Klein, Dominik; Marx, Johannes; Scheller, Simon

Rational Choice and Asymmetric Learning in Iterated Social Interactions : Some Lessons from Agent-Based Modeling

Date of secondary publication: 02.12.2024

Accepted Manuscript (Postprint), Bookpart

Persistent identifier: urn:nbn:de:bvb:473-irb-1051499

Primary publication

Klein, Dominik; Marx, Johannes; Scheller, Simon (2019): Rational Choice and Asymmetric Learning in Iterated Social Interactions : Some Lessons from Agent-Based Modeling, in: Karl Marker, Annette Schmitt, und Jürgen Sirsch (Ed.), Demokratie und Entscheidung : Beiträge zur Analytischen Politischen Theorie, Wiesbaden: Springer VS, pp. 277–294, doi: 10.1007/978-3-658-24529-0_18.

Rational Choice and Asymmetric Learning in Iterated Social Interactions. Some Lessons from Agent-Based Modeling

Dominik Klein, Johannes Marx und Simon Scheller

1 Introduction

In this contribution we analyze how the actions of rational agents feed back on their beliefs. We present two agent-based computer simulations studying complex social interactions in which agents that follow utility maximizing strategies thereby deteriorate their own long-term quality of beliefs. We take these results as a starting point to discuss the complex relationship between rational action couched in terms of maximizing utility and the emergence of informational inequalities. Our discussion of this relation is informed by literature on informational cascades, social epistemology and epistemic game theory.

The effects of information asymmetries are well researched in the contexts of principal-agent problems and contract theory (Akerlof 1970; Spence 1973), in accounting (Chaney and Lewis 1995), and in the study of non-economic behavior, *e.g*, in political science (Banks and Weingast 1992). By *information asymmetry* we mean situations where one party of a transaction has more information about the situation than the other parties. Principal agent theory and the economic theory of contracts offer some insights on the effects differences in informational access may trigger. Asymmetric information may, for instance, translate into power asymmetries between players, leading to inefficiencies in markets (see Kunreuther and Pauly 1985). We will argue that such asymmetries can also play a crucial role in the explanation of the stability of political systems or the emergence of inequality.

Much of the existing literature focuses primarily on the effects of informational asymmetries. In many approaches informational asymmetries are either exogenously given or emerge as an effect of the different roles held by the agents (buyer vs seller or, in contexts of repeated interactions, the informational advantage of certain sellers). In this contribution, we take a different perspective. Analyzing the iterated interactions of equally endowed rational agents, we study informational effects that emerge as a side result of the player's strategic choices. Our focus is thus not so much on effects of informational asymmetries, but on their origins, i.e. how differences in information emerge. It is only in a second step that we study, how such informational feedbacks impact the agents' long-term success. By focusing on the role of strategies for information acquisition as well as information collection as a side effect of strategic action, we add a new perspective to the existing literature on the emergence of informational asymmetries. We

A version of this text appeared as:
Klein D., Marx J., Scheller S. (2019) Rational Choice and Asymmetric Learning in Iterated Social Interactions – Some Lessons from Agent-Based Modeling. In: Marker K., Schmitt A., Sirsch J. (eds) Demokratie und Entscheidung. Springer VS, Wiesbaden

show that different strategies have non-neutral informational consequences that trigger considerable long-term effects. Shortly put: We model the emergence of information asymmetries as an endogenous property of a social system of interacting rational agents.

We start our discussion from Zimmerling's characterization of Rational Choice Theory. She characterizes Rational Choice Theory with the following three principles (1994: 16):

(i) Principle of methodological individualism:
Social change supervenes on individual actions, i.e. all changes on the macro level of social relations result from changes on the micro level constituted by individual actions.

(ii) Principle of choice:
Individual actions are based on individual decisions that satisfy the standards of rationality.

(iii) Principle of rationality:
A choice of action is defined as rational if and only if the action is caused by the desires and beliefs of an agent. More precisely, this action must be the best available alternative in terms of maximizing utility, given the agent's desires and beliefs. Zimmerling further points out that an action can only count as best choice if all its relevant consequences have been considered by the agent.

In the following, we will elaborate on what should be understood by 'given the desires and beliefs of an agent' and 'all relevant consequences'. In a nutshell, we will argue that in certain contexts rationality standards for belief formation can come to conflict with rationality standards for maximizing expected utility. As a consequence, Rational Choice Theory should take the informational underpinnings of rational action more seriously. At least in some social situations the quality of an agent's belief cannot be treated as fixed, but is a relevant consequence of rational actions that should be taken into account by agents when forming rational decisions.

We will structure our argument as follows: First, we report the result of two computer simulations where information asymmetries occur as side-effects of rational actions. We will then relate our findings to the literature on social epistemology, before sketching a broader concept of rationality that takes into account the epistemic dimensions of rational action.

2 Two Agent-Based Simulations on Informational Side Effects

In the following we present the results of two agent-based simulations in which rationality standards for beliefs come into conflict with rationality standards for actions. Our first simulation models the epistemic dynamics preceding the emergence of mass movements.

2.1 Epistemic Sources of Political Revolutions

Our first example of asymmetric learning concerns the prerequisites for political revolutions.[1] In particular, we are interested in revolutionary uprisings in authoritative regimes. Voicing discontent is a dangerous activity for citizens of oppressive regimes. Hence, rational actors will only do so if they see a certain prospect for their protest to trigger significant change. Part and parcel of such considerations is, of course, a public goods problem. Even for those that favor overcoming the status quo, it is an individually dominant strategy to stay silent, thereby leaving the costs and potential risks of protesting to others.

The main focus of our simulation, however, is a step prior to this public goods problem. Before even considering creating the public good[2] of a protest, agents that are, *prima facie*, interested in change need to assess whether there is sufficient discontent in society to make successful protest feasible in principle.[3] Acquiring this information, however, is far from trivial. Due to the oppressive nature of the regime, agents cannot simply start a public dialogue about general political attitudes. To control risk, they might bring up the topic in private conversations only. Even there, doing so is risky, as individuals cannot be sure of the others' motivations. Thus, agents won't raise the topic unless they see the possibility of overall discontent being sufficiently high.

Within the model, each agent is determined by two parameters: their assessment of overall discontent, and their own stand towards the regime. While the first parameter changes as agents receive new information, the latter is fixed: individual agents do not alter their attitude towards the regime in the short time span that is represented by our simulation. They remain critics or supporters throughout. We stipulate, however, that assessing overall discontent is of interest to critics and supporters of the regime alike. For both, the emergence of protest may play a role for deciding what actions to take and what to expect from near

[1] For a detailed description and analysis of the model see Klein and Marx (2018). The computer simulation was implemented in Netlogo (Wilensky 1999). The simulation can be downloaded from the CoMSES Computational Model Library under the url: https://www.comses.net/codebases/5825/releases/1.3.0/.

[2] We use 'public good' in a slightly technical sense here. By calling a protest or regime change a public good, we merely imply that it is non-exclusive, i.e. the effects apply to all citizens alike. We do not make any assumptions on whether individual citizens perceive this public good as desirable or not.

[3] In the classic, rational choice informed view on political revolutions the number of intrinsically motivated agents must be high enough to start the process of regime destabilization (Opp 2013), or at least sufficient selective incentives must be provided (see for example Tullock 1971). Once the process has started, latently motivated agents may observe the viability of protests and join in (Lohmann 1994). Our simulation, however, takes a different route that does not rely on the presence of unconditionally motivated agents willing to start the process, no matter what. Instead, we focus on social situations in which a critical number of discontent agents is present but no public awareness of this fact has yet been established. The simulation models communicational mechanisms that can give rise to public awareness of the possibility of a revolution.

future. Hence, all agents should be interested in learning about attitudes towards the regime, albeit for different reasons.

In the model, agents move randomly at a pre-determined speed on a toroidal grid. Along these movements, they regularly engage in conversations with others they encounter along the way. In each such discussion, one of the two agents is singled out who has to decide on whether or not to bring up the topic of political discontent. If she does so, the opponent reveals her own attitude to the regime.[4] In such an exchange, both sides can learn. Of course, the agent eliciting her counterpart's attitude gains a new piece of information. But also the agent being asked can acquire additional information. By learning *whether* the other raises the topic, she receives an indirect clue about the counterpart's assessment of discontent, reflecting the source's evidential state. Believing that agents only address the topic if they judge it relevant, i.e. if there is a sufficient level of discontent, the fact whether others addresses the topic provides a clue of their assessment. Hence, both agents may receive a piece of information they can use to update their own estimate of overall discontent: They raise this estimate if the other voiced discontent or brought up the topic. They lower their assessment if the other voiced content with the regime or did not raise the topic.

There is a crucial asymmetry in this learning rule: while the responding agent always receives a new piece of information, the topic being addressed or not, the inquirer can only learn if she brings up the topic.[5] This, in turn, happens only if she judges overall discontent to be high enough. In particular the *content* of the agent's belief impacts her information collection and thus the prospect of altering her belief. Agents judging overall discontent to be sufficiently low will not bring up the topic. They will thus not receive any information whenever it is up to them to raise the topic. The only way they can learn is by observing whether others address the topic. As agents only assume the observer role in half of their interactions, the belief that others are supporting the regime can only change slowly. On the other hand, agents that sense a predominantly negative attitude towards the status quo willingly address the topic, thereby collecting new information on the level of public support. This information might, of course, lead them to change their own beliefs. Hence, the judgment that others are discontent will be exposed to a higher frequency of testing through collecting evidence than its regime supporting counterpart.

---

[4] Crucially, we do not want to assume that all agents reveal their *true* attitude towards the regime. For obvious reasons, individuals may prefer to hide their true stance. In assessing overall discontent, agents will need to work with whatever utterances they receive independent of whether the source holds a divergent private opinion. In fact, the model only relies on which attitude agents are willing to utter, not on their true attitude.

[5] For a similar model of the epistemic consequences of trust see Baurmann 2010, p. 210.

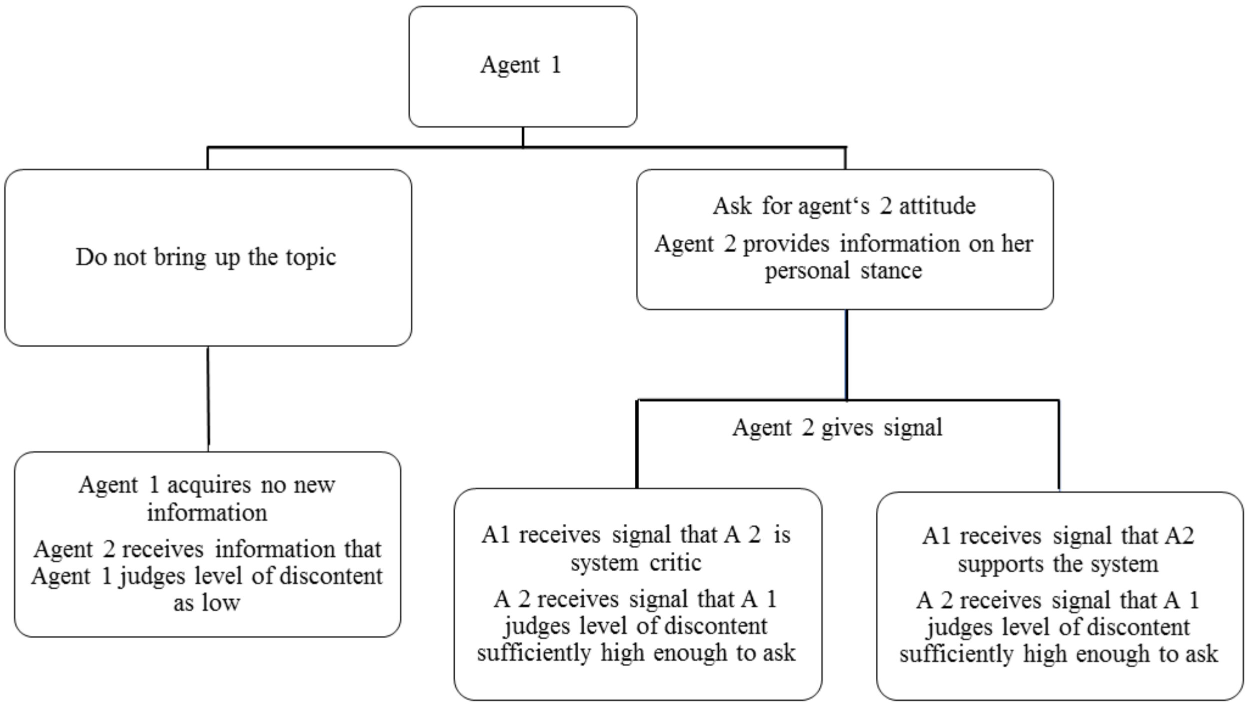


FIGURE 1: POLITICAL REVOLUTIONS AND THE EPISTEMIC STRUCTURE OF THE GAME

There is, moreover, a second factor we should address here: only agents eliciting the attitudes of others collect genuinely new information about individual attitudes. Assessing discontent by observing whether others address this topic, on the other hand, merely leads to updating beliefs on the basis of others' beliefs, without providing any new direct evidence. Hence, once members of a society univocally belief the level of overall discontent to be sufficiently low, no agent ever collects any new information about individuals' real attitudes. Even if a majority of agents happened to be critical towards status quo, such a society could never learn this fact, at least not by the mechanisms studied here. This stands in a stark contrast with a situation where everybody believes overall discontent to be sufficiently high. In such a society, all agents set out to collect new information about the stances of individuals. If it happens to be the case that only few agents are actually dissatisfied by the status quo, this second group will learn so and abandon their overall belief of dissatisfaction being sufficiently high. In a nutshell: False beliefs in a high level of discontent cannot be stable. Whenever the actual level of dissatisfaction, i.e. the share of individuals uttering a preference for change is sufficiently high, society will inevitably learn so, independent of which belief they start with. False belief in support of the regime, on the other hand, can be stably entertained. A society where everybody falsely beliefs overall public to be satisfied with the status quo will never abandon this belief, even if, as a matter of fact, sufficiently many agents are individually discontent with the political situation.

Within our paper we demonstrate that this asymmetry has some unexpected repercussions (Klein and Marx 2018):

- Firstly, we find that a correlation between the agent's own attitude towards the regime and her assessment of overall discontent (see Figure 2).

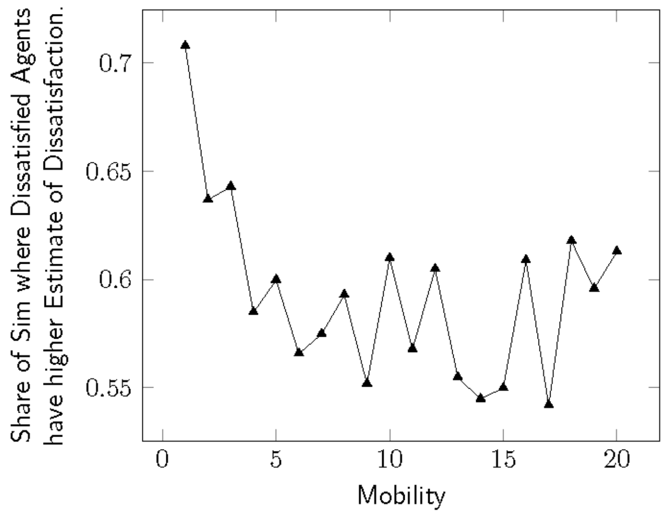


FIGURE 2: SHARE OF SIMULATIONS IN WHICH SYSTEM CRITICS HAVE A HIGHER ASSESSMENT OF GENERAL DISSATISFACTION THAN SYSTEM SUPPORTERS.

  Agents supporting the status quo have, in average, a lower estimate of general discontent than those opposing the regime. Crucially such finding is not brought about by any psychological mechanism driving the agent's belief formation. Within the model, both types of agents operate with exactly the same behavioral and epistemic rules. The difference is exclusively caused by informational feedback effects. Each agent, of course, bears on the beliefs of her surroundings, as her attitude may trigger belief updates of some peers. These updated beliefs of the surroundings may then, later, feed back into the agent's assessment, for instance when these decide on whether to bring up the topic.

- Second, the informational asymmetry described interacts with the agent's mobility in unexpected ways (see Figure 3).

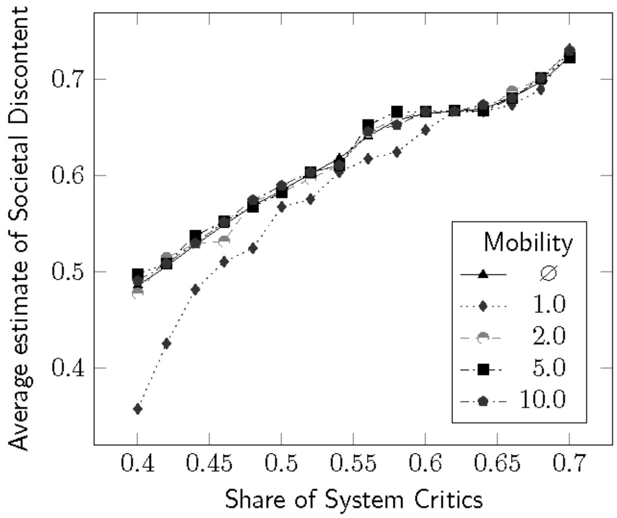


FIGURE 3: AVERAGE ASSESSMENT OF SOCIETAL DISTRUST RELATIVE TO SHARE OF SYSTEM CRITICS AND MOBILITY IN THE SOCIETY

Within our model, all agents move around at a fixed, pre-determined velocity. This velocity, i.e. the distance agents cover each step, stands for their geographical mobility. As it turns out, this mobility has a major influence on the long-term assessment of overall discontent. *Ceteris Paribus*, a lower geographical mobility leads to a significant decrease in likelihood that a majority of agents will judge overall discontent to be sufficiently high (see Figure 3). The driving factor behind this, again, is the asymmetry in the learning rule described, paired with local feedback effects under low levels of mobility. The asymmetry of learning, hence, creates a causal connection (Williamson 2009) between mobility and the perception of discontent, a major condition for successful societal change.

To summarize, by means of an agent-based simulation, we study the informational processes generating those shared attitudes towards a political system that are necessary preconditions for mass movements. We show that societies of lower mobility will structurally underestimate the potential for political change. Moreover, we find that system critics will, over time, develop higher estimates of the potential for change than their less discontent peers. This informational asymmetry is an endogenous result of a social system where information collection is driven by deliberate actions of rational agents.

2.2 Economic Bargaining and Inequality

The second example of asymmetric learning is taken from a simulation on distributive processes in economic interactions. In Klein, Marx, and Scheller (2018), we study iterated bargaining games as a mechanism for settling distributional questions within society. In a highly interdependent society, agents often produce benefits through interaction. Prime examples are an employer hiring an employee for a certain job, or a group of agents combining their respective specializations to produce a complex product. Whenever agents encounter the potential to jointly produce a benefit, they will have a distributional problem to solve: Who is to get how much of the benefit produced. This task may prove especially difficult, as individual contributions might be incommensurable, for instance when agents differ in specialization or assume different roles such as employer and employee. Hence, there often won't be any focal distribution, but agents will need to allocate benefits through bargaining on a case to case basis. In our model, we simulate rational agents' engaging in such distributional processes.

Within our simulation, again agents are paired up in teams of two. Each such team has the opportunity to jointly produce a benefit. However, they will only do so after having agreed on a distribution for the prospective gains. The agents' bargaining is described through an iterated chicken game[6]: In each bargaining round, players either place a high or a modest demand on the potential benefits to be distributed. We assume that two high claims are incompatible. In such a case no agreement is found and bargaining continues. As soon as at least one of the players agrees to a moderate demand, an agreement is found and a production can begin[7]: We assume that players start with a high demand and decide for how many rounds to keep up this demand, and when to reduce their demand to the low amount. We describe the strategies of players through the parameter 'toughness', which states after what time an agent 'gives in'. There is, however, a twist to the bargaining process: Agents overall interaction time is fixed. The longer they spend on bargaining, the less time they have for the actual production of a surplus. Bargaining agents have to balance between bargaining hard to secure a high share of profits and giving in early enough that sufficient benefits can be produced in the first place.

The processes studied here, we argue, stand at the heart of many economic exchange relations, they should be envisaged as a central driving force of a society's wealth distribution. Hence, our study offers a novel perspective for research on the causes of inequality, taking a game theoretic approach, rather than locating the sources of inequality on an individual level.

In our model, we assume that players choose their toughness based on what they know about their opponents' behavior. Within the simulation, agents encounter several collaboration partners: once the maximal interaction time with their current collaborator has expired, agents move on to find a new partner for joint production. Each collaboration begins with a bargaining phase. Agents are hence incentivized to gradually learn about optimal bargaining behavior. They do so via assessing the distributions of toughness in society. Based on their past encounters, agents can form beliefs about the relative frequencies

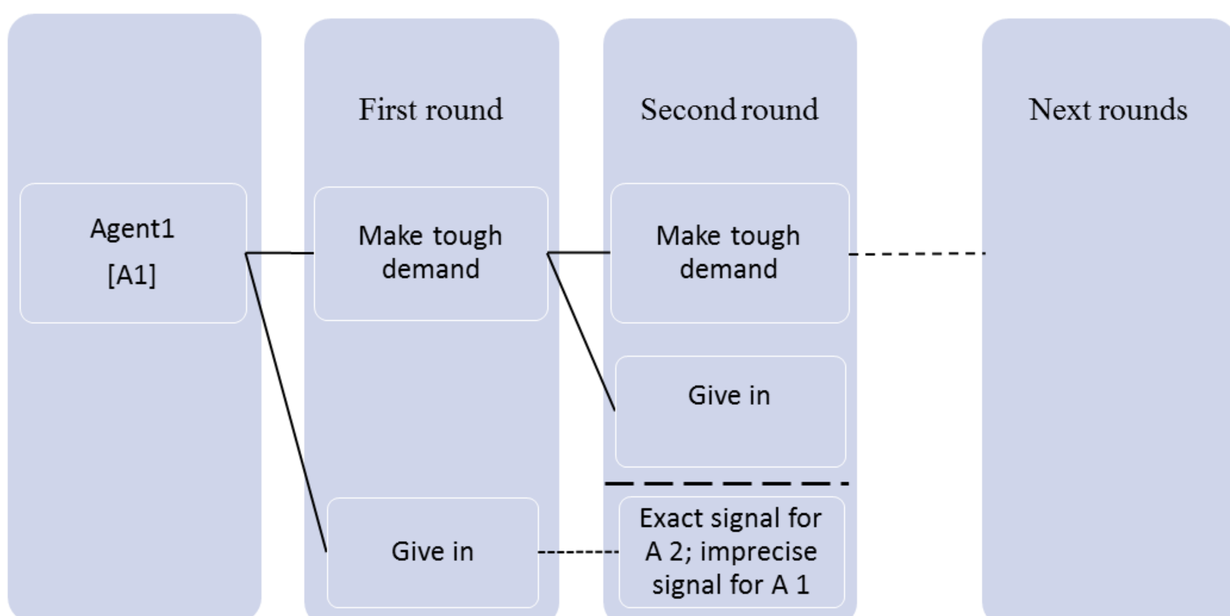
[6] An alternative model of bargaining is provided by Güth and Kliemt 2004. They model bargaining as an ultimatum game. For our purposes the structure of an iterated chicken game is better suited to capture the strategic dimension of the negotiations. We justify this reconstruction of the bargaining process in Klein, Marx, Scheller 2018.

[7] If both players make modest demands, the share of the benefit that remains unclaimed is distributed equally among the two.

of different toughness values among opponents. These beliefs then guide rational agents in selecting a toughness level for their next bargaining interaction.

For the present purpose, let us highlight a central asymmetry of the learning process in iterated bargaining, as figure 4 illustrates: First, note that a player who *wins* a bargaining game receives an exact signal about her opponent's toughness. By upholding a high demand until the other side gives in, such an agent receives precise information after how many rounds her opponent reduces her demand. Second, note that players who give in first receive only incomplete information. The losing player merely learns *that* the opponent's toughness was higher than her own, but not *how long* this opponent would have kept up the high demand. Losing players hence receive structurally less information than winning players.[8]

Moreover, this asymmetry can have structural influences on the content of beliefs. To see this, let us consider a player that sticks with a medium level of toughness. Whenever such a player encounters low toughness opponents, she will win the game and thereby receive exact information about her counterpart's strategies. Against high opponents, on the other hand, this agent will loose and thus only collect imprecise information. In the long run, the agent will hence end up with highly accurate information about low toughness levels, but imprecise information at best about high levels of toughness. This is a structural imbalance in beliefs, emerging as unintended side-effect of her own choices of toughness.



[8] For a detailed description and analysis of the model see Klein, Marx and Scheller 2018. The computer simulation was implemented in Netlogo (Wilensky 1999).

FIGURE 4: ECONOMIC BARGAINING AND THE EPISTEMIC STRUCTURE OF THE GAME

Rational agents, however, will not simply stick to some level of toughness throughout. Rather, they may adapt this level over time, depending on their collected information on the behavior of others. The gradual collection of new information, in combination with the asymmetry just described, will lead to self-enforcing feedback loops. Once her rational choice calculus identifies a low toughness strategy as maximizing expected utility, a player will not collect much information about higher levels of toughness. Hence, she may not update her beliefs on whether higher choices of toughness are beneficial, or become worthwhile when others change their strategies. Tougher players, on the other hand, will receive more precise information, enabling them to constantly optimize their moves. These considerations become especially prominent, when many others change their toughness, changing the calculus of optimization. Receiving precise information about most levels of toughness, tougher players will learn soon about such changes and can adapt accordingly. Lower toughness players, on the other hand, may miss out on subtle changes among tougher strategies, preventing them from optimally adjusting their own strategic choices.

To illustrate the importance of information, learning and the endogenous relation between these two, we compare the success of the different bargaining strategies in a population of different strategy types.[9] These strategy types describe how to select toughness values, based on past experience.

In accordance with a classic notion of rationality, MaxEU agents (simple utility maximizers) always choose the toughness value that maximizes expected utility, given their current beliefs about the distribution of toughness. Contrary to what one might expect, these agents are consistently outperformed by another strategy type, which we call experimenters (see Figure 5). Experimenters follow the EU maximizing principle in 90% of the cases, yet in the remaining 10%, they set their toughness to the maximally possible value. These agents hence sometimes probabilistically decide to not give in early, but to keep up the high demand until the very end. The reason for this, briefly, is information collection. When playing a maximal toughness, these agents receive precise, unbiased information about the behavior of others. Incorporating this information renders the agents' beliefs more adequate, thereby improving performance when choosing for a subjectively EU maximizing toughness. Within all simulations analyzed, experimenters accumulate larger average gains, compared to their pure utility maximizing counterparts. Also, both strategies differ structurally in their long-term beliefs. Experimenters identify larger toughness values as optimal, compared to MaxEU agents. They will hence make different strategic choices, explaining significant differences in payoffs (see Klein, Marx and Scheller 2018).

[9] In particular, now all strategies follow the maximizing expected utility-paradigm in a strict sense.

To show that experimenter's success is directly connected to its informational advantage, we introduce a third strategy type called 'NoLearnEx'. This last type chooses toughness according to the same 90%-10% rule as Experimenter do. Yet, it does not update her beliefs about other agents in the 10% cases where she plays maximal toughness values. That is, NOLearnEx foregoes any information collected in her experimenting phase, leaving her with beliefs that are exclusively based on her EU maximizing moves. This modification produces a drastic drop in average income. Experimenter outperforms NoLearnEx, hence it clearly is the informational advantage that is responsible for the experimenter's high performance. In the controlled setting of our simulation we can thus rule out that experimenter's superior performance is purely due to choosing a better strategy by accident.

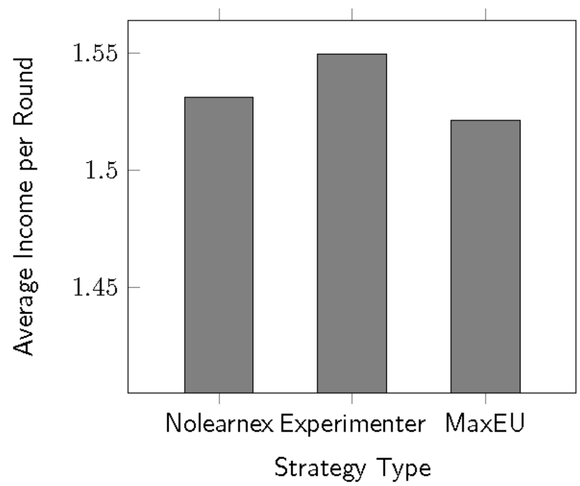


FIGURE 5: AVERAGE INCOME PER ROUND OF DIFFERENT STRATEGIES

Our simulational findings carry far-reaching implications for how we conceptualize rational behavior in iterated bargaining contexts. While classic conceptions of Rational Choice Theory understand one's beliefs as independent from one's choices, the presented example illustrates that they are intricately connected. Being dependent on adequate beliefs, long-term utility maximization must thus take into account how one's action can influence the quality of one's beliefs. As we will argue later in more detail, asymmetric learning complicates this issue beyond a simple choice between explorative and exploitative strategies. In iterated interaction, it's not just the adequacy, but also the unbiasedness of beliefs that is at stake. The beliefs of agents who play comparatively lower toughness strategies are not just less reliable, but are systematically skewed towards making a player believe that small toughness values are more beneficial. Moreover, this bias is produced by an endogenous connection between actions and beliefs. This connection is asymmetric because one type of action (giving in early) produces structurally biased beliefs (towards identifying a low toughness as optimal).

3 How Social Epistemology can Help Towards a Better Understanding of Rationality for Actions and Beliefs

From a classic rational choice perspective, an agent's actions are determined by three main factors: the agent's beliefs, her preferences, and the strategic nature of the underlying decision problem, such as the number of agents involved, their possible moves or how combinations of moves translate into outcomes. The first of these, the agents' beliefs, usually receives little attention. Often, it is simply assumed that agents' have collected sufficient evidence and entertain reasonable beliefs. However, both case studies presented here suggest that this perspective might be missing out on crucial factors. In the following, we will discuss some relevant aspects of belief formation in social contexts.

Often, as in the first case study on informational dynamics preceding revolutions, it might be difficult for agents to collect sufficient direct evidence. In such cases, agents may attempt to learn from the beliefs and experience of others. Such social learning, while justified in many contexts, has a number of potential pitfalls. As the first case study illustrates, social learning can lead to informational cascades when applied by many parties in parallel: self-reinforcing belief formation phenomena that are based on echo-chamber effects, without importing new factual evidence. While social learning might be an epistemically reliable strategy for the individual, it can lead to detrimental effects if applied by a large group in parallel. Hence, an adequate analysis in the social sciences cannot stop at the individual level, but needs to represent a multitude of agents in context.

Belief formation is classically assumed to be independent of the agent's current beliefs or actions. Even when we are reasoning about when and whether agents invest in collecting new evidence, this should be independent of her strategic actions or the *content* of her current belief. Both our case studies indicate that this picture is inadequate for many cases of long-term iterated involvement. In our first case study, informational dynamics preceding revolutionary uprisings, the amount of evidence collected, depended on the agent's information seeking which, in turn, was directly influenced by the *content* of her belief. Hence, the agent's current beliefs impact evidence collection in an epistemically non-neutral way. In fact, this dependence creates an asymmetry between two possible beliefs states corresponding to high and low levels of overall dissatisfaction. This asymmetry, in turn, gives rise to novel dependencies, for instance between social mobility and starting conditions for social unrest.

In our second example, the emergence of distributional inequality, belief formation was impacted by the agent's actions.[10] Here again, the different actions not only impacted the amount of new information collected over time, but also its content. For certain strategies, this new information turned out structurally skewed, hence gradually deteriorating their players' long-term adequacy of beliefs. Expected utility strategies, however, rely on having adequate beliefs. Undermining the quality of belief hence translated into a poor performance of these strategies, leading, in turn, to the emergence of economic inequality. Here again, it is a non-neutral dependence of information collection on the agent's actions that forms the basis of relevant societal patterns, the emergence of inequality.

To sum up, our case studies serve to illustrate that epistemic considerations, pertaining to both, agent's beliefs and information collection, play a crucial role for social processes and deserve explicit attention. We have shown that natural and rational approaches to belief formation *i)* cannot be separated from the agent's current belief or actions, *ii)* need to be studied on a societal level rather than focusing on single agents, and *iii)* can impact societal patterns such as emergent inequality or a causal relation between mobility and the propensity for societal uprisings. In the next section we will discuss the consequences of these findings for our concept of Rational Choice Theory.

4 Exploration-Exploitation-Trade-Off, Asymmetric Learning and Rationality

In economics, but also in broad parts of the social sciences, the concept of rationality is employed in a narrow sense and understood as the ability to adopt the best actions to achieve given goals. This understanding has, however, been challenged by various authors (see for example Elster 1988, Davidson 1980, Marx and Tiefensee 2015). Elster (1988), for example, argues that it is insufficient to understand Rational Choice Theory merely as a theory about the rational selection of actions. Instead he argues that Rational Choice Theory should also provide standards of rationality for the beliefs and desires held by an agent. Elster identifies three places in a theory of rational choice where rationality standards come into play:

1. The set of beliefs an agent holds must be internally and externally consistent (rationality standard for beliefs).
2. The set of desires an agent holds must be internally consistent (rationality standard for desires).

[10] Of course, these actions are, at least in parts, influenced by beliefs again.

3. An agent's beliefs and desires must cause an action in the right way. Also, this action must be the best choice given the individual desires and beliefs (rationality standard for choice of action). Rational Choice Theory formalizes this last standard of rationality in an axiomatic way (see for example von Neumann and Morgenstern 1953).

Characterizing an agent as irrational is thus equivalent to the proposition that she fails on one or more of conditions (1), (2), and (3). Much of Rational Choice Theory concentrates on the third condition exclusively, taking the first two conditions as independent of (3) and given. In particular, Rational Choice Theory conceptualizes information and its acquisition as independent from one's action.

The results from the two simulation experiments reported above indicate that this picture may be in parts wrong and that there is need for a broader concept of rationality. As illustrated by the simulation results, some classic assumptions, in particular about beliefs being independent from actions, are shortsighted and lead to a neglect of systematic feedback effects. Our results show that the different standards of rationality can be intertwined in a complex manner and, *a fortiori,* that they may conflict with each other.

A well-known connection between information and action that has been amply studied is the classic exploration-exploitation divide.[11] Briefly this divide describes the necessary trade-off between exploring one's environment to identify new strategies and assess their prospects, and to exploit these findings for the production of gains. The question of how to optimally allocate one's resources between exploration and exploitation is frequently discussed yet remained without a general answer so far.

Classically, the Exploitation-Exploration-trade-off conceptualizes the exploration of the strategy-space as a risky endeavor, as both potential strategies and their outcomes are unknown. The selection between exploration and exploitation may hence be difficult to assess with the classic instruments of rational choice. Yet it is often implicitly assumed that risks and benefits are randomly distributed, rendering exploration a stochastic optimization problem. Often, one can imagine (as in Weisberg and Muldoon's model) exploration as the search in an unexplored territory of a landscape.

In our two simulations on political revolutions and bargaining, the case is arguably more complex than that. In those cases, exploration (or non-exploration) is strategically non-neutral in two very distinct ways: On the one hand, (1) there is a connection between strategic actions and information collection. Exploration is not a move in itself. Rather, it is different strategic actions within social interaction that provide varying degrees of information as side effect. On the other hand, (2) every choice of action

[11] For an overview in the behavioral and brain sciences we refer to Cohen, McClure, and Yu (2007), in philosophy see Weisberg and Muldoon (2009) and Zollman (2010).

influences the underlying system as it is observed by other agents who also adapt their strategic decisions to the system's behavior.

Both sources of non-neutrality hence stem from the fact that in our case studies, learning takes place in a game-theoretic setting. Every strategic action in such a setting affects (and, in combination with all other actors, even makes up in the long-run) the environment for future decision. Let us briefly relate the discussion of non-neutrality to our case studies:

> Ad (1): In bargaining, gathering more information about the behavior of others is tantamount to aggressive bargaining behavior. This is so because those who give in early in our endurance-competitions will never know when an opponent would have given in. Vice versa, those who hold out long will, sooner or later, find out when the opponent gives in, thereby receiving precise information about her toughness. The aggressive behavior of information collection, however, has far-reaching consequences well beyond its informational implications: In fact, it changes the very nature of the iterated bargaining game agents are involved in. To illustrate this point with a simple example, imagine a game of poker. When some opponent makes a high bet, there is only one way to find out about her cards and whether or not she is bluffing: you need to make a large bet yourself, i.e. you need to play aggressively. In fact, the actual point of making large bets often is to make the acquisition of information more costly for one's opponents. In both cases, it is the aggressive players who know more, hence creating a tight connection between aggression and exploration. This connection, in particular renders exploration a non-neutral strategy.

> Ad (2): As illustrated by the simulation on political revolutions, one often may not be able to acquire information about a system without changing the very thing one wants to assess. To see this, recall that agents employ the fact of whether others address the topic of political attitudes as an indicator on whether general dissatisfaction is high or low. For agents that expect low level of dissatisfaction this is, in fact, the only access to new information. However, if being asked, agents will have to reveal their own attitudes (truthfully or not). By doing so, they may, of course, impact their sources assessment of overall discontent. Uttering political dissatisfaction will fuel the source's perception of discontent. Reaffirming the status quo, on the other hand, will lead the source to lower her estimate. Changing the perception of its surroundings, this updates assessment might feed back to the original agent, explaining why dissatisfied agents in our simulation have a higher estimate of the overall societal dissatisfaction with the political system.

5 Conclusion

With this contribution we showed that various social settings require a concept of rationality that is broader than current standard approaches. By means of two agent-based simulations, we illustrated that the standard notion of rationality as maximizing expected utility is too narrow for a class of social situations that is characterized by iterated interaction, social feedback loops and an intricate connection between actions and belief formation. Instead we argued for a substantial account of rationality that keeps in mind both, the needs for optimal actions as well as the acquisition of new information. In this broadened perspective, the quality of future beliefs should be considered as a relevant consequence of rational actions.

Literature

Akerlof, George A. (1970): “The Market for 'Lemons': Quality Uncertainty and the Market Mechanism”. *Quarterly Journal of Economics*. The MIT Press. Vol. 84: 3, pp. 488-500.

Banks, Jeffrey S.; Weingast, Barry R. (1992): “The political control of bureaucracies under asymmetric information”. *American Journal of Political Science*. Vol. 36: 2, pp. 509-524.

Baurmann, Michael (2010): “Kollektives Wissen und epistemisches Vertrauen: der Ansatz der Sozialen Erkenntnistheorie“, in: Gert Albert/Steffen Sigmund (eds.), Soziologische Theorie kontrovers. Sonderheft 50 der Kölner Zeitschrift für Soziologie und Sozialpsychologie, pp. 185-201.

Chaney, Paul K.; Lewis, Craig M. (1995): “Earnings management and firm valuation under asymmetric information”. *Journal of Corporate Finance*. Vol. 1: 3-4, pp. 319-345.

Cohen, Jonathan D; McClure, Samuel M.; Yu, Angela J. (2007): “Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration”. *Philosophical Transactions of the Royal Society B: Biological Sciences*. Vol. 362: 1481, pp. 933-942.

Davidson, Donald (1980). Actions, Reasons and Causes. In: idem: Essays on Actions and Events, pp. 3-20. Oxford: Clarendon Press.

Klein, Dominik; Marx, Johannes (2017): “Wenn Du gehst, geh ich auch! Die Rolle von Informationskaskaden bei der Entstehung von Massenbewegungen”. *Politische Vierteljahresschrift, PVS*. Vol. 58: 4, pp. 560-592.

Güth, Werner; Kliemt, Hartmut (2004): “Perfect or Bounded Rationality? Some Facts, Speculations and Proposals“. *Analyse & Kritik.* Vol. 26: 2, pp. 364-381.

Klein, Dominik; Marx, Johannes (2018). “The Informational Dynamics of Regime Change” (Version 1.3.0). CoMSES Computational Model Library. NetLogo Code. Retrieved from: https://www.comses.net/codebases/5825/releases/1.3.0/

Klein, Dominik; Marx, Johannes; Scheller, Simon (2018): “Rationality in Context. On Inequality and the Epistemic Problems of Maximizing Expected Utility”, *Synthese*, forthcoming.

Kunreuther, Howard; Pauly, Mark (1985): “Market Equilibrium with Private Knowledge”. In: Dionne G., Harrington S.E. (eds) Foundations of Insurance Economics. Huebner International Series on Risk, Insurance and Economic Security, vol 14. Springer, Dordrecht, pp. 424-443.

Lohmann, Susanne (1994): The dynamics of informational cascades: the Monday demonstrations in Leipzig, East Germany, 1989–91. World politics. Vol. 47:1, pp. 42-101.

Neumann, John von; Morgenstern, Oskar (1953): Theory of Games and Economic Behavior. Princeton, NJ. Princeton University Press.

Opp, Karl Dieter (2013): Die enttäuschten Revolutionäre: politisches Engagement vor und nach der Wende. Springer-Verlag.

Spence, Michael (1973): “Job Market Signaling”. *Quarterly Journal of Economics*. Vol. 84: 3, pp. 355-374.

Tullock, Gordon (1971): “The Paradox of Revolution”. *Public Choice*. Vol. 11 (Fall, 1971), pp. 89-99.

Weisberg, Michael; Muldoon, Ryan (2009): “Epistemic landscapes and the division of cognitive labor”. *Philosophy of science*. Vol. 76.2, pp. 225-252.

Wilensky, Uri. (1999): NetLogo. Center for Connected Learning and Computer-Based Modeling: Northwestern University, Evanston, IL. Download link: http://ccl.northwestern.edu/netlogo/ [last accessed: 28.02.2018].

Williamson, Jon (2009): “Probabilistic theories of causality”. In: Beebee H., Hitchcock C. (eds) The Oxford handbook of causation, Oxford University Press, pp. 185-212.

Zimmerling, Ruth (1994): “Rational Choice“ Theorien: Fluch oder Segen für die Politikwissenschaft?”. In: Druwe, U./Kunz, V. (eds), Rational Choice in der Politikwissenschaft. Opladen, pp. 14-25.

Zollman, Kevin JS. (2010): “The Epistemic Benefit of Transient Diversity”. *Erkenntnis* 72:1, pp. 17-35.