1996, ch. 1).

I would like to thank Cailin O'Connor, Jeff Barrett, Brian Skyrms, Justin Bruner, Joerg Oeschssler, and an anonymous referee for their helpful comments on earlier versions of this paper. Audiences at the 2024 PPE Society Annual Meeting and the Social Dynamics Seminar at UC Irvine also provided valuable feedback, for which I am grateful.

These and similar findings have been interpreted as evidence that human choice reflects a concern for more than material self interest: we also care about *fairness*. Although this is a familiar feature of our psychology, it raises puzzles for economists and biologists, since fair behavior is often costly—for example, in ultimatum games, where acting fairly sometimes involves turning down freely-offered resources.[1] What explains the prevalence of a psychological disposition that appears to straightforwardly conflict with its bearer's material interests?

One approach to explaining our concern for fairness uses tools from evolutionary game theory to identify conditions under which fair behavior is likely to be selected for under biological or cultural evolution. Evolutionary game theoretic models consider large populations of individuals each pre-programmed to choose some strategy in a base game. Members of the population repeatedly play the base game with randomly assigned partners. Over time, some individuals reproduce and others die off, with the Darwinian fitness for each strategy being proportional to the average payoff for its adopters in the base game, relative to the current population state. Evolution is modeled with a selection dynamics that describes rates of change for the proportion of the population playing each strategy in terms in terms of that strategy's fitness. Random mutation might also be modeled. Standard evolutionary game theoretic analysis is concerned with identifying population states that are evolutionarily significant because they have special stability properties or are reliably reached and settled on from a large range of initial population states.

A well-known example from the evolution of fairness literature is Nowak et al. (2000). They show that fair behavior in the ultimatum game can evolve in populations where proposers have access to information about the past play of responders. The evolutionary success of fair behavior is due to fairminded responders developing a reputation for rejecting small offers which induces Responders to make larger offers against them than they would against players with a history of accepting low offers. For an evolutionary game theoretic approach to fairness in the Nash demand game, a classic reference is Skyrms (1996). Skyrms uses computer simulations to compare the evolutionary significance of the various stable population states for the Nash demand game under the *replicator dynamics*, a popular model of evolution in games. He finds that, under a broad range of conditions, the state in which all players demand half evolves from a larger share of initial states than the other asymmetric stable states (Skyrms 1996, 20-22).

In these models, selection operates directly on strategies in the base game. The standard interpretation, mentioned above, is that members of the population are equipped with all-or-nothing dispositions to choose a particular strategy. What if we want to model populations of more sophisticated agents, whose behavior is dictated flexibly by their preferences and

[1]For an overview of the puzzles concerning rationality and evolution raised by these findings, see Debove et al. (2016)

beliefs? In considering fairness, one motivation for developing such a model is that a popular approach explains fair behavior as a consequence of preferences that reflect a “taste for fairness” (see, e.g., Fehr and Schmidt 1999, Fehr and Schmidt 2002, Bolton and Ockenfels 2000). On such accounts, fair behavior is explained as expected-utility-maximizing choice for an agent whose utilities are sensitive to fairness-relevant features of distributive outcomes, in addition to their own material interests. How might we model the evolution of these sorts of preferences?

The *indirect evolutionary approach*, introduced in Güth and Yaari (1992), provides a framework for doing just this. The key innovation, explicated in greater detail in the next section, is to define a base game in terms of two distinct payoff functions: first, the players’ *subjective utilities*, which represent the profiles of preferences on which selection operates, and, second, the underlying *material payoffs* the players receive under each strategy profile, which are assumed to be proportional to Darwinian fitness. While players are assumed to choose *intra*generationally according to their subjective payoffs, the *inter*generational rate of change for each preference type’s share of the population is determined by the average material payoff to players with those preferences.

The present project uses the indirect evolutionary approach to model the evolution of a taste for fairness in the context of the Nash demand game. In the model, a taste for fairness is operationalized in terms of *inequity aversion*, that is, preferences that disfavor material inequality among players. We find that, in populations comprising both inequity averse agents and materially self-interested ones, a large class of selective dynamics strongly favors the inequity averse type. More precisely, the inequity averse type will take over the population under any payoff monotone dynamics, as long as the inequity-averse type has non-zero proportional representation in the initial population state. This result is further supported by simulations of evolution in finite populations, reported in the Appendix.

As is discussed in the next section, applying the indirect evolutionary approach framework to issues of fairness is not novel. But unlike extant work in this vein, my project uses a model of inductive learning in games (specifically, fictitious play) to estimate the fitnesses of different types of preferences in a given population state. This is in contrast to the standard approach, which uses solution concepts from classical game theory—in particular, the Nash equilibrium and its refinements—to predict intragenerational play and calculate the average material payoffs players with different kinds of preferences will receive in different population states. The central evolutionary result, mentioned above, depends crucially on the use of a learning model to estimate intragenerational payoffs. There is a methodological moral here: in the context of the indirect evolutionary approach, considering intragenerational learning can have important consequences for a model’s evolutionary results.

Section 2 introduces the indirect evolutionary approach and reviews extant applications of that approach to issues of fairness. Section 3 defines the base bargaining game, as well as the two kinds of preferences whose evolution we will consider: *fairminded* preferences, which exhibit inequity aversion, and *Homo Economicus* preferences, which reflect material self-interest alone. Section 4 provides a static analysis of the base game as played by agents with these two kinds of preferences. Section 5 introduces *fictitious play*, the model of learning used to estimate intragenerational material payoffs, and gives results for simulations of fictitious play learning for all preference-type pairings. Section 6 uses these simulation results to model evolution in populations of inequity averse and selfish individuals and shows that the central evolutionary result, mentioned above, follows from a well-known theorem in evolutionary game theory. Section 7 discusses some interpretational issues and concludes.

2. Fairness and the Indirect Evolutionary Approach

The indirect evolutionary approach is a framework for modeling the evolution of preferences in populations of rational choosers. An *indirect evolutionary game* is played by a large population of rational agents, each typed according to the preferences governing their choice behavior in a specified base game. In each generation, the members of the population are randomly paired to play the base game, with players choosing according to the preferences associated with their type. It is assumed that material resources are at stake in the base game, and that the players' preferences, represented by numerical utilities, need not be identical with these.[2] We assume further that these material payoffs are proportional to Darwinian fitness and that intergenerational selection operates on the preference-defined types, so that the rate at which a type grows in the population is a function of the average material payoff to individuals of that type in a given population state. An indirect evolutionary game thus involves a dual payoff structure: while the type-dependent subjective utilities determine individuals' choices in intra-generational play, it is the material consequences of those choices that dictate the rate at which types are reproduced intergenerationally.

Typically, solution concepts from classical game theory, especially the Nash equilibrium and its refinements, are used to predict which strategies individuals of each type will choose against each other type in the base game (see Königstein and Müller 2000, p. 239). Crucially, the chosen solution concept is applied to the game defined by the players' *subjective* utilities. With determinate predictions about how individuals will play under each possible pairing of types, the modeler can construct an *evolutionary game* in which the space of strategies is the set of types in the population and the payoff matrix is filled in with the material payoffs induced by the strategies predicted by the classical equilibrium analysis. Then, standard

[2]Since standard measures of subjective (preference-representing) utility are interval-scaled, "identical with" in this context is to be read as "linear in."

evolutionary game theoretic methods (e.g., identifying evolutionary stable strategies) can be applied to the derived evolutionary game to study how the population will evolve in terms of the proportional shares of the preference-defined types.

The literature applying the indirect evolutionary approach to issues of fairness dates back to Güth and Yaari (1992). They consider the ultimatum game, showing that preferences supporting rejecting small offers and proposing relatively large offers can outcompete materially self-interested preferences in populations comprising the two types. In their model, the evolutionary success of the fairness-concerned type depends crucially on the fact that players can observe their partner's type. Huck and Oechssler (1996) relax this assumption, showing that punishment of lowball offers in a simplified variant of the ultimatum game can persist when players know only the empirical distribution of types in the population. Subsequent work has extended these models in multiple directions. Berninghaus et al. (2007), for example, model the evolution of reciprocity and inequity concerns in populations playing both an ultimatum game and the closely-related dictator game, while Güth and Pull (2004) consider evolution in populations comprising "equity-driven" and "strategic" individuals playing the Nash demand game, on the assumption that players always choose strategy profiles corresponding to Nash's axiomatic bargaining solution (Nash 1950).

Kim and Lee (2023) offer a different model of the evolution of fairness in the Nash demand game. They consider populations playing a variant of that game in which the total stock to be divided is unknown to the players and alternating offers are allowed. Following earlier work, populations comprise two types, one self-interested and the other fairness-concerned. Kim and Lee explicate concern for fairness in terms of *inequity aversion.* Inequity aversion, introduced by Fehr and Schmidt (1999) as an explanation for experimental behavior in ultimatum games, refers to any pattern of preferences which reflects a distaste for material inequality. Kim and Lee find that inequity aversion evolves when players know their partner's type but that this result is not robust against introducing uncertainty about partner type.

As in Kim and Lee's project, the present paper considers the evolution of fairness in the context of a variant of the Nash demand game, where fairness is explicated in terms of inequity aversion. The main innovation concerns the method used to predict intragenerational play among types. Extant work predicts intragenerational play using classical solution concepts. The present model instead assumes players *learn* how to play with randomly-chosen partners with whom they repeatedly interact in blocks of bargaining problems. An individual's fitness is proportional to the average material payoffs they receive in these blocks. Both the relative speeds at which different pairs of types learn equilibria in the base game and the relative frequencies with which the different equilibria are learned are crucial to the central evolutionary result of the paper. Both are consequences of intragenerational learning. There is a methodological moral here: in the context of indirect evolutionary models, introducing learning can lead to population-level outcomes that could not have been predicted with static solution concepts alone.

Appealing to learning in explaining fair behavior in resource-division games is not without precedent. Roth and Erev (1995) show that fair behavior can emerge and stabilize in pairs of simple reinforcement learners playing repeated ultimatum games. Camerer and Ho (1999) suggest that their Experience-Weighted Attraction Learning, a form of reinforcement learning that is sensitive to counterfactual as well as actual payoffs, could similarly explain behavior observed in ultimatum game experiments. Mengel (2012) considers how subgame-imperfect play in the ultimatum game could arise in a model where players facing multiple similar bargaining games learn how to categorize those games as they play. The novelty of the present approach lies in its integration of learning and evolutionary dynamics in explaining how fairness might evolve. Rather than showing that learning produces fair behavior directly, it shows that learners with fairminded preferences tend to fare better materially than selfish ones in blocks of repeated bargaining games, and so those preferences are selected for evolutionarily.

It is worth noting that there is significant scholarly disagreement about the role of inequity aversion in explaining fair behavior (see Bergh (2008) and Binmore and Shaked (2010) for critical treatments). Reasons of space prevent a detailed discussion of the difficult empirical questions raised in that literature. Happily, the interest of the present project does not depend on settling those questions. If inequity aversion is indeed responsible for important kinds of fair behavior, then the question of its evolutionary origins is of obvious scientific interest. But even if inequity aversion is very uncommon or explanatorily insignificant in connection with fair behavior, investigating its evolutionary prospects in competition with selfishness may be of interest to social and political philosophers. There are longstanding debates in the history of political thought about the status of material self-interest as a motive for action. There are views on which some form of selfishness is an immutable feature of human nature that ought to constrain institutional choice. Others see the motivational role of self-interest as itself significantly shaped by social and economic institutions, holding that egalitarian or altruistic attitudes could (or did, or would) proliferate under the appropriate social or material conditions. Investigating the possibility of egalitarian preferences surviving natural dynamic processes like social learning or Darwinian evolution in competition with selfishness is a useful first step in approaching these debates. Moreover, the methodological moral mentioned above, concerning the significance of intragenerational learning vis-a-vis the indirect evolutionary approach, does not depend on the prevalence or behavioral significance of inequity aversion.

3. The Stage Game

The resource-division game we consider is a variant of the *Nash demand game*.[3] In our game, two players make competing demands on a fixed stock of some good. The set of possible demands is the same for each player, and we assume it contains just three elements: l, m, and h (for "low," "medium," and "high"). The value of m is always equal to half the quantity of the good available, and we assume that l and h together perfectly divide the resource. Each player chooses how much to demand without knowing what their partner will choose. We will use (d_i, d_j) to denote the pair of demands chosen by i and j.

Each player then receives a share determined by both their own demand and that of their partner. To facilitate developing an indirect evolutionary model, we must describe the game in terms of both the players' material payoffs and their subjective utilities. The material payoffs are described by a pair of *objective payoff functions* π_i and π_j defined on the space possible pairs of demands, where $\pi_i(d_i, d_j)$ is the number of units of the good i receives in the event that i demands d_i and j demands d_j. Each player gets the share they demanded if the sum of the players' demands does not exceed the total quantity of the good available; otherwise, the players have made *incompatible demands*, thereby realizing the *disagreement point* in which neither player gets anything.[4]

The *subjective utilities* are captured by a pair of functions u_i and u_j, also defined on the possible pairs of demands, where $u_i(d_i, d_j)$ gives the subjective utility i attaches to the outcome resulting from i demanding d_i and j demanding d_j. Unlike the material payoffs, which represent quantities of a physical good, the payoffs captured by u_i are just a numerical representation of an agent's preferences. The kinds of distributive preferences evolution acts on correspond will be captured by different assumptions about how a player's subjective utility function related to the underlying objective payoffs in the game.

We consider two types of players: *Homo Economicus* (HE) and *Fairminded* (FM). An HE player's subjective utilities depend only on the objective payoffs they receive for each pair of demands; that is, if player i is an HE type, then u_i is a function of π_i. This entails that an HE player's utilities are independent of the objective payoff received by their bargaining partner. This is a key feature of the standard homo economicus model of choice: "economic man" cares only about the material payoffs he receives in different outcomes.

For HE players, we assume the simplest functional relationship between objective and subjective payoffs, such that for HE player i, $u_i = \pi_i$. Note that since VNM utilities are interval scaled, any linear transformation of u_i is equivalent to u_i as a representation of player

[3]For more on the Nash demand game, see Nash 1950, Skyrms 1996.

[4]More general models of bargaining allow for non-zero payoffs at the disagreement point. Payoff asymmetries between the players at the disagreement point can also be modeled. See, e.g., O'Connor (2019), Bruner (2022).

		FM		
		l	m	h
	l	$(2.5, 2.5)$	$(2.5, 4.375)$	$(2.5, 6.25)$
HE	m	$(5, 1.875)$	$(5, 5)$	$(0, 0)$
	h	$(7.5, 1.25)$	$(0, 0)$	$(0, 0)$

TABLE 1. Subjective payoff matrix for the bargaining game defined by $(l, m, h) = (2.5, 5, 7.5)$, played by an HE and a fairminded player ($\alpha = 0.25$).

		FM		
		l	m	h
	l	$(2.5, 2.5)$	$(1.875, 4.375)$	$(1.25, 6.25)$
FM	m	$(4.375, 1.875)$	$(5, 5)$	$(0, 0)$
	h	$(6.25, 1.25)$	$(0, 0)$	$(0, 0)$

TABLE 2. Subjective payoff matrix for the bargaining game defined by $(l, m, h) = (2.5, 5, 7.5)$, played by two FM players ($\alpha = 0.25$).

i's preferences. So, we say an HE player's utility function is *linear in units of the good to be divided*.[5]

By contrast, if player i is Fairminded, their subjective utilities cannot be represented as a function of their own material payoffs alone. An FM player's utilities are a function of *both* their own material payoff *and* the material payoff to their partner. If player i is Fairminded, then for every pair of demands (d_i, d_j) (where j is i's bargaining partner),

$$u_i\,(d_i, d_j) = \pi_i(d_i, d_j) - \alpha|(\pi_i\,(d_i, d_j) - \pi_j\,(d_i, d_j)\,|$$

where $\alpha > 0$. The idea is that, other things equal, an FM player likes an outcome less the larger the inequality between what they receive and what their partner receives in that outcome. More precisely, this rule discounts the utility of an outcome by an amount proportional to the magnitude of the material inequality prevailing in that outcome, with no discounting applied to outcomes in which the players receive equal amounts of the good.

A *Nash equilibrium* of a game is an assignment of strategies, one for each player, such that no player could realize a strictly greater payoff by unilaterally deviating from their assigned strategy. In bargaining games, the strategies are demand levels, so in our game a

[5]I also studied a variant of the model with subjective utility functions that reflect diminishing marginal utility in objective payoff. In this treatment, subjective utilities for HE players are given by $u_i(d_i, d_j) = ln(\pi_i(d_i, d_j))$ and subjective utilities for FM players are given by $u_i\,(d_i, d_j) = ln(\pi_i(d_i, d_j)) - \alpha|ln(\pi_i\,(d_i, d_j)) - ln(\pi_j\,(d_i, d_j))|$. See Appendix 8.4 for details concerning the indirect evolutionary game played by a population of players with such utility functions.

		HE		
		l	m	h
HE	l	$(2.5, 2.5)$	$(2.5, 5)$	$(2.5, 7.5)$
	m	$(5, 2.5)$	$(5, 5)$	$(0, 0)$
	h	$(7.5, 2.5)$	$(0, 0)$	$(0, 0)$

Table 3. Subjective payoff matrix for the bargaining game defined by $(l, m, h) = (2.5, 5, 7.5)$, played by two HE players.

strategy profile is just a pair of demands. We are interested in the Nash equilibria defined by players' *subjective utilities*, rather than objective payoffs, since the subjective utilities are what guide their choices. When $\alpha = 0.25$ for the FM type, there are three Nash equilibria for all type pairings: (m, m), (h, l), and (l, h). When $\alpha \geq 0.5$, the equilibria for FM–HE and FM–FM pairings change. Appendix 8.2 provides details.

4. Static Analysis

Before modeling learning—and, later, evolution—for HE and FM agents, it will be helpful to consider the choice behavior of each type of player as a function of their beliefs about their partner's strategy. All results reported here are for the bargaining game with demand levels $(l, m, h) = (2.5, 5, 7.5)$. Games with demand levels given by $(l, h) \in \{(1, 9), (2, 8), (3, 7), (4, 6)\}$ (with $m = 5$) were also tested.[6]

We assume our players are expected utility maximizers. That is, a player's choice in the stage game is determined by their subjective utility function and a probability function describing their uncertainty concerning how their partner will choose, such that they always choose a strategy that maximizes the utility they expect to receive, given their beliefs. Figures 1 and 2 plot the optimal strategy relative to each possible set of beliefs about one's partner's demand for the two types we consider. The visualization works as follows. Consider the set containing all triples defining possible probability distributions $(p_i(l), p_i(m), p_i(h))$ of player i's degrees of belief over her partner j's available strategies in the bargaining game. Since probabilities must sum to one, we can infer the probability of the third strategy if we know the probabilities for the other two. This allows us to represent the space of possible probability distributions two-dimensionally by projecting each point (x_1, x_2, x_3) down to the point (x_1, x_2). The result is an equilateral triangle with vertices corresponding to belief states in which the player is *certain* that their partner will make a certain demand. Points in the interior of the triangle represent belief states in which they are uncertain about their partner's demand.

[6]See Appendix 8.4 for details.

For a given utility function u and demand level d, we can represent the set of belief states relative to which d maximizes expected utility with the set of all points $(x_1, x_2) \in \mathbb{R}^2$ such that d maximizes expected utility relative to the probability function given by: $p(m) = x_1$, $p(h) = x_2$, $p(l) = 1 - (x_1 + x_2)$. Figures 1 and 2 record, for each type, the ranges of credal states in which players of that type will choose each strategy (given the specified parameters). This information is represented by filling in different regions of the triangle representing possible belief states with different colors corresponding to the available demand levels. The yellow region includes all the belief states in which a rational chooser makes the low demand, the orange region covers all the belief states in which they make the medium demand, and the red region consists of the belief states in which they make the high demand. In Figure 1, for example, the point at $(\frac{1}{3}, \frac{1}{3})$ is orange, meaning that if the focal agent thought it equally like that their partner would choose any of the three available demands, they would make the medium demand.

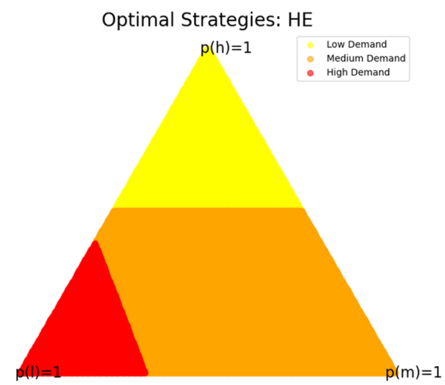

FIGURE 1. Expected-utility-maximizing choices for an HE agent in the stage game with demand levels $(l, m, h) = (2.5, 5, 7.5)$.

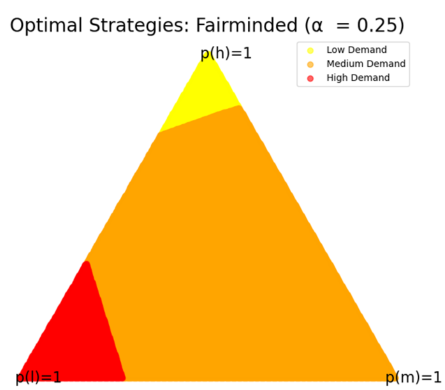

FIGURE 2. Expected-utility-maximizing choices for a Fairminded agent $(\alpha = 0.25)$ in the stage game with demand levels $(l, m, h) = (2.5, 5, 7.5)$.

Figure 1 assumes an HE-type utility function and Figure 2 assumes an FM type with $\alpha = 0.25$. Comparing these figures reveals that there is a smaller range of credal states relative to which an FM player will choose either l or h as compared with an HE player. This is intuitive: since an FM player prefers outcomes that minimize material inequality, they will choose the unique demand compatible with an equal split more often than will a player that cares only about their own material gain. As α increases, the share of the space of possible credal states for which m maximizes expected utility grows. When $\alpha \geq 0.5$, the low demand maximizes expected utility for an FM player only when they are *certain* their partner will choose high.

5. The Learning Model

To model evolution in a population of learners, we must specify a learning dynamics and estimate how much material payoff each type will get, on average, in each possible type pairing. These estimates will allow us to calculate mean objective payoffs for each type relative to any population state and so to model the evolutionary dynamics of a population of FM and HE agents, on the assumption that fitness is determined by objective payoffs. The main advantage in moving to a learning model in estimating within-generation objective payoffs is psychological realism. Classical solution concepts make strong rationality and knowledge assumptions about players, while simple models of inductive learning require much less of the agents that implement them. In particular, the model of learning we adopt here does not require that players have knowledge of their partner's rationality, their partner's type, or the type proportions in the current population state.

It is assumed that players learn by *fictitious play*, a simple and widely used model of inductive learning in games. Fictitious play is a *best-response* dynamics (see Fudenberg and Levine 1998). This means that a fictitious play learner is rational in the Bayesian sense: they always choose a strategy that maximizes expected utility given their beliefs about their partner's strategy. The dynamical content of fictitious play is given by a rule specifying how each player's credences evolve over time as a function of their experience. The rule can be modeled with balls and urns. Maintaining the assumption of a two-player game, a fictitious play learner's probabilities can be represented with an urn containing balls with labels $1, 2, ..., k$, where $q_i(t)$ denotes the number of i-labeled balls in the urn at time t. The agent's credence that their partner chooses strategy s_i at t is given by

$$p(s_i) = \frac{q_i(t)}{\sum_{j=1}^{k} q_j(t)}$$

The initial contents of the urn $q_1(0), q_2(0), ..., q_k(0)$ correspond to the learner's prior probabilities, which represent their pre-play assumptions about how their partner will choose. As the game is repeated, the contents of their urn are updated according to the following rule:

$$q_i(t+1) = \begin{cases} q_i(t) & \text{if partner did not choose } s_i \text{ at } t. \\ q_i(t)+1 & \text{if partner chose } s_i \text{ at } t. \end{cases}$$

In terms of balls and urns, the rule just says that the player adds an i-labeled ball to her urn whenever her partner chooses strategy s_i. At each timestep, she chooses a strategy that maximizes expected (subjective) utility relative to the beliefs given by the contents of her urn.

The main motivation for using fictitious play as a model of learning is that it is equivalent to Bayesian inference under two conditions: (1) the learner assumes their partner's choices are well-modeled by a sequence of independent and identically distributed random variables with a fixed but unknown distribution, and (2) the learner's prior over candidate partner-play-generating distributions is Dirichlet distributed.[7] It is worth noting that at every step in the repeated game, each player chooses only the demand to they will make at that step. In particular, they cannot adopt more sophisticated diachronic strategies that require planning.

Another property of fictitious play is significant for our model. Every strict Nash equilibrium is an *absorbing state* of a fictitious play process. This means that, if a strategy profile is a Nash equilibrium, then once that strategy profile is realized on some timestep, it will be realized on every subsequent timestep (see Proposition 2.1 in chapter 2 of Fudenberg and Levine (1998)).

We use fictitious play learning to estimate lifetime objective payoffs for HE and FM individuals in an evolving population. We assume that each agent's lifetime comprises a fifty-period repeated game with a partner chosen at random from the population. The agents learn by fictitious play. And at the end of their lifetime, they reproduce at a rate determined by the total quantity of objective payoff they accrued with their partner, with offspring inheriting their distributive preferences.

To estimate lifetime material payoffs, we ran computer simulations. In considering their results, we are mainly interested in comparing the cumulative objective payoffs that players of each type typically receive in a given run. Of course, the absolute values are not meaningful here, since the objective payoffs have not been given a concrete interpretation. But their relative values will be crucial for getting results for the evolutionary model.

Each simulated run consisted of a series of 50 plays of a repeated bargaining game in which players learned by fictitious play. As above, results are reported for the game with demand levels $(l, m, h) = (2.5, 5, 7.5)$. For each run, the players' prior probabilities are chosen randomly and independently. Specifically, for each player, three random numbers in the unit interval are chosen to serve as initial weights $q_l(0), q_m(0), q_h(0)$, where the prior probability

[7] See Chapter 2 of Fudenberg and Levine 1998 for details.

p is defined by

$$p(s) = \frac{q_s(t)}{\sum_{s' \in S} q_{s'}(t)},$$

for all strategies $s \in S = \{l, m, h\}$. The evolution of the player's beliefs is described by the fictitious play dynamics, so that

$$q_i(t+1) = \begin{cases} q_i(t) & \text{if partner did not choose } s_i \text{ at } t. \\ q_i(t) + k & \text{if partner chose } s_i \text{ at } t. \end{cases}$$

with $k = 0.02$.[8] Holding initial propensities fixed, larger k corresponds to faster learning and smaller k corresponds to slower learning. The discounting factors α for FM players were varied between the values $\{0.25, 0.5, 0.75, 1\}$.[9]

For each assignment of parameters, we collected results for 10^5 runs. As long as $\alpha = 0.25$, the Nash equilibria (relative to the players' subjective utilities) of the base game for all pairings of player types (HE–HE, HE–FM, FM–FM) are the same. In particular, (m, m), (h, l), and (l, h) are the Nash equilibria. For pairings involving an FM player with $\alpha > 0.25$, the equilibria change. See Appendix 8.2 for details. Since every Nash equilibrium is an absorbing state under fictitious play, once the players play a pair of strategies that constitute a Nash equilibrium, they will always choose that strategy profile in subsequent play.

We recorded the strategy profile chosen on the final play of each run to determine whether an equilibrium had been reached on that run and, if it had, whether that equilibrium was more favorable to one or the other of the players.[10] We also recorded the cumulative objective payoffs for each player on each run. For a given player and run, the cumulative objective payoff is calculated by taking the sum of the objective payoffs the player received in each play of that run. This value reflects more than just the final state a pair of players reached in a run; it also depends on the payoffs realized in early play prior to the players reaching an equilibrium.

When an HE agent plays another HE agent, the (m, m) equilibrium is realized on about 82% of runs. The rest of the time—aside from a handful of exceptions in which the players fail to learn any equilibrium–an (l, h) or (h, l) equilibrium is reached, with each occurring in roughly 9% of runs. On average, an equilibrium is learned in about 6.1 timesteps. The mean cumulative payoff for an HE player paired with another HE player is approximately 228.

[8]If we want to preserve the aptness of the balls-and-urns interpretation, we can set $k = 1$ and multiply the initial weights by 50.

[9]See table 7 in the Appendix for estimated lifetime material payoffs for the parameter settings not discussed in this section.

[10]If the strategy profile on the final play was not a Nash equilibrium, then the players had never chosen an equilibrium strategy profile on that run; otherwise, they would have played that equilibrium profile on every play after the first time it was chosen.

The dynamics are very different when the game is played by a HE player and an FM one. Results for these pairings are recorded in Table 3. The *equilibrium realized* column reports the proportion of runs on which the players converged to each Nash equilibrium, while the *mean cumulative objective payoffs* column records mean total of objective payoffs received across the 50–play run for agents of each type, rounded to the nearest whole number. The *mean time-to-convergence* column records how many timesteps, on average, it took players in each batch of runs to learn an equilibrium. Notice that, in these runs, the egalitarian convention is again the most common outcome, occurring in about 90% of runs for $\alpha = 0.25$. Larger α corresponds to a larger share of runs ending in an (m, m) equilibrium. When $\alpha \geq 0.5$, no run ended in an (h, l) strategy profile favoring the HE player. This is unsurprising, since choosing l never maximizes expected utility for an FM player when $\alpha \geq 0.5$.

Table 4. Homo Economicus/Fairminded Pairings

α	*equilibrium realized*: (d_{HE}, d_{FM})				*mean cumulative objective payoffs*		*mean conv. time*
	(m,m)	(h,l)	(l,h)	other	HE	FM	
0.25	0.898	0.021	0.072	0.009	223	236	4.87
0.5	0.943	0	0.048	0.009	226	239	4.39
0.75	0.971	0	0.024	0.005	230	240	4.18
1	1	0	0	0	235	240	4.12

Results for simulated fictitious play learning in 50–play blocks of repeated bargaining games (HE/FM pairings).

Pairings comprising two FM players exhibited the strongest tendency toward the egalitarian equilibrium. For $\alpha = 0.25$, about 95% of runs ended with equal splits. That value rises as α increases, with all runs ending in the egalitarian equilibrium when $\alpha = 1$. This is because making the medium demand maximizes expected utility unless the player is absolutely certain that their partner will demand either low or high, for an FM player with $\alpha = 1$.

Notice that, for any fixed α, FM players receive a higher mean cumulative payoff when paired with other FM players than do HE players paired with FM players. And in FM–HE pairings, the FM players receive a higher mean cumulative payoff than HE players as well.

FM agents outperform HE agents in HE–FM matchups because the players are more likely to learn the equilibrium that favors the FM player than the one that favors the HE player, in the unlikely event that they do not learn the egalitarian equilibrium. The reason for this asymmetry has to do with differences in how the strategic behavior of players of the two types depends on their beliefs. In order for the players to land in an inequitable equilibrium, there has to be some timestep on which the high demand maximizes expected

Table 5. Fairminded/Fairminded Pairings

α	proportion of runs ending in (m,m) equilibrium	mean cumulative obj. payoff	mean time to convergence
0.25	0.949	234	2.76
0.5	0.984	241	1.04
0.75	0.996	247	0.31
1	1	250	0

Results for simulated fictitious play learning in 50–play blocks of repeated bargaining games (FM/FM pairings).

utility for one player and the low demand maximizes expected utility for the other. Figures 1 and 2 show that the high demand is a best response relative to a slightly larger set of belief states for an HE player than for an FM one. Similarly, there is a larger range of credal states for which the low demand is a best response for an HE player as compared with an FM one, only here the difference is much larger—relative to an HE player, an FM player must place *very* high probability on their partner making the high demand in order for demanding low to be a best response. As a result, the HE–FM pairs more often realize a state in which the FM player demands high and the HE player demands low than vice versa.

The reason for the difference in mean objective payoff between FM–FM and HE–HE pairings is different. Here, the explanation has to do with *how quickly* each kind of pairing reaches an equilibrium under the learning dynamics. In FM–FM pairings (with $\alpha = 0.25$), the players reach an equilibrium strategy profile just over three timesteps earlier than players in HE–HE matchups do, on average. The reason this difference matters is that when a pair plays a non-Nash strategy profile at some time step, some amount of the resource being split goes to waste. At the disagreement point, the players jointly receive zero units of objective payoff; in the (l, l) outcome, $2m - 2l$ of the resource goes to no one. So, slower convergence to an equilibrium in pairings of agents of the same type corresponds to smaller mean payoff for players of that type.

The quicker convergence in FM–FM matchups is explained by the fact that the medium demand is the optimal choice relative to *most* possible credal states for an FM player, on all values for α tested. Only relatively extreme credal states (i.e., those which place a large probability mass on their partner making a low demand) support a low or high demand as a best response. As a consequence, when two FM players meet, they are likely quickly coordinate on the egalitarian equilibrium. By contrast, the best response for a HE player is more sensitive to the player's beliefs (as reflected in the more equal distribution of colors in Figure 1), resulting in more failures to coordinate on a Nash strategy profile in early play.

6. Evolutionary Analysis

The next step is to use the simulation results from the learning model to construct a payoff matrix in which the strategies labeling rows and columns are the HE and FM types and the payoffs in each cell are the expected per-block payoffs for the corresponding pair of types. Then, following the usual practice on the indirect evolutionary approach, we apply standard evolutionary game theoretic tools to study the evolution of the two types of preferences. We begin by qualitatively describing the evolutionary setting.

We assume consider an effectively infinite population comprising both HE and FM individuals. Each individual is initially assigned a random partner from the population, and their lifetime comprises a fifty-period repeated bargaining game played with that partner. When two players meet, their prior probabilities over each other's strategies are chosen at random, and each player learns by fictitious play. The proportions of HE and FM players evolve under a selection dynamics, with fitness determined by material payoffs. In particular, the fitness associated with each type is equal to the expected material payoff to an individual with that type's preferences over a 50-play block against a randomly-selected member of the current population. A type's fitness thus depends on two factors: first, how much material payoff an individual of that type can expect to get in a block of play against an individual of each possible type, and, second, the proportion of the population comprising each type.

Since pairing is random and the population is very large, we assume that the average per-block objective payoff for, say, HE-type individuals in interactions with, say, FM-type individuals is equal to the mean per-block cumulative objective payoffs reported for HE–FM pairings in the simulations of the learning model discussed above. Of course, this will hold for all type pairings (HE–HE, FM–HE, FM–FM). Thus, we will use the simulation results for the learning model as estimates of the average cumulative per-block objective payoffs for each pairing of types in the evolutionary model. Let $E[\pi(i,j)]$ denote the expected objective payoff for an i-type agent in a 50-play iterated game played with a j-type agent. Estimates of this value for each type pairing, with $\alpha = 0.25$, are as follows:

$$E[\pi(HE,HE)] = 228$$
$$E[\pi(HE,FM)] = 223$$
$$E[\pi(FM,HE)] = 236$$
$$E[\pi(FM,FM)] = 234$$

Given random pairing, the mean per-block objective payoff to each type in a given population state is a weighted average of the mean payoffs members of that type receive against their own type and against the other type, with weights given by the corresponding type proportions.

The next step is to consider the population dynamics for HE and FM types in the evolutionary setting described above by applying standard evolutionary game theoretic tools to the game defined in Table 5. Notice that (FM, FM) is the only Nash equilibrium in pure strategies, and that it is strict. This is the case for all values of α tested in both discounting treatments (with and without diminishing marginal utility in objective payoffs; see footnote 5 and Appendix 8.4).

	HE	FM
HE	$(228, 228)$	$(223, 236)$
FM	$(236, 223)$	$(234, 234)$

Table 6. Payoff matrix (in expected lifetime objective payoffs) for indirect evolutionary game with $\alpha = 0.25$.

Since FM–FM is the unique strict Nash equilibrium, the evolutionary analysis is straightforward. In particular, since FM strictly dominates HE[11], it follows from a classic result that the fairminded type will eventually take over the population entirely, as long it has non-zero proportional representation in the initial population, under a broad class of selection dynamics (see Weibull 1997, p. 147).[12] The relevant class is the set of *payoff monotone* dynamics, where a selection dynamics is payoff monotone if it is such that the growth rate of every strategy is increasing in the average payoff to players who use that strategy.

To sum up, in populations comprising both selfish and inequity averse players, the inequity averse type will eventually take over the population under any payoff monotone selection dynamics as long as it has non-zero representation in the initial state.[13] The dominance result underlying this conclusion is robust against varying the intensity of the FM type's preference for equal outcomes and varying whether the agent's preferences reflect diminishing marginal utility in objective payoffs (see footnote 7).

[11]As mentioned above, we also ran simulations for games in which $h \in \{0.6, 0.7, 0.8, 0.9\}$ and $l = 1 - h$ (for all parameter settings considered, m was set to 0.5). Estimated lifetime material payoffs are reported in Appendix 8.4, with the same simulation protocol as described in Section 5. For all treatments except $l, h = 4, 6$, the strict dominance of FM over HE holds. For $l, h = 4, 6$ (with payoff estimates rounded to the nearest whole number), FM only weakly dominates HE.

[12]The relevant result: for any game with finitely many strategies, the population share of any pure strategy which is strictly dominated by another pure strategy will go to zero in the limit under any payoff monotonic selection dynamics, as long as every strategy has a non-zero population share in the initial state. For details on the relationships among dominance concepts and selection dynamics, see Nachbar (1990), and Dekel and Scotchmer (1992), and Hofbauer and Weibull (1996).

[13]Results for the finite-population simulations described in the Appendix are qualitatively similar: the FM type took over the population in all but a handful of runs in which it represented a tiny proportion of the initial population.

7. Discussion

We have considered the evolution of inequity aversion in the context of a discretized variant of the Nash demand game. The model makes use of the indirect evolutionary approach, incorporating a dual payoff structure where individual choice is directed by subjective payoffs and Darwinian fitness is proportional to material payoffs. But unlike extant work adopting the indirect evolutionary approach, the average material payoffs for each type in a given generation are predicted using a model of learning, rather than a classical solution concept. On a broad range of parameter settings, inequity averse players receive higher material payoffs, on average, than selfish players, both when paired against their own type and when paired against the selfish type. So, when the two types compete in a population over evolutionary time, any payoff monotone selection dynamics will lead to inequity aversion taking over the population, as long as it has non-zero proportional representation in the initial state. The evolutionary dominance of inequity aversion is further reinforced by simulations of evolution in finite populations, reported in the Appendix. What's more, the asymmetries in mean cumulative material payoff between fairminded and selfish players that drive the evolutionary result depend crucially on the fact that players *learn* their strategies.

Does these results suggest that should we *always* expect to observe inequity aversion in human populations? If so, that might be a serious problem for the model, given the substantial evidence of cultural variation in attitudes about fairness (Henrich et al. 2001). Fortunately, there are at least two strong reasons to resist this interpretation. Although the model is very general with respect to the class of selection dynamics considered, it makes rather special assumptions about the types of distributive preferences represented in the population and the range of interactions that determine an individual's fitness in a given generation. When these assumptions are relaxed, we have good reason to think other evolutionary outcomes are possible.

The first reason is that our central evolutionary result only applies to populations comprising two types of agents: Fairminded and Homo Economicus. Results for these populations tell us little about how evolution would go in a more general setting including other types of distributive preferences. Second, in the setting we consider, fitness depends only on material payoffs realized in a version of the Nash demand game. But humans engage in many different kinds of strategic interactions. The fitness contribution of a given profile of distributive preferences will be sensitive to the consequences of having those preferences in social interactions that are not well-modeled by the Nash demand game. While inequity aversion is adaptive in the version of that game we consider here, it may well be maladaptive in other resource-division games, and in a population that plays such games alongside (or instead of) the game considered here, different kinds of distributive preferences may be favored evolutionarily.

While our results do not support any particular prediction about the prevalence or extent of inequity aversion in real human populations, they have two significant upshots. First, our results establish the possibility of the proliferation and persistence of egalitarian preferences in competition with selfish ones in a particular kind of bargaining game. Second, they highlight the role of learning as a potential mechanism by which such a taste for fairness might evolve.

8. Appendix

8.1. **Evolution in Finite Populations.** We ran simulations of evolution in finite populations of Fairminded and Homo Economicus agents under a variant of the replicator dynamics. The model works as follows. At the beginning of each generation, each member i of an n-agent population is randomly paired with a different member j. Players i and j then play 50 iterations of the stage game (the game is the same for all pairs of players and all plays in the generation). Each generation thus consists of 50 timesteps, at each of which every assigned pair plays a single round of the stage game. The players learn by fictitious play, with the same parameters as in the learning simulations described in Section 4. Priors are assigned by the same procedure as in the earlier model as well.

The rule describing how proportions of types in the population change between generations is a finite-population variant of the replicator dynamics. Fitness scores for each type are equal to the mean cumulative objective payoffs realized by individuals of that type in a given generation. Let the players in the population at generation g be enumerated as $\mathrm{P}(g) = \{1, 2, ..., n\}$, and let $T_i(g)$ denote the subset of $\mathrm{P}(g)$ containing all and only individuals of type $i \in \{\mathrm{HE, FM}\}$. Use $\Pi_i(g)$ to denote the *total payout* to type i in generation g, where Π_i is calculated by summing over the cumulative payoffs for each individual of type i in generation g. Formally:

$$\Pi_i\,(g) = \sum_{j \in P_i} \left(\sum_{t=0}^{49} (\pi_j(t)) \right),$$

where t ranges over the timesteps in g and $\pi_j(t)$ denotes the objective payoff received by player j at timestep t. Dividing $\Pi_i(g)$ by the total number of i-type individuals in the population yields the fitness $f_i(g)$ of type i, given by the *mean cumulative payoff* to i-type players in generation g:

$$f_i(g) = \frac{\Pi_i(g)}{|T_i(g)|}.$$

Let $pr_i(g) =$ denote the proportion of the population comprised of i-type individuals in generation g. Let $\boldsymbol{\mu}(g)$ denote the mean cumulative payoff for the entire population:

$$\boldsymbol{\mu}(g) = \sum_i \left(pr_i\,(g)\, f_i\,(g)\right),$$

where i ranges over types in the population. The time-evolution of the type proportions is governed by a discretized version of the standard replicator dynamics. Let $near : \mathbb{R}^+ \rightarrow \mathbb{N}$ be the function that takes any non-negative real number r and returns the natural number n nearest to r (i.e. the n that minimizes $|n - r|$). The dynamics is given by

$$pr_i(g+1) = near\left(pr_i(g)\frac{f_i(g)}{\boldsymbol{\mu}(g)}\right).$$

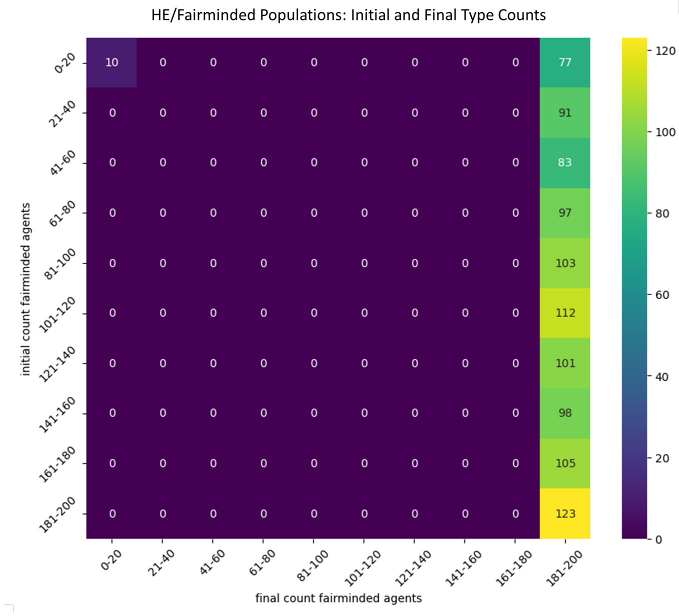

Figure 3. Initial and final type distributions for populations comprised of HE and Fairminded agents.

I simulated 1000 generations of evolution (each generation consisting of 50 timesteps) for 1000 randomly generated 200-member populations comprising both HE and Fairminded agents. Initial type proportions for each population were determined by generating a random number $0 < k < 200$ and setting the number of Fairminded individuals in the population to k and the number of HE individuals to $200 - k$.

These runs typically ended with the Fairminded type having taken over the population entirely; this occurred in 990 runs. In the remaining 10 runs, the HE type took over the entire population. In no run did a mixture of types persist to the end of the simulation. Figure 3 illustrates this by plotting initial and final counts of Fairminded agents for each run. Each square at coordinates (x, y) records the number of runs in which the initial type distribution lay in range y, among those in which the final distribution lay in range x. Notice that every run that ended with an HE-only population also *began* with a very large proportion of HE agents in the population. And even among those runs with 20 or fewer Fairminded agents in the initial population, the majority (77 out of 87) ended in a complete takeover by the Fairminded type.

8.2. **Stage Game Payoff Matrices.** Below are payoff matrices for the stage game with $l, m, h = 2.5, 5, 7.5$, in various type pairings with α varied among $\{0.5, 0.75, 1\}$ for the FM type. Payoffs are given in subjective utilities. Nash equilibria are bolded.

$\alpha = 0$

$\alpha = \frac{1}{2}$		l	m	h
	l	$(2.5, 2.5)$	$(1.25, 5)$	$\mathbf{(0,7.5)}$
	m	$(3.75, 2.5)$	$\mathbf{(5,5)}$	$(0, 0)$
	h	$\mathbf{(5,2.5)}$	$(0, 0)$	$(0, 0)$

$\alpha = \frac{3}{4}$

$\alpha = \frac{3}{4}$		l	m	h
	l	$(2.5, 2.5)$	$(0.625, 3.125)$	$(-1.25, 3.75)$
	m	$(3.125, 0.625)$	$\mathbf{(5,5)}$	$(0, 0)$
	h	$(3.75, -1.25)$	$(0, 0)$	$\mathbf{(0,0)}$

$\alpha = \frac{1}{2}$

$\alpha = \frac{1}{2}$		l	m	h
	l	$(2.5, 2.5)$	$(1.25, 3.75)$	$\mathbf{(0,5)}$
	m	$(3.75, 1.25)$	$\mathbf{(5,5)}$	$(0, 0)$
	h	$\mathbf{(5,0)}$	$(0, 0)$	$\mathbf{(0,0)}$

$\alpha = 0$

$\alpha = 1$		l	m	h
	l	$(2.5, 2.5)$	$(0, 5)$	$(-2.5, 7.5)$
	m	$(2.5, 2.5)$	$\mathbf{(5,5)}$	$(0, 0)$
	h	$\mathbf{(2.5,2.5)}$	$(0, 0)$	$(0, 0)$

$\alpha = 0$

$\alpha = \frac{3}{4}$		l	m	h
	l	$(2.5, 2.5)$	$(0.625, 5)$	$(-1.25, 7.5)$
	m	$(3.125, 2.5)$	$\mathbf{(5,5)}$	$(0, 0)$
	h	$\mathbf{(3.75,2.5)}$	$(0, 0)$	$(0, 0)$

$\alpha = 1$

$\alpha = 1$		l	m	h
	l	$\mathbf{(2.5,2.5)}$	$(0, 2.5)$	$(-2.5, 2.5)$
	m	$(2.5, 0)$	$\mathbf{(5,5)}$	$(0, 0)$
	h	$(2.5, -2.5)$	$(0, 0)$	$\mathbf{(0,0)}$

8.3. **Learning Results for Other α Values.** Below are estimated lifetime objective payoffs for pairs of players with various degrees of inequity aversion in 50-period repeated Nash demand games with $l, m, h = 2.5, 5, 7.5$. Players learned by ficitious play with all details of the simulation protocol as described in section 5.

a	0	0.25	0.5	0.75	1
0	228,228	223,236	225,239	231,239	235,240
0.25	236,223	234,234	236,239	239,241	243,244
0.5	239,225	239,236	241,241	244,244	247,247
0.75	239,231	241,239	244,244	247,247	249,249
1	240,235	244,243	247,247	249,249	250,250

8.4. **Indirect Evolutionary Games: Variants.** Below are payoff matrices for indirect evolutionary games with demand levels given by $m = 5$, $l \in \{1, 2, 3, 4\}$, and $h = 10 - l$. Results are also reported for populations of players with log-utility functions as described in footnote 5, with $l, m, h = 2.5, 5, 7.5$. Lifetime objective payoffs for the various type-pairings are calculated according to the same protocol as above. For all treatments, $\alpha = 0.25$ for the FM type.

	HE	FM
HE	$(179, 179)$	$(192, 194)$
FM	$(194, 192)$	$(205, 205)$

$l, m, h = 1, 5, 9$

	HE	FM
HE	$(234, 234)$	$(230, 243)$
FM	$(243, 230)$	$(240, 240)$

$l, m, h = 3, 5, 7$

	HE	FM
HE	$(212, 212)$	$(212, 222)$
FM	$(222, 212)$	$(224, 224)$

$l, m, h = 2, 5, 8$

	HE	FM
HE	$(222, 222)$	$(222, 222)$
FM	$(222, 222)$	$(223, 223)$

$l, m, h = 4, 5, 6$

	HE	FM
HE	$(237, 237)$	$(234, 245)$
FM	$(245, 234)$	$(244, 244)$

$l, m, h = 2.5, 5, 7.5$, log utilities

References

[1] Bergh, A. (2008). A critical note on the theory of inequity aversion. *The Journal of socio-economics*, 37(5), 1789-1796.

[2] Binmore, K., & Shaked, A. (2010). Experimental economics: Where next?. *Journal of economic behavior & organization*, 73(1), 87-100.

[3] Bolton, G. E., & Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *American economic review*, 91(1), 166-193.

[4] Bruner, J. P. (2021). Nash, bargaining and evolution. *Philosophy of Science*, 88(5), 1185-1198.

[5] Camerer, C., & Hua Ho, T. (1999). Experience-weighted attraction learning in normal form games. *Econometrica*, 67(4), 827-874.

[6] Debove, S., Baumard, N., & André, J. B. (2016). Models of the evolution of fairness in the ultimatum game: a review and classification. *Evolution and Human Behavior*, 37(3), 245-254.

[7] Dekel, E., & Scotchmer, S. (1992). On the evolution of optimizing behavior. *Journal of Economic Theory*, 57(2), 392-406.

[8] Engel, C. (2011). Dictator games: A meta study. *Experimental economics*, 14(4), 583-610.

[9] Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. The quarterly journal of economics, 114(3), 817-868.

[10] Fehr, E., & Schmidt, K. M. (2006). The economics of fairness, reciprocity and altruism—experimental evidence and new theories. in *Handbook of the economics of giving, altruism and reciprocity*, 1, 615-691.

[11] Fudenberg, D. and Levine, D. (1998) *The theory of learning in games.* MIT press

[12] Gale, J., Binmore, K. G., & Samuelson, L. (1995). Learning to be imperfect: The ultimatum game. *Games and economic behavior*, 8(1), 56-90.

[13] Güth, W., Schmittberger, R., & Schwarze, B.. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization*, 3(4), 367–388.

[14] Güth, W., & Kocher, M. G. (2014). More than thirty years of ultimatum bargaining experiments: Motives, variations, and a survey of the recent literature. *Journal of Economic Behavior & Organization*, 108, 396-409.

[15] Güth, W., & Yaari, M. (1992). Explaining reciprocal behavior in simple strategic games: An evolutionary approach. *Explaining process and change: Approaches to evolutionary economics*, 23-34.

[16] Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & McElreath, R. (2001). In search of homo economicus: behavioral experiments in 15 small-scale societies. American economic review, 91(2), 73-78.

[17] Hofbauer, J., & Weibull, J. W. (1996). Evolutionary selection against dominated strategies. Journal of economic theory, 71(2), 558-573.

[18] Huck, S., & Oechssler, J. (1999). The indirect evolutionary approach to explaining fair allocations. *Games and economic behavior*, 28(1), 13-24.

[19] Kim, J. Y., & Lee, K. M. (2023). Evolution of fairness in the divide-a-lottery game. *Scientific Reports*, 13(1), 7048.

[20] Königstein, M., & Müller, W. (2000). Combining rational choice and evolutionary dynamics: The indirect evolutionary approach. *Metroeconomica*, 51(3), 235-256.

[21] Mengel, F. (2012). Learning across games. *Games and Economic Behavior*, 74(2), 601-619.

[22] Nachbar, J. H. (1990). "Evolutionary" selection dynamics in games: Convergence and limit properties. *International journal of game theory*, 19, 59-89.

[23] Nash, J. F. (1950). The bargaining problem. *Econometrica*, 18(2), 155-162.

[24] Nowak, M. A., Page, K. M., & Sigmund, K. (2000). Fairness versus reason in the ultimatum game. *Science*, 289(5485), 1773-1775.

[25] Nydegger, R. V., & Owen, G. (1974). Two-person bargaining: An experimental test of the Nash axioms. *International Journal of game theory*, 3, 239-249.

[26] O'Connor, C. (2019). *The origins of unfairness: Social categories and cultural evolution*. Oxford University Press, USA.

[27] Roth, A. E., & Erev, I. (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and economic behavior*, 8(1), 164-212.

[28] Skyrms, B. (1996) *Evolution of the social contract.* Cambridge University Press.

[29] Weibull, J. W. (1997). *Evolutionary game theory.* MIT press.