

Rationality in context

On inequality and the epistemic problems of maximizing expected utility

Dominik Klein^{1,2}  · Johannes Marx¹ ·
Simon Scheller³ 

Received: 19 September 2017 / Accepted: 23 March 2018 / Published online: 31 March 2018
© Springer Science+Business Media B.V., part of Springer Nature 2018

Abstract The emergence of economic inequality has often been linked to individual differences in mental or physical capacities. By means of an agent-based simulation this paper shows that neither of these is a necessary condition. Rather, inequality can arise from iterated interactions of fully rational agents. This bears consequences for our understanding of both inequality and rationality. In a setting of iterated bargaining games, we claim that expected utility maximizing agents perform suboptimally in comparison with other strategies. The reason for this lies in complex feedback effects between an agents' action and the quality of beliefs used to calculate expected utility. Consequentially, we argue that the standard notion of rationality as maximizing expected utility is insufficient, even for certain standard cases of economic interaction.

Keywords Inequality · Rationality · Bargaining · Agent-based modelling · Rational choice

✉ Dominik Klein
dominik.klein@uni-bamberg.de

Johannes Marx
johannes.marx@uni-bamberg.de

Simon Scheller
simon.scheller@lrz.uni-muenchen.de

¹ University of Bamberg, Bamberg, Germany

² University of Bayreuth, Bayreuth, Germany

³ Munich Center for Mathematical Philosophy, Munich, Germany

1 Introduction

Inequality has recently become a major issue again in philosophy, economics, and social sciences (Atkinson and Bourguignon 2014; Salverda et al. 2009). It is discussed as a source of conflict between countries of the global south and north (Wood 1995), but also as a problematic issue within developed western democracies (Gottschalk and Smeeding 2000). An extensive literature identifies a broad range of social, cultural, educational and long-term-economic *effects* of inequality and poverty (McLeod et al. 2014).¹

In the present paper, however, we are concerned with the *causes* of inequality rather than its consequences. More concretely, we are interested in the relationship between the rationality of the agents involved and the emergence of inequality. By studying how the interaction of individually rational agents relates to inequality we aim to acquire new insights into both, inequality and the nature of rationality. In particular, we will show that the narrow conception of rationality as maximizing expected utility falls short in crucial aspects.

Literature on the connection between inequality and rationality is scarce, but that which exists provides some crucial insights (see Sheehy-Skeffington and Rea (2017) for an overview). A common result of the recent empirical literature on poverty and rationality is that “the poor often behave differently from the non-poor” (Carvalho et al. 2016). Numerous articles trace back unequal outcomes to supposedly irrational individual behavior, arguing that the poor frequently fail to meet the high requirements of rationality. We aspire to show that this conclusion is short-sighted. The alternative narrative we provide illustrates how unequal outcomes can emerge from different strategies that are all contextually rational.

Similar results appear in the literature on risk-taking. Empirical research suggests that poorer people and those with lower incomes exhibit a higher degree of risk-aversion (Riley and Chow 1992; Grable 2000), but that their exposure to existential financial risks in everyday lives nonetheless remains higher compared with their wealthier counterparts (Rehm 2011). Others report more mixed results with regards to the relation between income and risk-preferences (Halek and Eisenhauer 2001; Blume and Easley 2008), also from a theoretical perspective (Banerjee and Newman 1991; Kanbur 1979; Kolodny and Stern 2017). Our first conclusion is that these differences in risk-taking behavior may reflect rational responses to different boundary conditions. Having to balance considerations of short term survival with long term economic prospects may play out very differently for different types of agents. As our simulation illustrates, agents may have justifiable non-neutral risk attitudes in the short run, reflecting, for instance, the urgency of short-term needs. In short, the standards of rational choice may be context sensitive and the notion of rationality as maximizing expected utility could be insufficient for evaluating strategic behavior.

Classically, rationality is understood here in terms of Rational Choice Theory. That is, rational behavior is defined as maximizing expected utility, given a fixed set

¹ While poverty is a property that can be ascribed to one person, inequality is relational in nature. By inequality we refer to a concept that captures one’s relative position compared to a group of relevant others (Haughton and Khandker 2009).

of internally and externally consistent desires and beliefs (Scott 1999; Briggs 2017; Sen 2008). As probably the most prominent notion of rationality, it has extended into virtually all domains of the social sciences and philosophy alike (Becker 2013; Riker 1995). However, various authors have taken issue with this notion of rationality, claiming that it exhibits significant empirical and methodological shortfalls (Green and Shapiro 1994).

Much criticism against rational choice hinges on the fact that certain rational choice models either fail to take into account relevant aspects of a situation (Grüne-Yanoff 2012) or are at odds with observed behavior of well-performing agents (Kahneman and Tversky 1979). Our simulation indicates that Rational Choice Theory's notion of rationality can fall short in an even more fundamental way. We show that even in a seemingly innocuous setting, maximizing the long term incomes of economic agents, rational choice rules may perform suboptimally. We identify contexts where strategy choice is structurally connected to the long term quality of agents' beliefs. Through this connection, strategies may undermine the basis of their own success and hence perform poorly in the long run. Consequentially, agents need not only reason about their short and long term payoffs, but also about the quality of their epistemic states. One manifestation of this is the 'exploration-exploitation trade-off': Short-term utility maximization must be balanced with the discovery of even better long-term strategies (Holland 1992; Goldberg 1989; Zollman 2010). In short, in the context of iterated interactions a substantive theory of rationality should take into account the substantial interactions between actions and the quality of future beliefs.

At the same time, we contribute to the literature on the origins of inequality. Mostly, empirical research on inequality employs a macro level perspective. Others take a decision-theoretic perspective and focus on the quality of individuals' decisions (Neckerman and Torche 2007; Dabla-Norris et al. 2015). We hold that both miss out on an important source of inequality. In an economy where agents interact with each other, revenue maximization must be situated in a game theoretic, rather than a decision theoretic context. The present paper complements existing literature with a simulative approach, studying the emergence of inequality from a *game theoretic* perspective.

We start from the assumption that bargaining constitutes a ubiquitous (if not defining) feature of economic interaction – especially when distributional issues are settled. Additionally, modern economy is best described as a highly interconnected system in which agents frequently interact with each other. As a result, inequality may thus emerge through the interaction of agents in a complex system, and cannot be reduced to the individual psychological traits of agents alone. This system of strategic dependencies should be expected to have a substantial effect on the distribution of incomes. We ask, therefore, to what extent strategic behavior in bargaining processes can be a source of lasting inequality in societies. Especially in interactive settings where the success of various strategies depends on the bargaining behavior of others, considerations of short and long term rationality gain extra weight. One's opponent's behavior, in turn, may vary over time, complicating the problem of learning about one's own optimal strategy even further. Moreover, an agent's first hand choices may impact the beliefs of others and hence their future actions. Here again, the choice of short term optimal actions may impact the long term prospects agents face.

In sum, we argue that the orthodox *thin* notion of rationality as maximizing expected utility (Ferejohn 1991; Green and Shapiro 1994; Yee 1997) is lacking in at least two dimensions. First, it misses out on context dependencies created by boundary conditions of the agents. Second, rational choice theory does not account for the potential interplay between actions and the quality of long term beliefs. Our simulation model of iterated interaction shows that some strategic choices can indirectly influence an agent's future beliefs and, thus, her future actions and their resulting success. In Sect. 3, we will show that this applies in particular to maximizing expected utility in the context of iterated bargaining games. In Sect. 4, we relate these findings to a thicker notion of rationality, taking the agents' beliefs and, more generally, their informational economy into account.

2 Bargaining as a generative mechanism for inequality: theory and model

In a highly interactive and integrated economy, inequality should be conceptualized as an emergent property of a complex system. By focussing on bargaining games, we aim to capture the distributional component of joint production processes relevant to the emergence of inequality.² Bargaining games are directly related to distributional matters and hence to the emergence of equality or inequality. In the present simulation, we represent the agents' bargaining problem by an iterated chicken game.

We are interested in situations where a mutually beneficial economic endeavor is feasible, for example a joint production process or an employer hiring an employee for a certain job. We do not assume that both agents have symmetric or interchangeable roles in the process, as one may represent a potential employer and the other an employee. All we assume is that a successful production creates a surplus that is to be divided among the two agents. However, before production can start, agents have to agree on the division of the expected benefits. Only after having done so will they engage in the production process and distribute the benefits according to their prior agreement. An extended bargaining process is costly in itself. More specifically, we assume that the time spent on bargaining cannot be used for production. The longer the bargaining process lasts, the less time remains for actual production. This leaves actors in a situation where both have a strong interest in shifting the outcome in their own favor. One crucial way of doing so is to bargain long and hard to obtain a favorable deal. That is, *not to give in* to the opponent's demands, but wait until she accepts your terms. Imagine, for example, a situation in which two partners are bound together by an incomplete contract. Such an assumption is, for example, put forward by Hart and Moore (1999), who argue that in fact all contracts can be seen as incomplete and are hence subject to renegotiation after initiation. In such a setting, agents could possibly generate large gains, but only if they succeed in agreeing to a potential distribution. Lipman (1986) asserts that within many such situations "each party prefers a poor

² In general, economic interaction consists of first-level economic problems addressing the coordination problems surrounding the production of goods (i.e., positive-sum games) and bargaining games as zero-sum games representing the second-level economic problems of allocating a surplus generated.

agreement to no agreement” (Lipman 1986, p. 317). Hence, agents involved are faced with the challenge of establishing a common bargaining solution within a competitive environment. The crucial decision they have to make is *when* (if ever) to agree on a successful distribution. We hold that these *endurance competitions* depict a relevant feature of bargaining problems in the real world.

Classically, bargaining behavior is determined by when and how far to adjust one’s own demands in reaction to the opponent’s behavior. *Ceteris paribus*, the later and the smaller the adjustments an agent makes, the better the bargaining solution will be for her, once found. The downside to a tough bargaining strategy, obviously, is that it increases the expected time and hence the cost until a common solution is found. With the present simulation, we limit ourselves to a simplified bargaining process. For reasons of tractability, we represent the agents’ toughness in bargaining by a single parameter, denoting after how many rounds of unsuccessful coordination they are willing to adjust their demands. By a slight idealization, we assume that the agent who adjusts first does so in such a way that common ground is found. If both agents adjust simultaneously, they split the difference and meet in the middle.³

2.1 The bargaining model

To represent these situations formally, we constructed an agent based model of iterated bargaining encounters. At the beginning of each simulation run, players are matched together in pairs randomly for a fixed number of bargaining rounds. Whenever the maximal interaction time is reached, all matchings are dissolved; new random pairs are formed and then interact for the same fixed number of rounds.⁴ Players can neither choose their partners freely, nor leave a partner prematurely or stay longer than the fixed number of rounds.

The bargaining process is similar to a classic game of chicken: both agents have to decide between making a *high* or a *modest* demand on the surplus generated. They make their claims simultaneously and without knowledge of the other player’s action. We assume that a high demand from both players is *irreconcilable*. In this case, no pay-offs are generated and they need to enter into a further round of bargaining of the same game structure. Bargaining is continued until one or both agents lower their demands so that reconcilability is reached or until the maximal interaction time is up.

Every other combination, i.e., a high and a modest or two modest demands is admissible. In this case, the players start producing the surplus and divide all benefits produced by the distribution agreed upon. They will use each subsequent round to produce a surplus, which is distributed according to the agreed solution. That is, once agents agree on a distribution, they will not renegotiate, but keep producing as long as they are matched together.

³ Hence, we allow for egalitarian solutions. However, in light of the competitive environment, these are not reached through egalitarian offers, but they arise if both partners have a similar bargaining strength.

⁴ Within each run, the number of interaction rounds is constant. Between different runs, we varied the values of this parameter systematically between 10 and 20.

		Player Y	
		<i>high</i>	<i>modest</i>
Player X	<i>high</i>	(0, 0)	(3, 1)
	<i>modest</i>	(1, 3)	(2, 2)

Fig. 1 Normal form of the baseline game

The one-round bargaining situation is represented by a game similar to a chicken game, which can be seen as a simplified version of Nash's demand game (Nash 1953). In this game a total common resource worth a utility of 4 for either agent is to be distributed (see Fig. 1). A high demand is represented by a utility of 3 and a modest demand by a utility of 1. Both players make their demands simultaneously. The demands are reconcilable if their sum does not exceed the value of the resource. If demands are irreconcilable (i.e., larger than 4 in sum), neither player receives anything. If demands are reconcilable, each agent receives what she had demanded. Additionally, by a small deviation from the original framework, if both agents make a modest demand, we assume the remaining resource is divided equally between both agents, i.e., both receive a utility of 2.

Given that there are only two possible moves in the one-shot game, only one of which guarantees successful cooperation, a player's strategy in the iterated game is determined entirely by specifying how long she is willing to maintain a high demand when no cooperation is achieved.⁵ For obvious reasons, we call this parameter an agent's 'toughness'. A toughness of 0 thus denotes agents who start off with a modest demand, while agents whose toughness is equal to the number of interaction rounds will always place high demands. Hence, one could describe the bargaining process between two players as the player with the higher toughness holding out until the other player gives in. The lower of the two toughness values determines after how many rounds this will be the case. Thus, the player giving in first determines how much time is spent bargaining, how many rounds are left for the production of a surplus and who gains what from the interaction. Finally, if both agents have the same level of toughness, this will result in an egalitarian distribution of the surplus generated in the remaining time.

The rationale behind the choice of toughness can be intuitively interpreted as follows: a high value aims at making the opponent give in eventually, so that the player receives a high pay-off from that point on. A high toughness, however, bears the risk of forgoing many rounds without income, before either player gives in. In the worst case, such a strategy can result in minimal gains if an agreement is found too late

⁵ Skyrms (2014) emphasizes the existence of a fair solution as a special focal point. However, in most economic interactions, no saliently egalitarian distribution exists. When an employer bargains about wages with an employee, the different possible agreements might benefit either the employer or the employee more. No possible wage offer could be regarded as obviously egalitarian in any way. Notably, we do not preclude that some distribution actually divides the achieved surplus fairly. We merely claim that this property is by no means transparent enough for both parties involved to qualify as a focal point. We hold that this is a common feature of many bargaining situations, from wage negotiations to partners with complimentary skills setting up a joint endeavor. While a numerical representation of pay-offs in the chicken game seems to suggest that there is an egalitarian solution, this might often be an artifact of the modeling.

or the opponent turns out to have an even higher level of endurance. An agent with low toughness, in contrast, rather gives in earlier in order to avoid enduring too many rounds without income. Accepting a lower pay-off in each successive round after having given in is thereafter the price that must be paid.

Another way to describe the situation is based on the recognition that the one-shot-game offers two Nash-Equilibria in pure strategies. These are when one player chooses the ‘high’ strategy while the other player opts for ‘modest’. Each player has an interest in establishing the equilibrium with pay-offs in her favor. For this, it is necessary to start with the high demand, and then to wait until the other player switches to ‘modest’. The rationale of the bargaining game therefore symbolizes a classic war of attrition.

2.2 Learning and information processing

Agents can use the experience gained in previous games for choosing their level of toughness. Crucially, we do not assume that agents learn specific information about certain individual others. Each pairing is resolved after a finite amount of time and agents do not expect to encounter the same opponent again. They can, however, form expectations about how unknown others in society behave or how frequent the different values of toughness exist. These expectations are formed purely on the basis of agents’ experiences in previous interactions.

To understand the mechanics of the game better, we need to distinguish between two different types of information a player can gain. First, if the agent ‘wins’ a game, i.e., if the opponent concedes first or both give in at the same time, she receives an exact signal about the toughness of her opponent: if an opponent gives in after 3 rounds, the agent knows for certain that the opponent’s toughness was 3. The second type of information concerns the case where a player ‘loses’ her current game, i.e., she decides to give in before her opponent does. In this case, she receives only an imprecise signal about the opponent’s toughness, namely that it is larger than her own. To incorporate both cases into a common framework, let us call a level of toughness t *compatible* if the assumption that the opponent’s real level of toughness was t is consistent with the agent’s information. That is, if our agent ‘won’ the game, the only compatible level of toughness is the one the opponent actually played. If the agent ‘lost’ the previous game, on the other hand, all levels of toughness t are compatible that are above the toughness our agent played herself.

While our agent may not learn exactly which toughness her opponent played, she can at least form probabilistic beliefs thereabout. Based on her observation and her initial beliefs about the distribution of toughness, she can calculate a probability distribution p_{obs} about the compatible levels of toughness, i.e., the possible levels of toughness the current opponent may have played. This distribution p_{obs} is derived from her initial distribution p_{old} by means of conditionalisation:

$$p_{obs} = p_{old} | \text{Observation.}$$

More explicitly, for any compatible level of toughness t we have $p_{obs}(t) = p_{old}(t)/T$, where $T = \sum_{t' \text{ compatible}} p_{old}(t')$.

The information gained is incorporated into the agent's subjective probability distribution of other agents' toughness by means of an update. Agents employ their own experience in bargaining situations to assess the overall distribution of toughness within society. However, this distribution may change over time, as other agents adapt their toughness or agents may enter or leave the scene. Hence, agents will reasonably discount older information in favor of their most recent pieces of information. Following these considerations, we can state our updating rule explicitly. The agent incorporates her new piece of information through updating her initial probability distribution p_{old} about the opponents level of toughness to p_{new} by the rule: $p_{new} := 0.9 \cdot p_{old} + 0.1 \cdot p_{obs}$. Plugging in the definition of p_{obs} , this amounts to:

$$p_{new}(t) = \begin{cases} 0.9 \cdot p_{old}(t) + 0.1 \cdot \frac{p_{old}(t)}{T} & \text{if } t \text{ is compatible} \\ 0.9 \cdot p_{old}(t) & \text{if } t \text{ is incompatible} \end{cases}$$

for all threshold levels t .⁶ The factor $\frac{1}{T}$ ensures that p_{new} again is a probability distribution, i.e., probabilities sum up to one. At the beginning of a simulation run, agents start with a uniform distribution, i.e., they consider all values of toughness equally likely.

2.3 Strategies

Toughness is an agent's key strategy parameter. The choice of toughness reflects a variety of different strategic considerations, such as reducing possible losses, maximizing possible gains, maximizing expected utility, or, on the more epistemic side, finding out when an opponent is likely to give in. In general, agents can adapt their toughness from game to game, depending on which value of this parameter they find most promising. However, in choosing the level of toughness, different strategies might be cognitively more or less demanding. Normally, the choice of strategy may depend on the information available to agents, their current financial situation and their available cognitive resources.

In the current model, we explore five different strategy types that guide the choice of toughness. Later, we will add two further types for the sake of conceptual exploration. By a strategy, we mean the way in which an agent updates her toughness after having encountered another agent. In no way do we claim that these choices exhaust the realm

⁶ This rule is conceptually related to Erev and Roth (1998) type reinforcement learning. Also in this framework, past information is temporally discounted in order to facilitate learning about a changing target system. Unlike in our model, in most approaches to reinforcement learning (Bush and Mosteller 1953; Skyrms 2014) agents do not form beliefs about their environments or the behavior of an opponent, but simply learn which actions have proven more or less successful. For the present model, however, agents form and update *beliefs* about the toughness of others. This has two reasons. First, our agents assess the toughness distributions of an entire set of opponents. Different opponents will employ different strategies, hence even bounded agents will likely conceptualize and keep track of behavioral differences mirrored by the distribution. Second, the relationship between the agent's probability distribution and her optimal toughness is intricate. Even small variations in the distribution may trigger major differences in optimal toughness. Hence, we hold that forming beliefs about the toughness of others increases the agent's potential to select for optimal action.

of possible strategies, even in this already simplified scenario. Nevertheless, we seek to capture intuitions, arguments and heuristic rationales about how the game could reasonably be played, and to incorporate the strategies that are assumed to be the most prominent in such situations.

- **MaxEU** This first type always chooses the toughness that maximizes her subjective expected utility. Naturally, the utility gained depends on the player's own toughness as well as that of the opponent. A utility maximizer needs to estimate therefore how likely an opponent is to play the various toughness levels. She does so with the learning mechanism outlined above. Starting with a uniform initial distribution, i.e., no information at all, she gradually learns about the behavior of others and updates her distribution accordingly. This strategy type embodies the optimal strategy for a rational, risk-neutral player. Furthermore, according to the law of large numbers, this strategy should be expected to perform best in terms of long term wealth accumulation.
- **Maximin** Agents of this type always play a toughness of 0, no matter what. That is, they give in immediately, thus ensuring that a maximal amount of surplus is produced. The rationale behind the Maximin's strategy is that this player type is 'infinitely risk averse'. The Maximin takes a guaranteed pay-off of 1 per round rather than risking any incompatibility. As the name suggests, this strategy embodies the classic *Maximin* principle.
- **Maximax** This type never gives in, and is prepared to outwait her opponent at any cost. That is, the Maximax's toughness is always set to the maximal possible level. The rationale for this strategy is the opposite of 'Maximin'. Maximax agents are willing to accept any number of incompatible rounds for the chance to obtain the maximal possible pay-off of 3, i.e., it represents the strategy of an agent who is maximally risk-seeking. This strategy follows an iterated *Maximax* principle.
- **Experimenter** Experimenter is a mixture between the types of MaxEU and Maximax. Before each game, this type chooses between two possible behavior types. With a probability of 90%, an Experimenter adopts the MaxEU strategy. However, with the remaining probability of 10%, she adopts the Maximax strategy. The reason for doing so is different though. Whilst the original Maximax strategy is simply prepared to sacrifice everything for the prospect of high per-round-income, the main motivation of experimenters is to gain information about their opponents' behavior, in order to make better choices while playing the MaxEU-strategy. The information gained is highest when outwaiting the opponent, since it is only then that an agent receives an unambiguous signal about the opponent's toughness. Hence, the choice of Maximax is the only strategy guaranteed not to give in first.
- **Increase–decrease** This player-type follows a simplistic way of updating her strategy: Whenever such a player loses a game, she reduces her toughness by one. Vice versa, whenever she wins a game, she increases her toughness by one. The reasoning of this heuristically-rational strategy can be explained as follows: When the player loses, she holds that she is not able to maintain her demand long enough to secure a high pay-off. She thus reasons that she must give in a little earlier, in order to receive at least the lower pay-off over a greater number of compatible rounds. After a victorious round, however, the agent learns that being tough paid

off in the end. She is thus encouraged to be even tougher in the next encounters. Increase-decrease describes a decision-making heuristic that is clearly suboptimal, yet may characterize the behavior of some real life agents appropriately, as argued by Mishra et al. (2015). Increase-decrease players start with a random toughness in the very first game.

2.4 Simulation experiments

We simulate the described model and analyze its output on the basis of three major experiments. The first series constitutes the baseline model, where we are interested in the average long term success of the different strategies only. The second experiment introduces evolutionary mechanisms, while the third studies whether or not evolutionary pressure is structurally different for rich and poor agents. To do so, we add an additional model parameter, *cost of living*. While players receive pay-offs in the games they play, they also incur a certain cost c for maintaining their lives through each round. In order to stay alive, each player must spend c from her accumulated wealth in every round. Agents that fall below a wealth of zero fall prey to evolutionary pressure. They are removed from the simulation and are replaced by a new agent that mimics one of the survivors in her choice of strategy. At the beginning of a simulation run, each agent is equipped with a small initial endowment. In the third experiment we introduce two classes of agents, rich and poor. These differ in their initial wealth endowments.

We use three different output measures for analyzing the model. First, we are interested in the wealth accumulated by agents of different strategy types. We take wealth as an indicator for the long term bargaining success of the different strategy types. Second, in those experiments with evolutionary pressure, we measure the proportions of player types at the ends of simulation runs. These proportions of player types are used as a measure for the bargaining success of different strategies. Under evolutionary pressure, well performing strategies will survive and reproduce, while inefficient strategies are more likely to die out. Thus, successful strategies will be played by many at the end of a run, while unsuccessful strategies will barely be present. Third, we are interested in the content and quality of information collected by the different strategies. Each agent collects new information while being engaged in bargaining situations. Since an agent's toughness impacts the quantity and quality of information gained, the various agent types might differ structurally in the content of beliefs they hold at the end of a simulation run. We compare the *accuracy* of beliefs held by the different agents by means of a proxy measure. Each simulation run in these three experiments starts with a total of 100 agents, 20 of each from the five types described above.⁷

⁷ To check robustness, we ran various simulations with other distributions of the agent types and for different numbers of bargaining rounds. We also varied the overall length of a simulation run up to 100,000 steps and studied control runs, where we initiated wealth measurements only after 1000 steps, so that agents had time to calibrate their behavior. For a broad variety of input parameters the results are similar to those reported here. Most notably, the performance of Maximax agents varies. Depending on agent distribution and the number of bargaining rounds per game, Maximax agents may perform very well or fall behind Experimenter

3 The poor performance of EU maximizers

This section presents the results of the three experiments detailed above. All data is based on the range of parameter values described earlier, with 100 model runs for each parameter combination. A simulation run lasts for 1000 interaction rounds, before final measures are taken. At all times, the 100 agents are paired up in 50 couples, playing the iterated bargaining game described above. One round for the model corresponds to a single step in these iterated bargaining games. As each iterated bargaining situation lasts for 10–20 steps, depending on the parameter *bargaining rounds*, every agent engages in 50–100 different bargain games throughout a simulation run. For the second and third set of simulations, we aborted the simulation prematurely if no agent had to exit due to negative funds for 100 rounds. If a group of agents died out completely, its mean wealth was set to zero.

3.1 The baseline model

The baseline model works without evolutionary pressure and no cost of living. All agents start with the same wealth of zero, which makes the final wealth at the end of a simulation a direct measure of different strategies' bargaining success. In this experiment, the types Maximax, MaxEU and Experimenter all perform well, while Maximin performs by far the worst in terms of average income per round, see Fig. 2. These outcomes demonstrate that, in basic terms, the maximin approach comes at a price. Each time a player makes a high demand, she risks a round of failed coordination and hence no production, without being guaranteed to gain more later in the interaction. The Maximin strategy avoids such risks. It settles for a secure pay-off of one unit per round, with the slight chance of receiving two, should both players give in at the first round.

We should, however, emphasize that these effects are in some way dependent on the distribution of strategy types. In a small class of distributions which are somehow degenerate, Maximin might perform extremely well. This can be the case, for instance, when a few agents of this type enter a society almost exclusively consisting of Maximax types. For the more regular distributions of agent types, however, the results are similar to those presented here. Notably, the performance order among the three best ranked strategies, Maximax, MaxEU and Experimenter, depends upon a variety of input parameters such as the exact distribution of agent types, the number of bargaining rounds or the overall number of simulation steps. Depending on these settings, either Maximax or Experimenter fares best. The Experimenter strategy, however, always outperforms MaxEU in all settings tested.

The second set of results concerns the role of information and the quality of agents' beliefs. Some strategies, such as MaxEU or Experimenter, adjust their toughness constantly, based on their beliefs about the behavior of others. Naturally, these strategies

Footnote 7 continued

and MaxEU. Our main finding, however, holds for a large class of parameter values: Experimenter agents always outperform MaxEU types and, in many settings, also Maximax types.

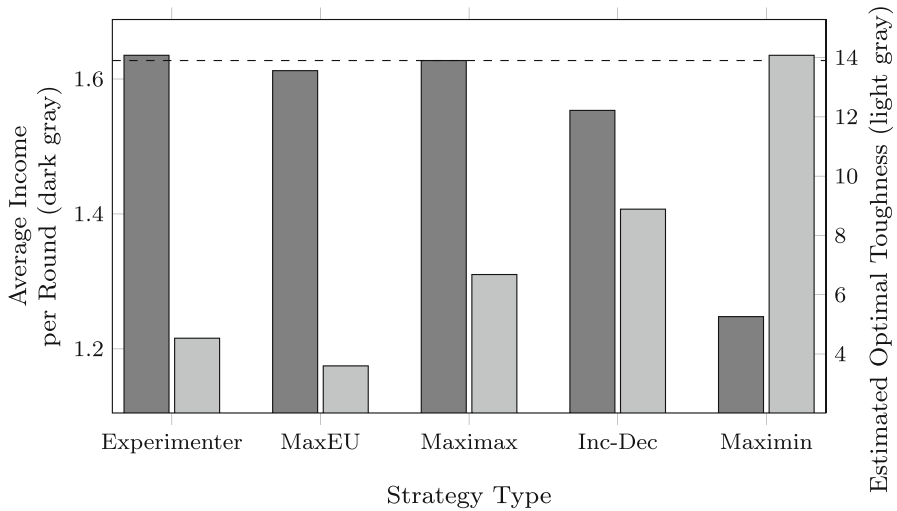


Fig. 2 Average income per round (dark gray) and EU maximizing toughness (light gray) of different agents

can only be successful if their beliefs are sufficiently accurate. The fact that MaxEU does only moderately well in terms of average income may hence be attributed to inaccurate beliefs. Notably, these beliefs are updated with every bargaining game the agent engages in. An agent's beliefs hence develop parallel to her game play. However, more is true: agents acquire precise information only of those opponents that give in before them. While playing a moderate toughness, the agent will be able to form adequate beliefs only about the distributions of low and moderate levels of toughness among other agents. In order to learn about the behavior of high toughness agents, the agent needs to play tough herself. That is, an agent's action has a structural impact on the *type* of information collected. Certain strategies may hence be structurally correlated with more adequate beliefs than others. Notably, such influence is not depicted in classic rational choice theory, where beliefs are treated as uninfluenced by the agent's actions. This action-belief dependence complicates the agent's strategic considerations. Rather than focusing on monetary gains alone, she needs to focus on both, ensuring a sufficient quality of beliefs while also maximizing utility income. We will show that this consideration, indeed, explains much of the deficits of MaxEU. The performance of Experimenter strategies will be traced back, largely, to the fact that these are better in acquiring adequate beliefs than classic MaxEU strategies.

Recall that only the 'winning' player gains exact information about her opponent's toughness, while the player giving in first merely receives imprecise information about the opponent's toughness being above hers. Hence, strategy types that tend to give in first (i.e., Maximin) will collect very little information. On the other end of the spectrum, strategies that never give in, such as Maximax, will collect the maximal amount of information possible. However, we are not so much concerned about the quantity of information than its content. As each agent's belief corresponds to a probability distribution over the different toughness values, we could compare these by their mean

or median. For the present purpose, however, we chose an operationalization that is closer to the agents' goal of selecting an optimal toughness level for themselves. For each belief, i.e., each probability distribution, we calculate which toughness would maximize expected utility in light of this belief. Hence, this measure specifies the toughness level an agent would choose, were she a MaxEU agent. As we are interested in the quality of an agent's belief over time, we average up this value – *estimated optimal toughness*' for short—over her playing career. We use this as a rough proxy to assess the content of an agent's beliefs or at least the beliefs an agent could have formed in light of the available information.⁸ We should emphasize here that most agents do not make use of the full information they collect. Only the type MaxEU and, in 90% of games, the Experimenter-types actually employ the information collected for calculating expected utilities. All other agents do not employ the information collected. In fact, for any interpretation of the model, we do not need to assume that those agents keep track of their incoming information at all. This may be important when discussing the cognitive demands of the different strategies. The current analysis can thus be described as comparing the different types' available information by discussing which beliefs they could have formed, had they processed the information available to them.

In our experiment, Maximin identifies the highest value of toughness on average for maximizing expected utility, see the light gray bars in Fig. 2. While this may initially seem counterintuitive, note that Maximin gains no information about any toughness higher than zero, and therefore never updates the higher parts of her probability distribution. Her optimal strategy is completely determined by her initial beliefs and thus by the uniform initial distribution we chose. Had we started with a different initial distribution, Maximin would have identified a different optimal strategy. Further, we note that all other agents obtain at least some information about others' toughness.

However, the relation between the amount of information gained and the optimal toughness value identified is not monotonic. Indeed, the types collecting least and most information, Maximin and Maximax respectively, are relatively close to each other in their assessments of optimal toughness, while both MaxEU and Experimenters locate the ideal toughness at a much lower level, see Fig. 2. Since Maximax agents never give in first and hence always learn the true value of each opponent's toughness, these results might be taken to indicate that more cautious agents underestimate systematically the potential of high toughness strategies.

This points to an intricate relationship. Not only do agents employing moderate levels of toughness collect less detailed information, but also the content of this information turns out to be skewed. Having skewed information leads to skewed beliefs which, in turn, may lead agents to significantly underestimate the benefits of insisting on high demands. This interpretation is further supported by analyzing the performance of the different strategies. By the law of large numbers, we *should* expect MaxEU to accrue the highest wealth in the long run. It does not. Rather, Maximax outperforms MaxEU significantly, see the dark gray bars in Fig. 2. One possible explanation for this shortcoming is that the collected evidence of 'MaxEU' agents is so strongly biased that

⁸ We have also considered different proxies for assessing the content of agents' beliefs, such as mean or median. These have turned out largely uninformative.

the subjectively optimal move differs widely from the objectively income maximizing toughness.⁹ We want to inquire deeper into the role of information in this process. As depicted in Fig. 2, Experimenters are apparently able to benefit from their strategy to be Maximax in 10% of the cases. There are at least two possible explanations for this. Either, an Experimenter benefits from the increased information that she gains in the 10% of cases where she plays Maximax (1), or an Experimenter merely benefits from the fact that Maximax perform better than MaxEU agents in some cases (2).

To decide between these two hypotheses and to carve out the benefits from improved information, we perform a further experiment with a new strategy type called *NoLearnex*. This type chooses her actions in the same way as Experimenter, i.e., a mixture of 90% MaxEU and 10% Maximax. On the epistemic side, however, the strategies differ. *NoLearnex* fails to incorporate any information acquired in the Maximax role. That is, *NoLearnex* combines the strategy choice of an Experimenter with the knowledge only from those games where she assumes the role of a MaxEU-agent. If this strategy fares as well as an Experimenter, we know that Experimenter's success is caused mainly by the higher benefits achievable as a Maximax. However, if Experimenter outperforms *NoLearnex*, we can infer that this difference must be caused by differences in the available information, i.e., the information collected as Maximax. As the dark gray part of Fig. 3 shows, Experimenter outperforms *NoLearnex* significantly (see dashed line), thus supporting the first hypothesis. Or to put it differently: It is beneficial to be very tough and act as a Maximax every now and then, not only for its own expected utility, but also for the sake of collecting information that makes it possible to act optimally in future interactions. For the sake of comparison, we also plot the income an EU maximizing strategy could make were it based on fully accurate information about the current distribution of toughness. This information is listed as MaxEUp_{perfect}. The light gray part of Fig. 3 shows which toughness values the different strategies identify as optimal. MaxEUp_{perfect} thereby constitutes the reference point for what would have been the optimal toughness choice in such a situation.¹⁰

We should highlight an intricate property of the strategy MaxEU that follows from this analysis: While this type maximizes expected utility in light of current beliefs, the information collected along the way leaves MaxEU not only poorly informed, but also actively misinformed. It is exactly this misinformation that allows MaxEU to be outperformed by Maximax, Experimenter and *NoLearnex*. In other words, MaxEU actively undermines the basis for its own success in the long run. The Experimenter's success, conversely, shows how it can pay off sometimes to invest in examining the behavior of others. Through its occasional adoptions of the Maximax behavior, this strategy has more accurate beliefs at its disposal, allowing it to perform far better in the long run. Conversely, proponents of MaxEU-type strategies face a classic trade-

⁹ A further candidate explanation for MaxEU's lack of performance is that the chosen uniform initial beliefs are far off. To rule out this interpretation, we ran an extended simulation run over 2000 interaction rounds and measured the wealth accumulated only in the last 1000 rounds. This setup allowed MaxEU agents to adjust their beliefs well before pay-off collection began. The observed wealth distribution was similar to the one depicted in Fig. 2, ruling out this explanation.

¹⁰ Between-type performance differences in comparison to the first experiment stem from the altered agent constellation, which slightly favors some types more than others relative to the prior experiment.

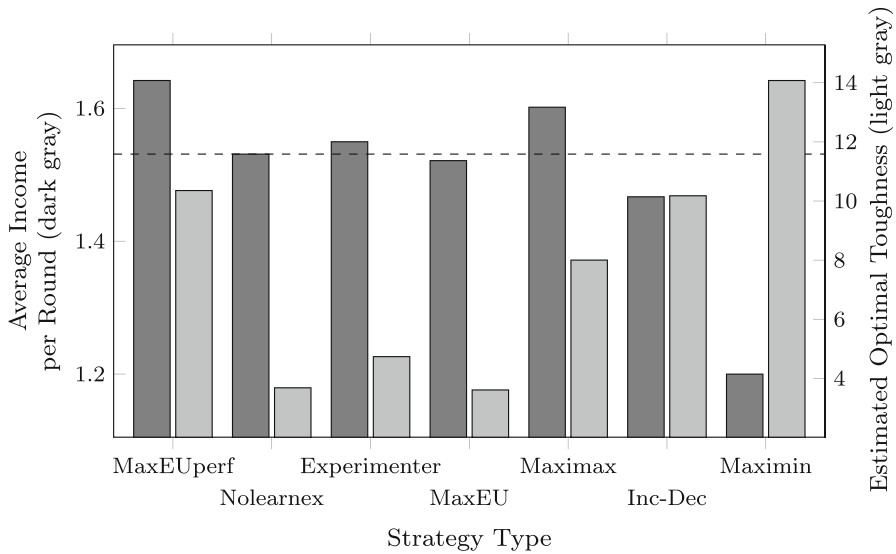


Fig. 3 Average income per round (dark gray) and estimated optimal toughness (light gray) of different agents

off between exploration and exploitation. They have to decide how often to engage in information searching and when to exploit the information gained. Or, to put things differently, such agents face a constant trade-off between short term rational action (in terms of maximizing expected utility) and long term performance due to rational collection of evidence.

3.2 Evolution in a world of poverty

We now shift focus from long term maximization to questions of survival. In a first extension of the baseline model, we introduce a slight evolutionary pressure on the individual agents. More specifically, a cost of living c is introduced that agents have to pay each round. This cost of living is constant within a simulation run and the same for every agent, yet varies between different runs. Once an agent has negative wealth when leaving an interaction cycle, this agent is removed from the simulation. To fill the gap, one of the remaining agents is chosen at random and duplicated, thus keeping the number of agents constantly at 100.¹¹ With this simulation, we address matters of survival and evolutionary success. If a strategy cannot guarantee survival,

¹¹ This framework is in line with a modest evolutionary mechanism. The new agent is an exact copy of the old with all her variables, including her wealth. The newcomer is not to represent a newly arriving, vulnerable agent, but should be seen as a further representative of the original society. Compared with most accounts in the literature on epistemic game theory (Easley and Kleinberg 2010, chapter 7), the evolutionary mechanism described here is still rather mild. It is not the absolute economic success of different strategies that determines their evolutionary success, but merely whether or not they generate sufficient funds for survival.

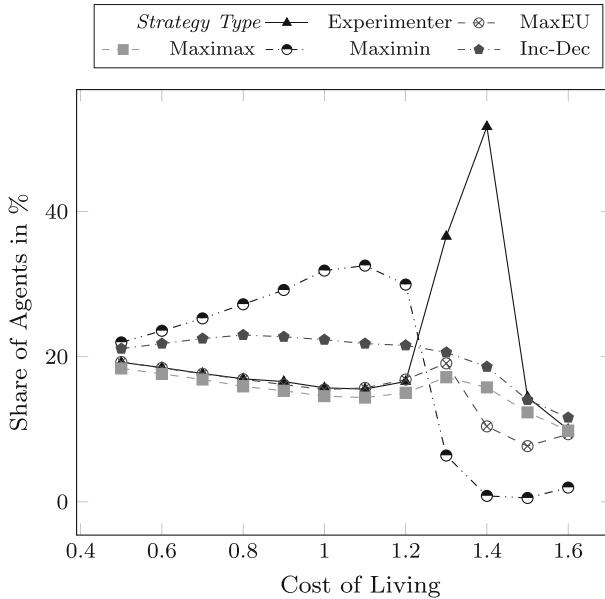


Fig. 4 Evolutionary fitness of different types

some players with this strategy might be forced out. Successful strategies, on the contrary, can increase their share through the replacing mechanism of evolution. At the beginning of each simulation run, there are 20 agents of every type, each with a starting wealth between 5 and 20 drawn from a uniform distribution. The number of agents per strategy at the end of a simulation run is thus a direct measure for its success.

When the cost of living is too low, below 0.3, there is no significant evolutionary pressure and all strategies maintain their initial share. For a moderate cost of living between 0.3 and about 1.1, the strategy type Maximin has a significantly higher chance of survival and hence constitutes a higher share of the population than all other agent types, see Fig. 4. Notably, this performance does not translate into expected wealth: in line with findings from the first experiment, Maximin still performs worse in terms of aggregate wealth than all other agent types. The explanation for this is as simple as it is instructive: although Maximin does not perform well in terms of average income, none of its adopters perform poorly enough to have her wealth fall below zero in which case evolutionary pressure would kick in. For a moderate cost of living, it is precisely the maximin approach that guarantees short- and medium-term survival. All other strategies may perform much better in terms of wealth, yet there are still some that do not survive due to unlucky circumstances, such as trying a high toughness strategy yet having to give in eventually.

Adopting the Maximin strategy maximizes the chances of survival, yet fares poorly in terms of long term expected gains. Hence, there is an inherent trade-off between two sets of goals a rational agent might pursue: survival and income maximization. To put this in a language of rationality: Different strategy choices can be rationalized

by weighing these objectives differently or by using one as a boundary condition for the other.

For a high cost of living, between 1.2 and about 1.4, the share of surviving Maximin decreases abruptly and dramatically. At high costs of living, the moderate yet guaranteed income of a maximin strategy can no longer provide sufficient means for survival. The Maximin strategy guarantees survival only as long as the costs of living are not too high. If the environment is more hostile, survival requires risk-taking and the risk-free Maximin strategy does not flourish. At the same time, all other strategies remain constant or perform slightly better than they do for a medium cost of living. We attribute this mainly to the poor performance of Maximin: as Maximin dies out, other strategies occupy a higher share of the population. There is, however, one notable outlier. The Experimenter-type exhibits a sharp increase in population size at a cost of living around 1.4. This can be explained by the previous findings: experimenters are more successful on average than all but the Maximax. They are therefore more likely to survive and reproduce compared with these others. Maximax, conversely, as the only strategy earning a higher average income than Experimenters, fall prey to their risk taking behavior. This type is willing to take arbitrarily high risks, hoping eventually to make the opponent concede. While this strategy might be successful on average, it produces a high variance with longer stretches of close to no income. If such stretches grow too long or too frequent while the cost of living is substantial, the agent's wealth might temporarily turn negative, at which point the evolution mechanism kicks in.

3.3 Third series of experiments: evolution in a heterogeneous world

The previous experiment revealed an inherent conflict between two dimensions of rationality: survival and maximization. Any choice of strategy needs to make a trade-off between these two, as none of the strategies studied so far fared well with regard to both dimensions simultaneously. Within a third set of experiments, we introduce a further dimension that could be responsible for long-lasting inequality: the initial wealth endowment of agents. For this, we distinguish two types of agents, *rich* and *poor*. We then inquire whether both types face similar problems of trading off survival and maximization, or whether the dimensions in strategy choice are significantly different for rich and poor types.

This experiment is based on the same scenario as the previous simulation. However, before starting the simulation, we randomly select a proportion of agents and label them *rich*. These agents receive a bonus of 50 on their initial wealth. Across different simulation runs, we vary the proportion of rich agents between 10 and 90%.

In terms of survival rates, poor agents display a similar behavior compared with the agents studied in the second series of experiments. This is far from surprising as these agents have exactly the same starting conditions as in the previous experiment. In particular, at a moderate cost of living, Maximin have a significantly higher chance of survival than other poor agents, see left side of Fig. 5. The mechanism behind this finding is the same as identified above: While other strategies might be superior in terms of long term gains, they carry the risk of falling below the poverty threshold of zero.

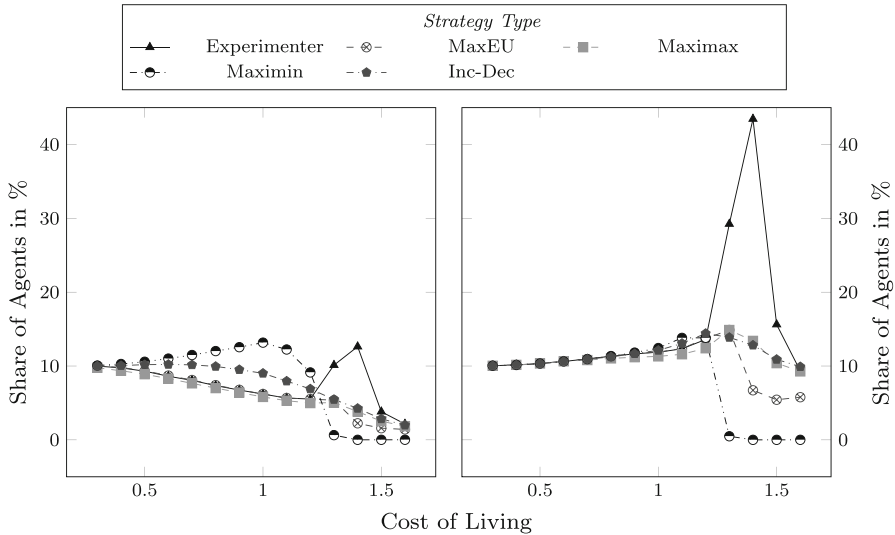


Fig. 5 Evolutionary fitness of different types starting as poor (left) and rich (right)

Rich agents, on the other hand, display a completely different evolutionary pattern. At moderate costs of living, all richer agents perform similarly well in terms of survival rates. Thus, the mechanism identified above does not apply to rich agents. Their prior endowment helps them to survive even moderately long spells of low income unharmed. Rich agents do not face the same trade-off between maximizing expected gains and ensuring survival as poor agents at moderate costs of living. Rich agents are thus less constrained in their choice of strategies. They have access to a set of potentially high gain strategies that might be inadmissible for poorer agents due to their risk of non-survival.

In terms of rationalization, this pattern is reversed. For poorer agents, many different strategies could qualify as rational, depending on the trade-off between maximization and chances of survival. In the extreme case, even the Maximin strategy can be understood as rational, for it maximizes the rate of survival. This is not the case for rich agents. For these, the Maximin strategy is always suboptimal, as it performs poorly in terms of expected gains and does not give an edge in terms of survival.

A second finding concerns the inclusion of a competitive cost of living between 1.0 and 1.4. In this case, rich players are also highly privileged compared with their poorer counterparts. Moreover, as Maximin perform poorly in this region, there is no longer a safe strategy for poorer agents. These cannot escape the increased evolutionary pressure and die out. Once again, there is one notable exception. Within a certain, narrow cost of living margin, the Experimenter strategy fares extremely well, as already noted above, indeed well enough to enable even poor Experimenters to have a sufficient chance of survival. Notably, this strategy is the most cognitively demanding, as it requires agents to assess expected utilities and balance between exploration and exploitation. In other words, poor agents may be able to survive in a competitive market only if they command sufficient cognitive resources.

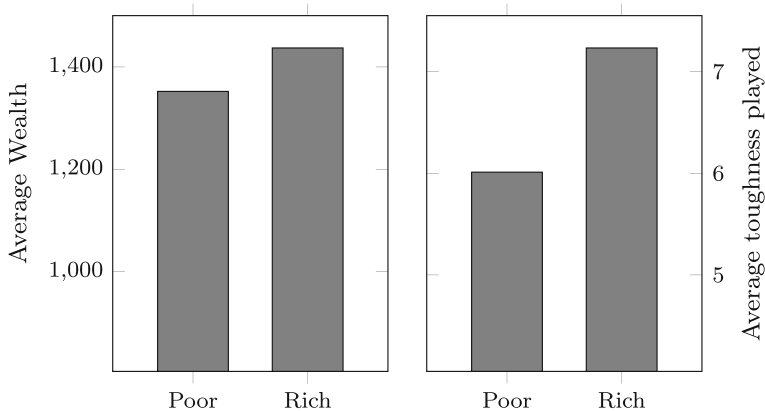


Fig. 6 Differences in wealth (left) and average toughness played (right) between rich and poor at the end of a simulation run

We should highlight two consequences from these results. Firstly within the third set of simulations, the rich–poor gap widens compared with inequality at the time of the model’s initialization. While the average initial wealth gap between rich and poor is 50 points, it grows considerably to 85 at the end of the simulation—see the left side of Fig. 6. As we have seen in the second experiment, evolutionary pressure among the poor agents affects everybody but the Maximin agents. Hence the latter will, in the long run, have a large share among the remaining poor agents. The first experiment, however, suggests that the Maximin strategy fares worst in terms of long term accumulated wealth. Hence, having a large share of Maximin is detrimental to the long term economic development of poor agents. No such reasoning applies to rich agents, as Maximin are not favored by evolutionary pressure there.

Secondly in the long run, rich agents develop a higher toughness on average than poor agents, see right hand side of Fig. 6. In other words, richer agents have higher bargaining power and they are more likely to settle distributional quarrels in their favor. Notably, this development is fully endogeneous to the simulation, as both classes start with the same average toughness. The driving mechanism, again, is the different compositions of poor and rich classes in terms of strategies at the end of a simulation run. Having a larger share of Minimax agents with a toughness of 0 causes the class of poor agents to have a lower toughness on average. Taking the evolutionary mechanism as representing the learning process of rational agents, our simulation demonstrates how and why rich agents learn to negotiate tougher.

4 Rationality revisited

There are three main findings from our analysis. First, contrary to the predictions of Rational Choice Theory, the MaxEU strategy is not optimal at maximizing pay-offs in the long run. In short, the information collected while playing this strategy is structurally biased and hence utility calculations rely on false beliefs. Second, in some circumstances, the worst strategy in terms of expected utility might turn out

to be the best for maximizing chances of survival and vice versa. Third and finally, the considerations and boundary conditions for strategy choice might be structurally different for rich and poor agents. In this section we want to discuss some implications from these findings for the concept of rationality as it is used in Rational Choice Theory.

In economics, rationality is defined instrumentally as the ability to adopt the best actions to achieve given goals. Elster (1988), on the other hand, argues that this conception may be too narrow. He identifies three places where rationality plays a role in the concept of Rational Choice Theory:

1. The set of beliefs an agent holds has to be internally and externally consistent.¹²
2. The set of desires an agent holds has to be internally consistent.¹³
3. The desires and beliefs of an agent must cause an action in the right way and this action has to be the best choice given individual desires and beliefs.

Rational Choice Theory formalizes the latter property in an axiomatic way. From the set of all available alternatives, the agent chooses (or should choose) the strategy that maximizes expected utility. For decisions under risk, agents hold probabilistic beliefs about the possible consequences, making it possible to calculate expected utilities. Agents compare the expected utility of the various available options in order to identify which action is optimal, given the agents' desires. Formally, the expected utility of an alternative a is defined as:

$$EU(a) = \sum_j p_j \cdot u_j$$

where u_j stand for the utilities of various possible outcomes of action a , whilst p_j denotes their respective likelihoods. Agents can (or should) then compare different actions by their expected utility and choose the option that offers the highest expected utility. Classifying an agent as irrational could classically be interpreted as saying that she fails on one or more of the conditions (1), (2), and (3). Much of rational choice theory focusses on condition (3) exclusively, taking (1) and (2) as independent of (3) and as given.

Additionally to condition (1), the amount of evidence available to an agent plays a role (Spohn 2002). What beliefs are adequate for a given body of evidence might not only depend on the content of that evidence, but also on the quantity of evidence available. In some situations, rationality might require agents to accrue sufficient levels of evidence for their beliefs, if feasible. Collecting additional evidence, however, may come at a cost. Thus, the agent is faced with a strategic problem of how much evidence to acquire. A fortiori, the different facets of rationality may not be independent, but they can be intertwined in a complex manner (Weisberg and Muldoon 2009). This shows in the present model, where the agents' information acquisition and their strategic choices are in a complex interdependency. Each move in a bargaining situation generates a new

¹² For example, on the internal side, probabilities need to sum to one, impossible events need to receive a probability of zero and so on (Ramsey 1931). On the external side, the agents beliefs have to respond to the available evidence in the right way.

¹³ This means, for example, that the order of preferences has to be transitive, complete and connected (Davidson et al. 1955).

piece of evidence about how the opponent reacts. Thus, different bargaining strategies generate different flows of evidence.

Furthermore, the different strategies not only differ in the amount of information acquired, but also in its content. The agent's action indirectly influences her future beliefs and thus her future actions in a substantial and possibly non-neutral way. Hence, in choosing a strategy, an agent must not only make sure that she maximizes expected utility, given her current beliefs, but must also take care to ensure sufficient quality of her future beliefs. An agent may need to sacrifice short term income for creating an adequate evidential basis for future actions. As poorer agents may find themselves unable to make such short term sacrifices, this trade-off may deepen existing inequalities. Only those who can afford to forgo current income for collecting information can hope for long term optimality. Those who cannot may be structurally challenged in their quest for a better future.

We now discuss our main findings in light of Elster's conceptualization of rationality. Our first result is that MaxEU fares structurally and significantly worse than other strategies. In focusing on short term maximization only, this strategy leads the agent to build up significantly false beliefs over time, thereby undermining the basis of its own success. The fact that MaxEU fares far from optimally suggests that a robust theory of rationality should encompass more than the single dimension of utility maximization. A second necessary dimension is constituted by rationality norms on the collection of evidence. Against the understanding of Rational Choice Theory that these dimension are independent of each other, our results give evidence that information search and strategic behavior are complexly intertwined. The choice of action impacts the amount and content of information collected, which in turn may influence future behavior. In particular, agents might need to balance between maximizing short term utility and collecting highly accurate information. A *thick* notion of rationality needs to take both these dimensions and their interplay into account.

The second finding sheds light on a further, often overlooked aspect of rationality. It invites the argument that long term maximization goals may need to be balanced against or are constrained by short term needs created, for instance, by the accruing costs of daily life. It is no use having the highest expected utility in the long run if short term random fluctuations impede ever reaching that point. Or to put it in more mathematical language, it is not enough to consider the regularities from the law of large numbers; one must also consider the random irregularities of small numbers. Moreover, the current simulation suggests that such constraining factors might be overproportionally strong for the poor, who cannot afford higher risk levels or longer periods without income. By systematically constraining the poor in their choice of available long term strategies, side constraints may widen rather than close an existing wealth gap.

Finally, the third finding illustrates that strategies are not in themselves rational or irrational (Galeazzi and Franke 2017). Rather, they depend on the context in which a player is situated, how other agents act and also the agent's endowment of information as well as resources. This finding is in line with Gigerenzer and Gaissmaier (2011) and their concept of 'ecological rationality'. What action should count as rational, they argue, can be assessed only relative to the situational context. In the context of our simulation, certain strategies that are highly beneficial in the long run may be admis-

sible for rich people, but carry too high a risk of failure for poorer agents. Conversely, what may be a defensible pick for poorer agents—sacrificing long term pay-offs in exchange for guaranteed short term income—might be completely unreasonable for richer agents not facing short term existential threats. Generally speaking, what is rational for one agent might not be rational for the next, even if both share the same goals and desires.

5 Conclusion

Our simulation experiments' findings contribute to two ongoing debates. The first concerns the nature of rationality as discussed in Sect. 4. By means of an agent-based model, we provide an example which illustrates that the standard conditions of rationality put forward by rational choice theory are too narrow for certain situations. In line with Elster (1988) we argue that a substantial account of rationality needs to address both principles of evidence collection as well as strategy and action choice. Both these dimensions are intricately connected.

Second, this paper contributes to a debate about the origins and structures of inequality. A recent body of work enquires into the various relations between inequality on the one hand, and rationality or cognitive resources and strategy choice on the other hand. With the present simulation, we show that even within societies of rational agents, existing inequalities might persist and aggravate over time. The driving force here is that rich and poor agents may be subjected to different boundary conditions of rationality. These findings directly relate to a variety of empirical results showing that poverty and inequality are correlated with a variety of distinct behavioral patterns concerning risk attitude (Carvalho et al. 2016). Our results suggest that, at least *prima facie*, these differences need not always be driven by a failure to act and choose rationally. Rather, they might be indicative of rationality's context-sensitive requirements that affect the rich differently from the poor.

Finally, our results invite some general normative conclusions about inequality. Even in a society of equally endowed rational agents, spontaneous local variations can generate momentary inequalities. As we have shown, these inequalities tend to aggravate rather than diminish over time due to the interplay between information acquisition and strategy choice. Put simply, one can succeed in bargaining over the long run only if one dares to take risks from time to time. Moreover, one must be able to afford experiments and failure, as one would otherwise not have access to relevant information about the behavior of other agents. Thus, those who are disadvantaged anyway, will also be systematically hindered in learning about their actual chances.

Acknowledgements We would like to thank Paolo Galeazzi, Frederik Van De Putte, two anonymous reviewers and audiences in Poznan, Oslo, Gent and Paris for valuable feedback and suggestions. The work of DK and JM on this project was partially supported by the Deutsche Forschungsgemeinschaft (DFG) and Agence Nationale de la Recherche (ANR) as part of the joint project Collective Attitude Formation [RO 4548/8-1]. The work of DK was partially supported by the Deutsche Forschungsgemeinschaft (DFG) and Grantov Agentura České Republiky (GAČR) as part of the joint project From Shared Evidence to Group Attitudes [RO 4548/6-1]. The work of SS was supported by the DFG through the Bamberg Graduate School of Social Sciences (BAGSS) and by the Humboldt-Foundation through the Munich Center for Mathematical Philosophy (MCMP).

References

- Atkinson, A. B., & Bourguignon, F. (2014). *Handbook of income distribution* (Vol. 2a–2b). New York: Elsevier.
- Banerjee, A. V., & Newman, A. F. (1991). Risk-bearing and the theory of income distribution. *The Review of Economic Studies*, 58(2), 211–235.
- Becker, G. S. (2013). *The economic approach to human behavior*. Chicago: University of Chicago press.
- Blume, L. E., & Easley, D. (2008). Rationality. In S. N. Durlauf & L. E. Blume (Eds.), *The New Palgrave Dictionary of Economics*. Basingstoke: Palgrave Macmillan.
- Briggs, R. (2017). Normative theories of rational choice: Expected utility. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Stanford: Metaphysics Research Lab, Stanford University.
- Bush, R. R., & Mosteller, F. (1953). A stochastic model with applications to learning. *The Annals of Mathematical Statistics*, 24(4), 559–585.
- Carvalho, L. S., Meier, S., & Wang, S. W. (2016). Poverty and economic decision-making: Evidence from changes in financial resources at payday. *The American Economic Review*, 106(2), 260–284.
- Dabla-Norris, M. E., Kochhar, M. K., Suphaphiphat, M. N., Ricka, M. F., & Tsounta, E. (2015). *Causes and consequences of income inequality: A global perspective*. IMF Staff Discussion Note SDN/15/13.
- Davidson, D., McKinsey, J. C. C., & Suppes, P. (1955). Outlines of a formal theory of value, I. *Philosophy of Science*, 22(2), 140–160.
- Easley, D., & Kleinberg, J. (2010). *Networks, crowds, and markets: Reasoning about a highly connected world*. New York: Cambridge University Press.
- Elster, J. (1988). The nature and scope of rational-choice explanation. In E. Ullmann-Margalit (Ed.), *Science in reflection: The Israel Colloquium: Studies in history, philosophy, and sociology of science* (Vol. 3, pp. 51–65). Dordrecht: Springer.
- Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *The American Economic Review*, 88(4), 848–881.
- Ferejohn, J. (1991). Rationality and interpretation: Parliamentary elections in early Stuart England. In K. R. Monroe (Ed.), *The economic approach to politics: A critical reassessment of the theory of rational action* (pp. 279–305). New York: Harper Collins.
- Galeazzi, P., & Franke, M. (2017). Smart representations: Rationality and evolution in a richer environment. *Philosophy of Science*, 84(3), 544–573.
- Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, 62, 451–482.
- Goldberg, D. E. (1989). *Genetic algorithms in search, optimization, and machine learning*. Boston: Addison-Wesley Professional.
- Gottschalk, P., & Smeeding, T. M. (2000). Empirical evidence on income inequality in industrialized countries. In A. Atkinson & F. Bourguignon (Eds.), *Handbook of Income Distribution* (Vol. 1, pp. 261–307). New York: Elsevier.
- Grable, J. E. (2000). Financial risk tolerance and additional factors that affect risk taking in everyday money matters. *Journal of Business and Psychology*, 14(4), 625–630.
- Green, D., & Shapiro, I. (1994). *Pathologies of rational choice theory: A critique of applications in political science*. New Haven, London: Yale University Press.
- Grüne-Yanoff, T. (2012). Paradoxes of rational choice theory. In S. Roeser, R. Hillerbrand, P. Sandin, & M. Peterson (Eds.), *Handbook of risk theory: Epistemology, decision theory, ethics, and social implications of risk* (pp. 499–516). Dordrecht: Springer.
- Halek, M., & Eisenhauer, J. G. (2001). Demography of risk aversion. *The Journal of Risk and Insurance*, 68(1), 1–24.
- Hart, O., & Moore, J. (1999). Foundations of incomplete contracts. *The Review of Economic Studies*, 66(1), 115–138.
- Haughton, J., & Khandker, S. R. (2009). *Handbook on poverty and inequality*. Washington, DC: World Bank Publications.
- Holland, J. H. (1992). *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. Cambridge, MA: MIT press.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 47(2), 263–291.
- Kanbur, S. M. (1979). Of risk taking and the personal distribution of income. *Journal of Political Economy*, 87(4), 769–797.

- Kolodny, O., & Stern, C. (2017). Evolution of risk preference is determined by reproduction dynamics, life history, and population size. *Scientific Reports*, 7(1), 9364.
- Lipman, B. L. (1986). Cooperation among egoists in prisoners' dilemma and chicken games. *Public Choice*, 51(3), 315–331.
- McLeod, J., Lawler, E., & Schwalbe, M. (2014). *Handbook of the social psychology of inequality*. Dordrecht: Springer.
- Mishra, S., Hing, L. S. S., & Lalumière, M. L. (2015). Inequality and risk-taking. *Evolutionary Psychology*, 13(3), 1–11.
- Nash, J. (1953). Two-person cooperative games. *Econometrica: Journal of the Econometric Society*, 21(1), 128–140.
- Neckerman, K. M., & Torche, F. (2007). Inequality: Causes and consequences. *Annual Review of Sociology*, 33, 335–357.
- Ramsey, F. P. (1931). Truth and probability (1926). In R. Braithwaite (Ed.), *The foundations of mathematics and other logical essays* (pp. 156–198). New York, London: Kegan, Paul, Trench, Trubner & Co., Harcourt, Brace and Company.
- Rehm, P. (2011). Risk inequality and the polarized american electorate. *British Journal of Political Science*, 41(2), 363–387.
- Riker, W. H. (1995). The political psychology of rational choice theory. *Political Psychology*, 16(1), 23–44.
- Riley, W. B., & Chow, K. V. (1992). Asset allocation and individual risk aversion. *Financial Analysts Journal*, 48(6), 32–37.
- Salverda, W., Nolan, B., & Smeeding, T. M. (2009). *The Oxford handbook of economic inequality*. Oxford: Oxford University Press.
- Scott, J. (1999). Rational choice theory. In Understanding Contemporary (Ed.), *Society: Theories of the Present* (pp. 126–138). London: SAGE Publications Ltd.
- Sen, A. (2008). Rational behaviour. In S. N. Durlauf & L. E. Blume (Eds.), *The New Palgrave dictionary of economics*. Basingstoke: Palgrave Macmillan.
- Sheehy-Skeffington, J. & Rea, J. (2017). *How poverty affects people's decision-making processes*. Joseph Rowntree Foundation. <https://www.jrf.org.uk/report/how-poverty-affects-peoples-decision-making-processes>. Accessed 31st January 2018.
- Skyrms, B. (2014). *Evolution of the social contract* (2nd ed.). Cambridge: Cambridge University Press.
- Spoehn, W. (2002). The many facets of the theory of rationality. *Croatian Journal of Philosophy*, 2(6), 249–264.
- Weisberg, M., & Muldoon, R. (2009). Epistemic landscapes and the division of cognitive labor. *Philosophy of Science*, 76(2), 225–252.
- Wood, A. (1995). *North-South trade, employment, and inequality: Changing fortunes in a skill-driven world*. Oxford: Oxford University Press.
- Yee, A. S. (1997). Thick rationality and the missing “brute fact”: The limits of rationalist incorporations of norms and ideas. *The Journal of Politics*, 59(4), 1001–1039.
- Zollman, K. J. (2010). The epistemic benefit of transient diversity. *Erkenntnis*, 72, 17.