IMT School for Advanced Studies, Lucca
Lucca, Italy

Essays on the Evolution of Prosocial Behaviors

PhD Program in Systems Science
Track in Economics, Networks and Business Analytics
XXXIV Cycle

By

Sedric Zucchiatti

2023

The dissertation of Sedric Zucchiatti is approved.

PhD Program Coordinator: Prof. Pietro Pietrini, IMT School for Advanced Studies Lucca

Advisor: Prof. Ennio Bilancini, IMT School for Advanced Studies Lucca

Co-Advisor: Prof. Leonardo Boncinelli, University of Florence

The dissertation of Sedric Zucchiatti has been reviewed by:

Daniele Giachini, Scuola Superiore Sant’Anna

Edgar J. Sanchez Carrera, University of Urbino Carlo Bo

IMT School for Advanced Studies Lucca
2023

To my family, for the constant support
To the friends who shared this journey with me
To those who are still finding their way

Contents

List of Figures

List of Tables

Acknowledgements

I'm so grateful to my advisor, Ennio Bilancini, and my co-advisor, Leonardo Boncinelli, for the guidance they have provided me during these years. I give them credit for the huge contribution they made as coauthors of the papers ″The Structure of Interaction and Modes of Reasoning Can Shape the Evolution of Conventions″ (Chapter 2) and ″The Co-Evolution of Cooperation and Cognition Under Endogenous Population Structure″ (Chapter 3).

I would like to thank the PRO.CO.P.E project within which the paper ″The Co-Evolution of Cooperation and Cognition Under Endogenous Population Structure″ (Chapter 3) has been conceived.

I would also like to thank the reviewers of this thesis – Daniele Giachini and Edgar J. Sanchez Carrera – for their precious comments and suggestions.

I also owe much to everybody else who spent time discussing my works with me, who suggested possible ways to improve them, or who simply listened to my complaints when things were not working out. I can't thank you enough.

Vita

October 17, 1994	Born, Udine (UD), Italy
2016	Bachelor's Degree in Economics Final mark: 110/110 cum laude Università degli Studi di Udine, Udine, Italy
2018	Master's Degree in Economics Final mark: 110/110 cum laude Università degli Studi di Udine, Udine, Italy

Publications

Presentations

1. Bilancini, Boncinelli, Zucchiatti, “The Co-Evolution of Cooperation and Cognition Under Endogenous Population Structure”, at *FIRST INTERIM PRO.CO.P.E. WORKSHOP*, Online, 2020.
2. Bilancini, Boncinelli, Zucchiatti, “The Structure of Interactions and Modes of Reasoning Can Shape the Evolution of Conventions”, at *61ma RSA della Società Italiana degli Economisti*, Online, 2020.
3. Bilancini, Boncinelli, Zucchiatti, “The Structure of Interactions and Modes of Reasoning Can Shape the Evolution of Conventions”, at *GAMES2020*, Budapest, Hungary, 2021.
4. Bilancini, Boncinelli, Zucchiatti, “The Structure of Interactions and Modes of Reasoning Can Shape the Evolution of Conventions”, at *SASCA PHD CONFERENCE 2021*, Sassari, Italy, 2021.
5. Bilancini, Boncinelli, Zucchiatti, “The Co-Evolution of Cooperation and Cognition Under Endogenous Population Structure”, at *LEG CONFERENCE 2022*, Lucca, Italy, 2022.

Abstract

Prosocial behaviors – such as helping others, donating, and cooperating – are often considered key to evolutionary success. Therefore, it is of great interest to understand under what conditions these behaviors can emerge and/or can be sustained at a population level. Following a dual process approach, I study whether and how cognition can affect the evolution of collaboration, cooperation, and generosity. I do this by employing stochastic stability analysis techniques and agent-based simulations. For each prosocial behavior considered, I find that cognition can play an important role in the diffusion of prosocial behaviors, sometimes fostering them and other times hampering them. These results shed light on recent experimental evidence and, at the same time, suggest new interesting research avenues.

Chapter 1

Introduction

This thesis is a collection of three papers studying the evolution of prosocial behaviors following a dual process perspective.

The term 'prosocial behaviors' covers a wide range of behaviors whose common feature is that they provide a benefit to others. Helping others, cooperating, collaborating, and donating are typical examples of prosocial behaviors. These behaviors are extensively studied in many fields such as economics, psychology, and biology.

From an economic perspective, prosocial behaviors are interesting for many reasons. In fact, given that performing prosocial behaviors often (although not necessarily) comes at a cost one might wonder why individuals engage in these behaviors in the first place: Is it because of individual preferences? If so, are these preferences by nature or by nurture? Or do individuals perform prosocial behaviors because of other external mechanisms such as repeated interactions, kinship, reciprocation, or even fear of punishment? Moreover, given that prosocial behaviors benefit others, if they come at no cost or at a cost lower than the benefit they provide then they are welfare-enhancing. With this perspective in mind, it is of particular interest to understand whether and how prosocial behaviors can be promoted, but also how to eliminate disruptive elements.

In each paper contained in this thesis, I develop and analyze models of the evolution of prosocial behaviors. Even though each model focuses on a distinct prosocial behavior and tries to answer different open questions in the relevant literature, they all aim at providing explanations of why prosocial behaviors may emerge even in societies in which individuals are self-interested, i.e. agents are always assumed to care merely about their own payoff. In addition, all models share a dual process approach to cognition and the use of evolutionary techniques.

In this thesis, the term cognition refers specifically to the process by which an agent chooses the amount and the type of information to exploit when revising (or implementing) its strategy. Moreover, all the papers adopt a dual process approach to cognition. According to dual process theories of cognition, human decision-making can be seen as the result of the interplay between two modes of reasoning: intuition, which is quick and relies on heuristics, and deliberation, which is slow and requires careful scrutiny of available information, costs, and benefits (Evans and Stanovich, 2013). Following recent applications of dual process theories of cognition (Bear and David G Rand, 2016; Jagau and Veelen, 2017), I model deliberative thinking as an individual's ability to obtain more accurate information. More precisely, given that in each model considered individuals engage in different strategic interactions, in the setups considered obtaining more accurate information via deliberation translates into the ability of an agent to recognize the specific type of interaction it is currently engaged in. This in turn allows the agent to behave differently in the various types of interactions (of course if this is beneficial). Even though the effect of deliberating is the same in all the models developed, the models differ in the way they assume agents incur into deliberation. On the one hand, the model in Chapter 2 assumes exogenous deliberation and, so, an agent either always or never deliberates depending on its type. On the other hand, the models in Chapter 3 and Chapter 4 assume that each individual chooses its cognitive mode after a cost-benefit evaluation and, so, in these models cognition is endogenous.

Evolutionary selection is studied within a population games approach. In fact, all the papers feature a population (or multiple populations) of agents that interact over time. More precisely, in every period of time, agents are randomly matched into couples – in case of an endogenous interaction structure, random matching may be restricted to sub-samples of the entire population – and each couple plays a stage game that is meant to capture the tension between adopting a prosocial behavior (cooperating, collaborating, being fair) and not doing so. After each interaction, agents revise their strategy with positive probability following a certain revision protocol (myopic best reply, imitation, reinforcement learning). After agents have revised their strategy (if they had the chance to do so), a new time period begins.

These kinds of setups are often characterized by inertia, i.e., agents are not guaranteed to revise their strategy in each period of time, and myopia, i.e., when revising their strategy, agents consider only current information about the system or about payoffs, but they do not account for possible future trajectories of the system. Under these conditions, the population games at hand behave as Markov chains, and consequently, they are usually studied via either stochastic processes or agent-based simulation techniques.

The population games considered present multiple equilibria, which significantly reduce the possibility of deriving clear predictions from the model. When this is the case, I employ stochastic stability analysis. Stochastic stability analysis (Foster and Young, 1990; H. P. Young, 1993; Kandori, Mailath, and Rob, 1993; Freidlin and Wentzell, 1984) is an equilibrium selection device. In short, under the assumption that agents make mistakes with a small probability ε, stochastic stability analysis selects the equilibria that are relatively easier to reach in terms of mistakes. In proving sufficient conditions for stochastically stable states, I often exploit the radius-coradius theorems in Ellison (2000).

When employing stochastic stability analysis, the population game is studied under two different dynamics, which are usually referred to as unperturbed and perturbed dynamics.

Under unperturbed dynamics, agents revise their strategy according to their revision protocol without making mistakes. Under these dynamics, the goal is to characterize all the recurrent classes or absorbing sets. An absorbing set of the system is a set of states such that if the system reaches one of the states belonging to the absorbing set, then (under unperturbed dynamics) the system will remain indefinitely in states belonging to the absorbing set.

Under perturbed dynamics, instead, agents revise their strategy according to their revision protocol most of the time, but they also make mistakes with positive probability ε. Under these dynamics, there are no absorbing sets, as the system may leave every set of states with positive probability (eventually via mistakes). Moreover, starting from any state, σ, the system can reach with positive probability any other state, so the population game constitutes an ergodic Markov chain. But then, it has a unique invariant distribution, denoted by μ^{ε}, which can be interpreted as a probability distribution over the state space Σ such that, for each $\sigma \in \Sigma$, the probability of being in state σ approximates the fraction of time spent by the system in state σ in the long run. This probability distribution exists and is unique for any given $\varepsilon > 0$; however, as it is customary in the literature, we focus on the limit distribution μ^* for $\varepsilon \to 0$.

The limit distribution μ^* corresponds to the limit of the invariant distribution as mutations become arbitrarily small, and it is an approximation of the invariant distribution for sufficiently small values of the mutation rate ε. States to which the limit distribution associates positive probabilities are said to be stochastically stable. Such states must belong to some absorbing set of the system in the unperturbed dynamics (H. P. Young, 1993).

The remaining part of the thesis is structured as follows: Chapter 2, Chapter 3, and Chapter 4 contain the three papers while Chapter 5 concludes.

The first paper, presented in Chapter 2, is a joint work with Ennio Bilancini and Leonardo Boncinelli. In the paper, we study the evolution of conventions in the Stag Hunt Game. In particular, we are interested in analyzing under which conditions the payoff dominant convention – in which players choose risky but socially optimal collaboration – is the expected evolutionary outcome. In this regard, previous studies have shown that the structure of interaction can foster the adoption of the payoff dominant convention (Oechssler, 1997; Anwar, 2002; Bhaskar and

Vega-Redondo, 2004; Pin, E. Weidenholzer, and S. Weidenholzer, 2017); moreover, recent experimental evidence (Belloc et al., 2019; Bilancini, Boncinelli, and Paolo, 2021) has found that cognition may play a role too. By combining these considerations, we consider an evolutionary game where: (i) agents choose a location to interact locally, (ii) interactions are sometimes global and sometimes local, and (iii) agents can be either fine or coarse reasoners, i.e., agents are able or not, respectively, to distinguish between global and local interactions. We show that the structure of interaction and the mode of reasoning affect the selection of social conventions. Further, we find that the coexistence of coarse and fine reasoning may favor or hinder the adoption of the payoff dominant action depending on the structure of interactions. In particular, if interactions are mostly local, then fine reasoning increases the diffusion of the payoff dominant action. Instead, if interactions are sufficiently global, then fine reasoners are never more collaborative than coarse reasoners and they may even disrupt the emergence of payoff dominant conventions.

The second paper, reported in Chapter 3, is also a joint work with Ennio Bilancini and Leonardo Boncinelli. In this paper, we study the co-evolution of cooperation and cognition with an endogenous interaction structure. We built upon an evolutionary model (Bear and David G Rand, 2016; Jagau and Veelen, 2017) in which agents interact over time and are randomly matched to play either an anonymous one-shot or an infinitely repeated prisoners' dilemma. However, agents can recognize the actual game they are playing only under deliberation which comes at a cost randomly sampled from a generic distribution. We introduce a set of identical locations and impose that only agents in the same location interact with each other (endogenous structure of interaction). We show that depending on the actual distribution of deliberation costs the system is characterized by either two or three types of absorbing sets: (i) an intuitive defection set, (ii) dual process cooperation states, and eventually (iii) dual process defection states. Moreover, we find that the presence of an endogenous interaction structure enlarges the parameter space in which dual process cooperation states are stochastically stable and, in particular, we show that if dual process cooperation states are absorbing then they are also stochastically stable.

The third paper is reported in Chapter 4. In it, I develop a model for the evolution of fair splits and harsh rejections in the Ultimatum Game. The model features two types of reinforcement learning agents, namely proposers and receivers, who are randomly matched to play either an Ultimatum Game or a simplified Bargaining Game. Agents know the relative frequency of each game, but they are unable to distinguish whether they are playing one game or another in a specific interaction unless they incur a deliberation cost randomly sampled from a known (uniform) distribution. I find that depending on the cost distribution considered, proposers' may evolve to follow different kinds of strategies while receivers' equilibrium behavior is independent of the cost distribution considered. These results can account for the emergence of fair offerings and frequent rejections in the Ultimatum Game. Further, the model provides several predictions regarding the effects of cognitive manipulations on proposers' and receivers' behaviors and underlines a potential limitation in the effectiveness of cognitive manipulations – namely, endogenous receivers' deliberation probabilities.

Chapter 2

The Structure of Interaction and Modes of Reasoning Can Shape the Evolution of Conventions

2.1 Introduction

We study how social conventions are shaped in the long run by the frequency of local social interactions (as opposed to global ones) when agents are mobile across the locations where local interactions take place (as in Oechssler, 1997; Ely, 2002; Bhaskar and Vega-Redondo, 2004) and the population consists of two types of myopic best responders: coarse reasoners (who cannot distinguish between local and global interactions) and fine reasoners (who can condition their actions on whether their current interaction is local or global).

Interest in this setting stems from a desire to understand whether and how the evolution of conventions may be affected by the heterogeneity of the mode of reasoning in the population when the mode of reasoning allows distinguishing between local and global interactions. While the degree of locality of interactions has been shown to play a role in the selection of conventions (see Newton, 2018, and references therein), its interplay with the mode of reasoning has not yet been investigated. In particular, taking into account parallel work on the evolution of cooperation (Jagau and Veelen, 2017), one can reasonably expect that the interplay between the degree of locality and the mode of reasoning leads to non-trivial results in terms of convention selection for both global and local interactions.

We focus on the evolution of conventions in the Stag Hunt game. The Stag Hunt game is often viewed as a paradigmatic representation of the obstacles to social cooperation (Skyrms, 2004). In fact, in the game, two players must simultaneously decide whether they want to collaborate with the other agent (hunt a Stag) or work on their own (hunt a Hare). Social cooperation can either provide the highest attainable payoff (if the other player cooperates too) or the lowest possible one (if the other agent does not cooperate), while individualistic behavior provides a higher expected payoff if the opponent's behavior is uncertain. It is of particular interest to understand under which conditions the payoff dominant convention is selected and, thus, social cooperation may thrive as coordination on the payoff dominant convention is socially optimal.

The structure of interaction is modeled as a mixture of random encounters in a chosen location and random encounters in the whole population. This is a convex combination of two interaction structures that have been extensively explored in the literature on the evolution of conventions. Under random encounters in the whole population (Kandori, Mailath, and Rob, 1993; Kandori and Rob, 1995; H. P. Young, 1993), the risk dominant convention tends to be established, instead under random encounters in a chosen location (Oechssler, 1997; Anwar, 2002; Bhaskar and Vega-Redondo, 2004; Pin, E. Weidenholzer, and S. Weidenholzer, 2017) there is room for the evolution of the payoff dominant convention. Under certain circumstances – such as limited location capacity or impossibility of re-optimizing together location choice and play – the payoff dominant and the risk dominant convention may co-exist. The literature has further investigated other structures of interactions that do not belong to the convex combination we consider. There are models where interaction occurs with exogenously given neighbors – as in Ellison (1993), where players are arranged on a circle and interact with the two immediate neighbors.[1] In these models, the risk dominant convention is selected in the long run if agents adopt a best response rule when updating behavior (Peski, 2010).[2] The literature has also considered models with endogenous network formation, where agents choose with whom to interact (Goyal and Vega-Redondo, 2005; Jackson and Watts, 2002; Staudigl and S. Weidenholzer, 2014; Bilancini and Boncinelli, 2018). In these models, the payoff dominant convention is shown to emerge in the long run if the single interaction is sufficiently costly or the total number of interactions per agent is sufficiently constrained. Our setting with local and global interactions combines some features characterizing models with exogenous network formation with features peculiar to models with endogenous network formation. In fact, under global interactions, each agent is connected with every other player in the society, and no agent can modify its pool of potential partners like in an exogenous fully-connected network. Instead, under local interactions, each agent interacts only with the other players staying in the same location, and it can partially decide with whom to interact by choosing one or another location. This feature of local interactions makes them resemble endogenous interaction structures despite having agents limited to choose between groups of agents (the ones staying in a given location) with whom to interact rather than between individual agents with whom to interact as in endogenous networks.

We assume that all agents in our model follow myopic best response, a widely adopted behavioral rule in evolutionary models (Newton, 2018). Behavioral rules other than myopic best response have also been considered in the literature. A prominent rule is imitation, which typically fa-

---

[1] See S. Weidenholzer (2010) for a survey on the evolution of social coordination under local interactions.

[2] If the behavioral rule is instead imitation, then the payoff dominant convention can emerge if interactions are neither global nor limited to the immediate neighbors (Alós-Ferrer and S. Weidenholzer, 2006; Alós-Ferrer and S. Weidenholzer, 2008) or if information transmission about average earned payoff is costly and agents have many neighbors (Cui, 2014). Long-run behaviors depend on the specification of the imitation rule and the communication structure (Chen, Chow, and Wu, 2013).

vors the long-run selection of the payoff dominant convention (Robson and Vega-Redondo, 1996; Alós-Ferrer and S. Weidenholzer, 2008).

Heterogeneity in the mode of reasoning is modeled by introducing two different types of myopic best responders (coarse reasoners and fine reasoners) who differ in their ability to condition their best response to information about the type of interaction. Specifically, coarse reasoners myopically choose a single best reply to the distribution of behaviors (over both local and global interactions), while fine reasoners myopically choose two best replies: one for local interactions and one for global interactions. Basically, coarse reasoners cannot distinguish if an interaction is local or global, although they are aware of the likelihood that it is one or the other. In other words, we model finer reasoning as the ability to obtain and process more accurate information about the features of the interaction, which translates into the possibility for fine reasoners to condition their action on the type of interaction (this is reminiscent of, e.g., Bear and David G Rand, 2016; Jagau and Veelen, 2017). This setup can be seen as a simplified version of dual process theories of reasoning where two modes of reasoning are possible for decision-making: intuition, which is quick and relies on heuristics, and deliberation, which is slow and requires careful scrutiny of available information, costs, and benefits (Evans and Stanovich, 2013). Our simplification is that an agent's reasoning mode is exogenous, i.e., it cannot switch from being coarse to being fine reasoner or vice versa. This is mainly done for the purpose of keeping the model tractable. In Section 2.5, we discuss the expected consequences of having an endogenous reasoning mode, which is also an interesting avenue for further research. Alternatively, the heterogeneity in the mode of reasoning characterizing our setup can be interpreted as the effect associated with the adoption of different heuristics by coarse and fine reasoners (Gigerenzer and Gaissmaier, 2011). Following this alternative approach, both types of agents adopt a heuristic as they myopically best reply and, so, they optimize their behavior without taking into account past information (only the current state of the system matters) and without considering potential strategy updates from other agents. However, coarse and fine reasoners adopt different heuristics as the former further discard information regarding the degree of locality of interactions.

Evolutionary selection is studied both by means of stochastic stability analysis and by performing agent-based simulations of the model. In proving sufficient conditions for stochastically stable states, we crucially exploit the radius-coradius theorems in Ellison (2000). We then provide simulation results indicating that the conclusions we obtain by performing stochastic stability analysis on sub-regions of the parameter space can actually be extended over the entire parameter space.

We find that different conventions can be selected in the long run depending on the probability of random local interactions and the fraction of fine reasoners in the population. More precisely, if the probability of local interactions is below a certain threshold, then independently of the fraction of fine reasoners in the population selection favors a semi-risk dominant convention in which coarse reasoners play the risk dominant convention while fine reasoners play the payoff dominant convention locally and the risk dominant one globally; moreover, agents are separated in distinct locations according to type. Instead, if the probability of local interactions is higher than that threshold, then the selected convention depends on the fraction of fine reasoners in the population. In particular, if the fraction of fine reasoners is higher than a given threshold, then selection favors a semi-payoff dominant convention in which coarse reasoners play the payoff dominant convention while fine reasoners play the payoff dominant convention locally and the risk dominant one globally; instead, if the fraction of fine reasoners is below that threshold, then selection favors a payoff dominant convention in which every agent plays the payoff dominant convention in each type of interaction.

The main contribution of this paper is to uncover the non-trivial interplay between the structure of interaction and the heterogeneity in the modes of reasoning for the evolution of conventions. In particular, it is found that more fine reasoners in the population do not necessarily lead to greater efficiency: in equilibrium coordination on the payoff dominant action may either increase or decrease, and this holds for miscoordination too. More specifically, the current paper brings together multiple streams of literature, providing contributions to each of them. First, it contributes to the literature on the evolution of conventions by showing that the heterogeneity in the mode of reasoning may matter in the selection of the ruling convention. Second, it contributes to the evolutionary literature on location-choice models by showing that heterogeneity in the mode of reasoning may give rise to the co-existence of conventions as well as to the full separation of agents' mode of reasoning even in the absence of frictions to mobility. Third, it contributes to the literature studying whether prosocial behaviors (such as social cooperation) are related to different modes of reasoning, showing that no mode of reasoning is always beneficial to cooperation in the Stag Hunt game but that, instead, what mode fosters cooperation depends on both the frequency of local interactions and the fraction of coarse reasoners. Finally, it provides interesting theoretical predictions – regarding, for example, the effects of cognitive manipulations on the likelihood that subjects coordinate on the payoff dominant convention – that can be empirically tested with laboratory experiments.

The remaining part of the paper is structured as follows: Section 2.2 illustrates the model, Section 2.3 and Section 2.4 present the main results, and Section 2.5 concludes. All proofs and intermediate results are collected in the appendices.

2.2 The Model

Consider a population of agents $\mathcal{N} = \{1, \ldots, N\}$ indexed by n and a set of locations $\mathcal{L} = \{1, \ldots, L\}$ with $L \geq 2$ indexed by ℓ. Time is discrete and denoted by $t \in \mathbb{N}$.

At the beginning of each period of time t, agents interact in a round-robin tournament fashion and the type of interaction they experience is determined probabilistically. More precisely, interactions can be either local or global in the following sense: with probability $p \in (0, 1)$ each agent interacts with every other agent staying in its current location (local interactions) while with probability $1 - p$ each agent interacts with every other agent in the population (global interactions). In either case, agents are randomly matched to play a Stag Hunt game as depicted in Figure 1. In the game, each player chooses an action between A (Hare, work individually) and B (Stag, cooperate). We assume that $b > a > c > d > 0$ and $a + c > b + d$ so that in a one-shot interaction A is risk dominant while B is payoff dominant.

		P2	
		A	B
P1	A	a, a	c, d
	B	d, c	b, b

Figure 1: The Stag Hunt game.

In the following, we will denote with $\alpha = \frac{a-d}{a-d+b-c}$ the probability with which action B is played in the mixed strategy equilibrium of the Stag Hunt game so that if a proportion α of agents plays action B then both A and B are best replies. We can also interpret α as the size of the basin of attraction of the risk dominant action A.

In this model, an agent can remain unmatched if interactions are local and if the agent is alone in its current location. We assume that an unmatched agent earns a reservation payoff of u. For simplicity, we set $u = 0$ although our results hold for any $u < d$, i.e., for any reservation payoff lower than the minimum payoff achievable in the Stag Hunt game.

There are two types of agents in the population, distinguished by their mode of reasoning. A fraction $q \in (0, 1)$ of the population is made of fine reasoners who can condition the action they play in the Stag Hunt game on the type of interaction (i.e., local or global) they are facing, while a fraction $1 - q$ of the population is made of coarse reasoners who play the same action in both types of interaction. We denote with $\mathcal{C}$ and $\mathcal{F}$, respectively, the set of coarse reasoners and the set of fine reasoners so that $\mathcal{C} \cup \mathcal{F} = \mathcal{N}$ and $\mathcal{C} \cap \mathcal{F} = \emptyset$. Moreover, we indicate with $i \in \mathcal{C}$ a generic coarse reasoner and with $j \in \mathcal{F}$ a generic fine reasoner.

We denote the strategy of agent $n \in \mathcal{N}$ at time t by means of a vector $\sigma_{nt} = (\ell_{nt}, l_{nt}, g_{nt})$ where $\ell_{nt} \in \mathcal{L}$ is the location chosen by agent n at time t, $l_{nt} \in \{A, B\}$ is the action agent n plays in local interactions, and $g_{nt} \in \{A, B\}$ is the action played in global interactions. By assumption it must be $l_{it} = g_{it}$ for every coarse reasoner $i \in \mathcal{C}$ as coarse reasoners cannot condition their action on the type of interaction they are facing. We indicate the state of the system at time t via a matrix $\Sigma_t = (\sigma_{nt})_{n \in \mathcal{N}} \in \Sigma$ where $\Sigma = \mathcal{L}^N \times \{A, B\}^N \times \{A, B\}^N$ is the state space. Finally, we denote the current state of the system excluding agent n' via the matrix $\Sigma_{-n't} = (\sigma_{nt})_{n \in \mathcal{N} : n \neq n'}$.

The system evolves according to synchronous myopic best reply with inertia and uniform random mistakes. More precisely, at the end of each period of time t, every agent $n \in \mathcal{N}$ has a fixed probability $\rho \in (0, 1)$ to revise its strategy. If given a revision opportunity, with probability $1 - \varepsilon \in (0, 1]$ the agent myopically best replies to the current state of the system Σ_t by choosing a strategy providing the maximum expected payoff and randomizing over best replies if the best reply is not unique; instead, with probability ε the agent makes a mistake and adopts a strategy uniformly at random. Formally, let

$$\pi_{nt}(\sigma) = \mathbb{E}[\pi(\sigma; \Sigma_{-nt})]$$

be the expected payoff of agent n associated to strategy σ given that the other agents adopt strategies Σ_{-nt}. With probability $1 - \varepsilon$ the agent best replies to the current state of the system, selecting with positive probability a strategy σ^* if and only if

$$\sigma^* \in \underset{\sigma}{\operatorname{argmax}} \ \pi_{nt}(\sigma)$$

Instead, with probability ε the agent makes a mistake and, so, it chooses a strategy uniformly at random.

The system described induces a Markov chain over the state space Σ. We say that the system evolves according to unperturbed dynamics if $\varepsilon = 0$, while we say that the system evolves according to perturbed dynamics if $\varepsilon > 0$. We study the two cases in turn.

2.3 Unperturbed Dynamics

In this section we study the dynamics of the system in the absence of mistakes: an agent myopically best replies to the current state of the system with probability one. In this dynamics every strict Nash equilibrium is an absorbing state: if the system reaches a state in which each agent adopts its unique best reply to the current state of the system at time t then it will stay there with probability one at any time $t + k$, for any $k > 0$. In the following, we characterize all the absorbing states of the system and we show that no other absorbing set exists.

We begin with some needed additional notation. We use an *ad hoc* labeling to refer to the states of the system where every coarse reasoner $i \in \mathcal{C}$ adopts strategy $\sigma_i = (\ell^*, X, X)$, $X \in \{A, B\}$, and where every fine reasoner $j \in \mathcal{F}$ adopts strategy $\sigma_j = (\ell^{**}, Y, Z)$, $Y, Z \in \{A, B\}$, made of the following three components:

- The action played by coarse reasoners: X;
- The spatial distribution of types: if coarse reasoners and fine reasoners stay in the same location, i.e., if $\ell^* = \ell^{**}$, then X will be followed by the symbol "-"; if, instead, coarse and fine reasoners are segregated according to type, i.e. if $\ell^* \neq \ell^{**}$, then X will be followed by the symbol "/";
- The actions played by fine reasoners in local (Y) and global (Z) interactions: YZ.

For example, the label A-BA refers to the states of the system in which $\sigma_i = (\ell^*, A, A)$ for every $i \in \mathcal{C}$ and $\sigma_j = (\ell^*, B, A)$ for every $j \in \mathcal{F}$ with everybody choosing the same location $\ell^* \in \mathcal{L}$. Note that there are L distinct states like this, one for each location in $\mathcal{L}$.

By adopting this notation we aim at stressing that the specific location chosen by agents is not of particular interest as all locations are identical:

what really matters is whether coarse and fine reasoners stay in the same or in a different location.

With this notation at hand, we can state our first result.

Theorem 1. *If the population is large enough, then all the states of the following types are absorbing:*

(1.1) A-AA, if $p, q \in (0, 1)$*;*

(1.2) A-AB, if $p \in (0, 1)$ *and* $q \in \big(\alpha, \min\big\{\frac{\alpha}{1-p}, 1\big\}\big)$*;*

(1.3) A-BA, if $p \in (0, 1)$ *and* $q \in \big(\alpha, \min\big\{\frac{\alpha}{p}, 1\big\}\big)$*;*

(1.4) A/BA, if $p \in \big(0, \frac{a-d}{b-d}\big)$ *and* $q \in (0, 1)$*;*

(1.5) B-AB, if $p \in (0, 1)$ *and* $q \in \big(1-\alpha, \min\big\{\frac{1-\alpha}{p}, 1\big\}\big)$*;*

(1.6) B-BA, if $p \in (0, 1)$ *and* $q \in \big(1-\alpha, \min\big\{\frac{1-\alpha}{1-p}, 1\big\}\big)$*;*

(1.7) B-BB, if $p \in (0, 1)$*;*

Further:

(1.8) there are no absorbing sets other than the states of types A-AA, A-AB, A-BA, A/BA, B-AB, B-BA, and B-BB.

The proof of Theorem 1 is reported in Appendix A. Here we provide a sketch of how we prove it. We first derive an expression for the expected payoff of an agent in a generic state of the system. Then, we obtain three conditions (one for coarse reasoners and two for fine reasoners) which allow us to easily determine the myopic best reply of an agent of either type in a generic state of the system. With these conditions at hand, we then show that each state of the types listed in Theorem 1 is a strict Nash equilibrium of the system and, consequently, an absorbing state. Finally, we show that there are no other absorbing sets of the system by providing an algorithm according to which, starting from any state, the system reaches with positive probability one of the absorbing states in Theorem 1, implying that all states different from those of the types listed in Theorem 1 are transient and not absorbing.

According to Theorem 1 the system is characterized by several types of absorbing states and most of them are such only in a sub-region of the parameter space (p, q) as it can be seen in Figure 2. A simple way to interpret this result is that in an absorbing state the following conditions must hold. First, in equilibrium agents of the same type must adopt the same strategy because they are identical. Second, given that the set of strategies available to coarse reasoners is a strict subset of the set of strategies available to fine reasoners, in equilibrium coarse reasoners cannot obtain a strictly higher expected payoff than fine reasoners. Third, in an absorbing state of the system, there must be some coordination between coarse reasoners and fine reasoners: they must coordinate at least in one among local and global interactions.

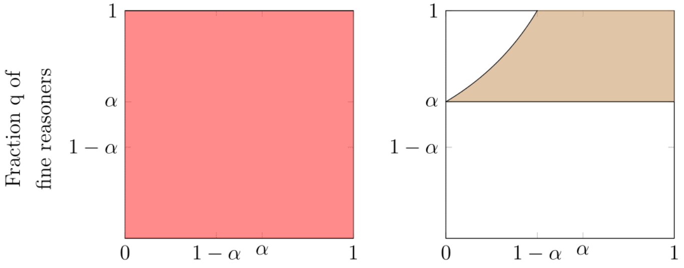
2.4 Perturbed Dynamics

In this section, we study the dynamics of the system in the presence of mistakes: an agent myopically best replies to the current state of the system with probability $1 - \varepsilon$ and adopts a new strategy at random with probability ε, with $\varepsilon > 0$. In the following, we focus on the long-run behavior of the system.

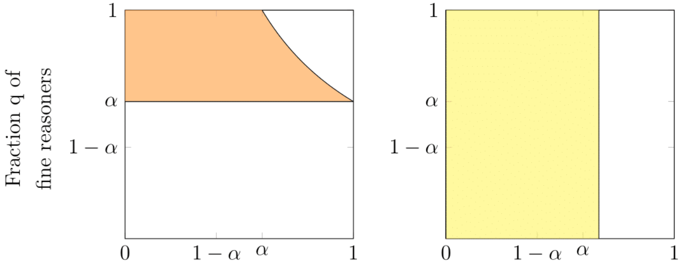
To study the long-run behavior of such a system, we employ two different techniques. In Section 2.4.1, we employ stochastic stability to identify the absorbing states (of the unperturbed dynamics) that may be selected in the long run, i.e., that may be stochastically stable for some value of p and q. Moreover, we provide sufficient conditions on p and q for an absorbing state to be selected. Then, in Section 2.4.2 we use agent-based simulations of the model to analyze the long-run behavior of the system in the entire (p, q)-space.

2.4.1 Stochastic Stability Analysis

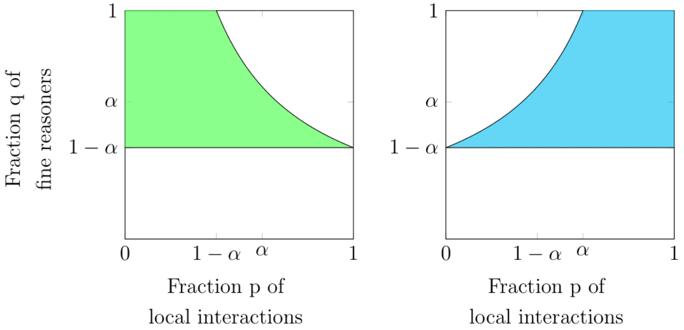
As already mentioned in Chapter 1, the stochastically stable states of the system must belong to some absorbing set or state of the system under unperturbed dynamics (H. P. Young, 1993). In our setup, this implies that only the absorbing states of the types listed in Theorem 1 are candidates to be stochastically stable states.


(a) A-AA, B-BB

(b) A-AB


(c) A-BA

(d) A/BA


(e) B-AB

(f) B-BA

Figure 2: Absorbing states according to Theorem 1. States that according to Theorem 1 can give rise to absorbing states of the system in the (p, q) parameter space in which they actually are absorbing states. Case $a = 4$, $b = 5$, $c = 3$, $d = 1$ and, consequently, $\alpha = 0.6$ and $(a - d)/(b - d) = 0.75$.

To determine the stochastically stable states of the system we employ radius-coradius arguments (Ellison, 2000). We apply them to opportunely selected partitions $\{\Omega, \Omega^{-1}\}$ of the set of absorbing states and then we compute (i) the minimum number of mistakes required to leave the basin of attraction of Ω, i.e., the radius of Ω or $R(\Omega)$, and (ii) the maximum minimum number of mistakes required to enter into the basin of attraction of Ω starting from an absorbing state belonging to Ω^{-1}, i.e., the coradius of Ω or $CR(\Omega)$. If $R(\Omega) > CR(\Omega)$ then all the stochastically stable states of the system are contained in Ω. We refer to stochastically stable *sets* to indicate sets containing only stochastically stable states. This is useful when we refer to the types of states listed in Figure 2 because if any single state of a given type is stochastically stable then all states of that type are so.

In Theorem 2 we identify the types of absorbing states of the system that are never stochastically stable.

Theorem 2. *If the population is large enough, then all the absorbing states of the types A-AA, A-AB, A-BA, and B-AB are never stochastically stable.*

The proof is reported in Appendix B and is based on radius-coradius arguments applied to two sets of absorbing states, namely Ω and {A-AA, A-AB, A-BA, B-AB}. The main idea is to show that the radius of the basin of attraction of Ω is strictly larger than one, while its coradius is equal to one. This implies that all the stochastically stable states are contained in Ω and, consequently, all the absorbing states of the types A-AA, A-AB, A-BA, and B-AB are never stochastically stable.

We repeat this reasoning by analyzing separately two following cases: the case $q \in \left(0, \min\left\{\frac{1-\alpha}{1-p}, 1\right\}\right)$ and the case $q \in \left(\min\left\{\frac{1-\alpha}{1-p}, 1\right\}, 1\right)$. In the first case we set $\Omega =$ {A/BA, B-BA, B-BB} while in the second case we consider $\Omega =$ {A/BA, B-BB}. We repeat the analysis twice because if $q \in \left(\min\left\{\frac{1-\alpha}{1-p}, 1\right\}, 1\right)$ then states of the type B-BA are not absorbing states and the parameters p and q are such that starting from a state of the type B-BA the system may reach with positive probability – i.e., without any mistake taking place – an absorbing state of the type A-BA which does not belong to the basin of attraction of Ω. Therefore, if we set $\Omega =$ {A/BA, B-BA, B-BB} in the entire parameter space, then the radius of Ω would be equal to zero and, hence, we could not establish whether absorbing states of the types A-AA, A-AB, A-BA, or B-AB are stochastically stable or not.

By stating that all the absorbing states of the types A-AA, A-AB, A-BA, and B-AB are never stochastically stable, Theorem 2 implies that only absorbing states of the types A/BA, B-BA, and B-BB are candidate stochastically stable states of the system. This result allows us to derive a first conclusion: in a stochastically stable state of the system fine reasoners must play the payoff dominant action, B, in local interactions. The main reason behind this result is that starting from any absorbing state in which fine reasoners play the risk dominant action in local interactions, i.e., absorbing states of the types A-AA, A-AB, and B-AB, one mistake is enough to lead with positive probability the system into another absorbing state in which fine reasoners play the payoff dominant action in local interactions. In fact, if one agent makes a mistake and adopts strategy (ℓ', B, Z) where ℓ' is an empty location and $Z \in \{A, B\}$, then every fine reasoner will strictly prefer strategy (ℓ', B, Y) with $Y \in \{A, B\}$ to its current strategy (ℓ^*, A, Y) involving playing the risk dominant action locally as this new strategy allows perfect local coordination on the payoff dominant convention.

In Theorem 3 we further analyze the set of stochastically stable states of the system, and we provide sufficient conditions for the stochastic stability of the absorbing states of the types A/BA, B-BA, and B-BB.

Theorem 3. *If the population is large enough, then all and only absorbing states of the following types are stochastically stable:*

(3.1) A/BA, if $p \in \left(0, \frac{2\alpha-1}{\alpha}\right)$ *and* $q \in (0, 1)$;

(3.2) B-BA, if $p \in \left(\frac{a-d}{b-d}, 1\right)$ *and* $q \in \left(2(1-\alpha), 1\right)$;

(3.3) B-BB, if $p \in \left(\frac{a-d}{b-d}, 1\right)$ *and* $q \in \left(0, 1-\alpha\right)$.

The proof is reported in Appendix C. The proof requires several steps. However, each step follows approximately the same logic: we first construct a partition of the set of absorbing states of the system into Ω and

Ω^{-1} such that the set Ω contains only one among the remaining candidate types of stochastically stable states (A/BA, B-BA, and B-BB), we then compute the radius and the coradius of Ω and we finally find the region in the (p, q) parameter space such that the radius is larger than the coradius. Note that if convenient we also include into the set Ω absorbing states that according to Theorem 2 are never stochastically stable. The inclusion of such absorbing states may simplify the computation of the coradius of Ω or it may be necessary to guarantee that the radius is larger than the coradius. In either case, this does not invalidate our conclusions as by Theorem 2 such states are never stochastically stable.

Sketch of the Proof. For the proof of point (3.1) we have to consider separately two regions of the parameter space and compute the radius and the coradius for different partitions of the set of absorbing states.

First, we analyse the case $p \in \big(0, \frac{2\alpha-1}{\alpha}\big)$ and $q \in \big(0, \min\big\{\frac{1-\alpha}{1-p}, 1\big\}\big)$ and we consider the following partition of absorbing sets: $\Omega = \{$A/BA, A-AA$\}$ and $\Omega^{-1} = \{$B-BA, B-BB, A-AB, A-BA, B-AB$\}$. We find that the radius of Ω is larger than its coradius if at least $p \in \big(0, \frac{2\alpha-1}{\alpha}\big)$. Therefore, if $p \in \big(0, \frac{2\alpha-1}{\alpha}\big)$ all the stochastically stable states of the system are contained in Ω; but then, given that states of the type A-AA are never stochastically stable, we can conclude that if $p \in \big(0, \frac{2\alpha-1}{\alpha}\big)$ then the set of absorbing states of the type A/BA is stochastically stable.

Second, we consider the case $p \in \big(0, \frac{2\alpha-1}{\alpha}\big)$ and $q \in \big(\min\big\{\frac{1-\alpha}{1-p}, 1\big\}, 1\big)$ and we set $\Omega' = \{$A/BA, A-AA, A-BA$\}$ and $\Omega'^{-1} = \{$A-AB, B-AB, B-BB$\}$. We again show that if $p \in \big(0, \frac{2\alpha-1}{\alpha}\big)$ then the radius of Ω' is strictly larger than its coradius and, so, we conclude that also if $p \in \big(0, \frac{2\alpha-1}{\alpha}\big)$ and $q \in \big(\min\big\{\frac{1-\alpha}{1-p}, 1\big\}, 1\big)$ then the set of states of the type A/BA is stochastically stable as by Theorem 2 absorbing states of the types A-AA and A-BA are never stochastically stable.

By combining the two previous results we can conclude that if $p \in \big(0, \frac{2\alpha-1}{\alpha}\big)$ then the set of absorbing states of the type A/BA is stochastically stable.

To prove point (3.2) we analyse the case $p \in \big(\frac{a-d}{b-d}, 1\big)$ and we consider the partition of absorbing states into $\Omega'' = \{$A-AA, A-BA, B-BA$\}$ and $\Omega''^{-1} = \{$B-BB, A-AB, B-AB$\}$. We find that if $q \in \big(2(1-\alpha), 1\big)$ then the radius of Ω'' is larger than its coradius and, consequently, all the stochastically stable states are contained into Ω''. But then, given that by Theorem 2 states of the types A-AA and A-BA are never stochastically stable, if $p \in \big(\frac{a-d}{b-d}, 1\big)$ and $q \in \big(2(1-\alpha), 1\big)$ the set of states of the type B-BA is stochastically stable.

Finally, point (3.3) of Theorem 3 is a direct implication of Theorem 1 and Theorem 2. In fact, by Theorem 1 only states of the types A-AA and B-BB are absorbing if $p \in \big(\frac{a-d}{b-d}, 1\big)$ and $q \in \big(0, 1-\alpha\big)$; while by Theorem 2 absorbing states of the type A-AA are never stochastically stable. $\square$

Note that Theorem 3 predicates the stochastic stability of the sets of all states of each type A/BA, B-BA, and B-BB, i.e., that the states of a given type are either all stochastically stable or none. This is because each state of a given type is equally likely to be reached, in the following sense: if the system is in an absorbing state, say σ', of a given type, then (i) either one or two mistakes can lead the system with positive probability into any other absorbing state of the same type, say σ'', and (ii) the same number of mistakes (either one or two) may lead the system with positive probability from σ'' to σ'. This fact is a direct consequence of the assumption that locations are identical: given that all locations are the same, there is no reason why agents should prefer one location over another in the long run and so, stochastic stability analysis cannot select among states of the same type characterized by different equilibrium location choices.

In Figure 3 we provide a graphical representation of the implications of Theorem 3 for a given set of payoffs ($a = 4$, $b = 5$, $c = 3$, and $d = 1$) by coloring the regions in the (p, q)-space where states of type A/BA, B-BA, and B-BB are or can be stochastically stable.

If the frequency of local interactions is sufficiently small, then the set of states of the type A/BA is stochastically stable. In other words, a semi-risk dominant convention is selected: if interactions are global all agents play the risk dominant action, instead if interactions are local coarse reasoners coordinate on the risk dominant action, while fine reasoners play the payoff dominant action. However, given that agents are segregated according to their type, agents perfectly coordinate among themselves in both local and global interactions. This guarantees the sustainability of the semi-risk dominant convention independently of the fraction of fine reasoners q in the population.

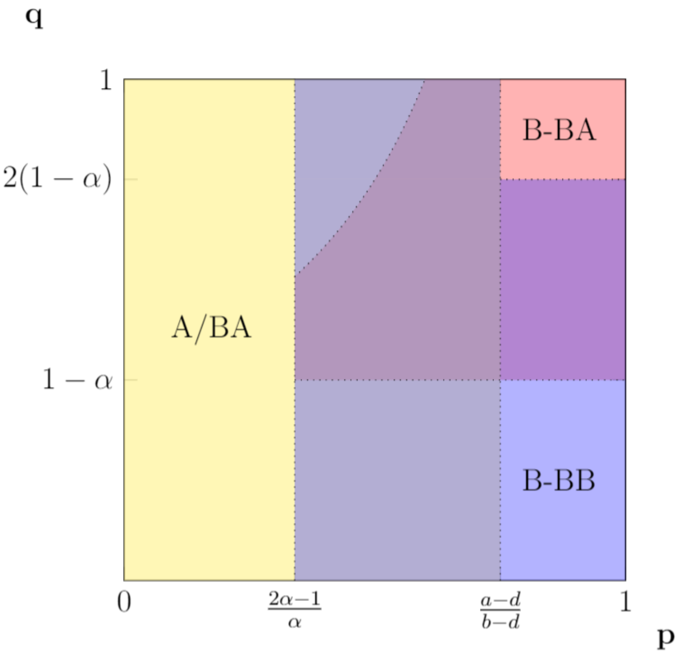


Figure 3: Stochastically stable states according to Theorem 3. Types of stochastically stable states according to Theorem 3 in the (p,q) parameter space together with uncertainty regions. Case $a = 4$, $b = 5$, $c = 3$, $d = 1$ and, consequently, $\alpha = 0.6$ and $(a-d)/(b-d) = 0.75$.

On the contrary, if the frequency of local random interactions is sufficiently high, then two different scenarios are possible and, more precisely, the actual long-run prediction depends on the fraction of fine reasoners in the society.

First, if the fraction of fine reasoners is sufficiently large, then the set of states of the type B-BA is stochastically stable, selecting a semi-payoff dominant convention. In such a convention if interactions are local then all agents play the payoff dominant action; instead, if interactions are global fine reasoners play the risk dominant action, while coarse reasoners stick to the payoff dominant action. Moreover, all agents stay in a single location. In this convention, miscoordination may arise. It happens in every global interaction in which a fine reasoner is matched with a coarse reasoner. As a consequence, the probability of miscoordination is decreasing both in the probability of local interactions p and in the fraction of fine reasoners in the society q.

If instead, the fraction of fine reasoners is sufficiently small, then the set of states of the type B-BB is stochastically stable. So, stochastic stability selects the payoff dominant convention in which all agents play the payoff dominant action independently of the type of interaction they face; moreover, all agents stay in the same location. This convention allows for perfect coordination on the payoff dominant for all levels of p and q that do not compromise stochastic stability.

Figure 3 also highlights that Theorem 3 characterizes stochastically stable states only for a portion of the parameter space. Broadly speaking there are two areas of uncertainty regarding what states are stochastically stable. For intermediate values of the probability of local interactions p we cannot say whether the semi-risk dominant convention or the two payoff dominant conventions are selected, while for intermediate values of the fraction of fine reasoners q and a sufficiently high probability of local interactions p it is unclear whether the semi-payoff or the payoff dominant conventions are selected.

Overall, these findings suggest that if interactions are mostly global social coordination is attained together with segregation by types of reasoners and with fine reasoners being more often collaborative in local interactions; instead, if interactions are mostly local social coordination is attained without segregation but possibly with fine reasoners being less often collaborative in global interactions.

Hence, on the one hand, a greater frequency of local interactions can be very beneficial for the diffusion of the payoff dominant convention since it can move the system from a semi-risk dominant convention to a (semi-)payoff dominant convention. On the other hand, a greater frequency of fine reasoners in the population may both be positive or negative for the diffusion of the payoff dominant action. In fact, if interactions are mostly global then the higher the fraction of fine reasoners in the population the more agents play the payoff dominant action in local interactions. Instead, if interactions are mostly local, then a higher fraction of fine reasoners in the population can have two distinct negative effects on the diffusion of the payoff dominant action: first, it may cause a switch from the payoff dominant convention (B-BB) to the semi-payoff dominant convention (B-BA), second if the semi-payoff dominant convention is in place a higher frequency of fine reasoners decreases the use of the payoff dominant action in global interactions.

Interestingly, there is an interaction between the frequency of local interactions and the fraction of fine reasoners in the population: the gains associated with an increase in the frequency of local interactions is weakly decreasing in the fraction of fine reasoners in the population. This holds because if the population of fine reasoners is large enough, then an increase in the frequency of local interactions implies a switch from the semi-risk dominant convention (A/BA) to the semi-payoff dominant convention (B-BA) rather than to the payoff dominant convention (B-BB).

2.4.2 Simulation Analysis

The stochastic stability analysis performed in Section 2.4.1 provides sharp predictions about the long-run behavior of the system in the perturbed dynamics, but only for a portion of the (p, q) parameter space. To complete the analysis we employ agent-based simulations exploring the long-run behavior of the system in the entire parameter space. An additional advantage of agent-based simulations is that they allow us to study the evolution of the system for fixed small population size and non-vanishing mistake probability, less demanding conditions with respect to those of Theorem 2 and Theorem 3.

We run simulations setting $N = 100$, $L = 3$, $\rho = 0.2$, and $\varepsilon = 0.1$. Moreover, we explore the (p, q)-space for $p, q \in \{0.1, 0.2, \ldots, 0.9\}$ for a total of $9^2 = 81$ combinations. For each of these combinations, we perform 10 independent simulations of the model and in each simulation we run the system for 10^5 iterations.

In each iteration every agent is given a revision opportunity with probability 0.2 and, if the case, the agent revises its strategy according to myopic best reply with probability 0.9 (selecting the strategy that maximizes the expected payoff and randomizing over best replies if the best reply is not unique); instead, with probability 0.1, the agent selects a strategy uniformly at random.

The results presented below are obtained by assuming the following payoffs in the Stag Hunt game: $a = 4$, $b = 5$, $c = 3$, and $d = 1$. We have selected these parameter values so as to avoid extreme values of the thresholds of interest, specifically getting $\alpha = 0.6$ and $\frac{a-d}{b-d} = 0.75$. In Appendix D we report robustness tests of the findings presented here by considering variations of these payoffs.

Applying Theorem 1 we know what types of states are absorbing in which regions of the (p, q) parameter space, but by Theorem 2 we can exclude that states of type A-AA, A-AB, A-BA, and B-AB are visited a significant amount of time in the long run. By Theorem 3 we obtain that the system spends most of the time in absorbing states of the type A/BA if $p \in (0, 0.33)$, of the type B-BA if $p \in (0.75, 1)$ and $q \in (0.8, 1)$, and of the type B-BB if $p \in (0.75, 1)$ and $q \in (0, 0.4)$. We summarize these conclusions in Table 1. Note that the system is never expected to spend most of the time in a specific absorbing state of a given type, but rather to move frequently from one absorbing state to another of the same type but characterized by a different location or locations.

Simulation results are illustrated in Figure 4. Specifically, for each pair $(p, q) \in \{0.1, 0.2, \ldots, 0.9\}^2$ we report the average fraction of agents playing the payoff dominant action by type of agent (coarse vs fine reasoner) and type of interaction (local vs global interaction). These average values have been computed as follows: first, for each simulation, we compute the fraction of coarse [fine] reasoners playing the payoff dominant action in local [global] interactions in each time period $t = 1000, \ldots, 100000$; second, we compute the average fraction of coarse [fine] reasoners playing the payoff dominant action in local [global] interactions to obtain a simulation-specific average; third, we average these values over the 10 independent simulations performed for a given pair (p, q).

The results in Figure 4 suggest that for $p \leq 0.6$ the system spends

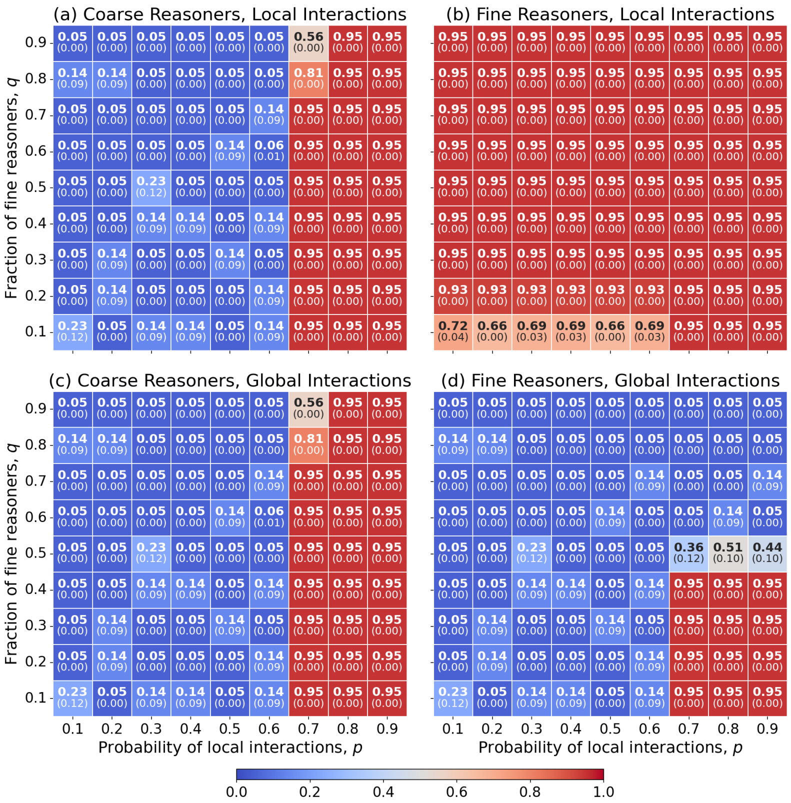
Type	Absorbing		Stochastically Stable	
	p	q	p	q
A-AA	$(0.00, 1.00)$	$(0.00, 1.00)$	$\emptyset$	$\emptyset$
A-AB	$(0.00, 1.00)$	$(0.60, \frac{0.60}{1-p})$	$\emptyset$	$\emptyset$
A-BA	$(0.00, 1.00)$	$(0.60, \frac{0.60}{p})$	$\emptyset$	$\emptyset$
A/BA	$(0.00, 0.75)$	$(0.00, 1.00)$	$(0.00, 0.33)$	$(0.00, 1.00)$
B-AB	$(0.00, 1.00)$	$(0.40, \frac{0.40}{p})$	$\emptyset$	$\emptyset$
B-BA	$(0.00, 1.00)$	$(0.40, \frac{0.40}{1-p})$	$(0.75, 1.00)$	$(0.80, 1.00)$
B-BB	$(0.00, 1.00)$	$(0.00, 1.00)$	$(0.75, 1.00)$	$(0.00, 0.40)$

Table 1: Theoretical predictions for the simulations setup. Values of p and q for which states of a given type are absorbing according to Theorem 1 and for which absorbing states of a given type are guaranteed to be stochastically stable according to Theorem 2 and Theorem 3. Case $a = 4$, $b = 5$, $c = 3$, $d = 1$.

most of the time in states of the type A/BA, independently of the fraction of fine reasoners q. Instead, for $p > 0.6$ two alternative scenarios are possible. More precisely, if $q < 0.5$ then the system spends most of the time in states of the type B-BB, while if $q \geq 0.5$ then the system spends most of the time in states of the type B-BA. We illustrate these findings graphically in Figure 5 together with the results in Section 2.4.1.

In Appendix D we report the results obtained for different values of the payoffs in the Stag Hunt game. By combining these results it appears that the threshold value of the probability of local interactions, p^* – which separates the area in which the semi-risk dominant convention (A/BA) is stochastically stable from the area in which the payoff dominant conventions (B-BA and B-BB) are stochastically stable – is always between α and $\frac{a-d}{b-d}$. Moreover, the threshold value of the fraction of fine reasoners in the population, q^* – that separates the region where the payoff dominant convention (B-BB) is selected from the region where the semi-payoff dominant convention (B-BA) is selected – approaches its lower-bound $(1-\alpha)$ if α is high, while it tends to be closer to its upper-bound $2(1-\alpha)$ if α is low.

Overall, these findings suggest that the qualitative results provided by Theorem 3 hold more generally: the threshold values for the suffi-



(a) Coarse Reasoners, Local Interactions

	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.9	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.56 (0.00)	0.95 (0.00)	0.95 (0.00)
0.8	0.14 (0.09)	0.14 (0.09)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.81 (0.00)	0.95 (0.00)	0.95 (0.00)
0.7	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.14 (0.09)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)
0.6	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.14 (0.09)	0.06 (0.01)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)
0.5	0.05 (0.00)	0.05 (0.00)	0.23 (0.12)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)
0.4	0.05 (0.00)	0.05 (0.00)	0.14 (0.09)	0.14 (0.09)	0.05 (0.00)	0.14 (0.09)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)
0.3	0.05 (0.00)	0.14 (0.09)	0.05 (0.00)	0.05 (0.00)	0.14 (0.09)	0.05 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)
0.2	0.05 (0.00)	0.14 (0.09)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.14 (0.09)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)
0.1	0.23 (0.12)	0.05 (0.00)	0.14 (0.09)	0.14 (0.09)	0.05 (0.00)	0.14 (0.09)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)

(b) Fine Reasoners, Local Interactions

	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.9	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)
0.8	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)
0.7	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)
0.6	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)
0.5	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)
0.4	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)
0.3	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)
0.2	0.93 (0.00)	0.93 (0.00)	0.93 (0.00)	0.93 (0.00)	0.93 (0.00)	0.93 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)
0.1	0.72 (0.04)	0.66 (0.00)	0.69 (0.03)	0.69 (0.03)	0.66 (0.00)	0.69 (0.03)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)

(c) Coarse Reasoners, Global Interactions

	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.9	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.56 (0.00)	0.95 (0.00)	0.95 (0.00)
0.8	0.14 (0.09)	0.14 (0.09)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.81 (0.00)	0.95 (0.00)	0.95 (0.00)
0.7	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.14 (0.09)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)
0.6	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.14 (0.09)	0.06 (0.01)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)
0.5	0.05 (0.00)	0.05 (0.00)	0.23 (0.12)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)
0.4	0.05 (0.00)	0.05 (0.00)	0.14 (0.09)	0.14 (0.09)	0.05 (0.00)	0.14 (0.09)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)
0.3	0.05 (0.00)	0.14 (0.09)	0.05 (0.00)	0.05 (0.00)	0.14 (0.09)	0.05 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)
0.2	0.05 (0.00)	0.14 (0.09)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.14 (0.09)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)
0.1	0.23 (0.12)	0.05 (0.00)	0.14 (0.09)	0.14 (0.09)	0.05 (0.00)	0.14 (0.09)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)

(d) Fine Reasoners, Global Interactions

	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.9	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)
0.8	0.14 (0.09)	0.14 (0.09)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)
0.7	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.14 (0.09)	0.05 (0.00)	0.05 (0.00)	0.14 (0.09)
0.6	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.14 (0.09)	0.05 (0.00)	0.05 (0.00)	0.14 (0.09)	0.05 (0.00)
0.5	0.05 (0.00)	0.05 (0.00)	0.23 (0.12)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.36 (0.12)	0.51 (0.10)	0.44 (0.10)
0.4	0.05 (0.00)	0.05 (0.00)	0.14 (0.09)	0.14 (0.09)	0.05 (0.00)	0.14 (0.09)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)
0.3	0.05 (0.00)	0.14 (0.09)	0.05 (0.00)	0.05 (0.00)	0.14 (0.09)	0.05 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)
0.2	0.05 (0.00)	0.14 (0.09)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.14 (0.09)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)
0.1	0.23 (0.12)	0.05 (0.00)	0.14 (0.09)	0.14 (0.09)	0.05 (0.00)	0.14 (0.09)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)

Figure 4: Simulations results. Average proportion of agents playing the payoff dominant action B by type of agent (coarse vs fine reasoner) and type of interaction (local vs global interaction); standard errors in parentheses. Case $a = 4$, $b = 5$, $c = 3$, $d = 1$.

cient conditions seem to extend naturally to nearby regions of the (p, q) parameter space and they hold also for non-vanishing mistakes probabilities and relatively small population size.

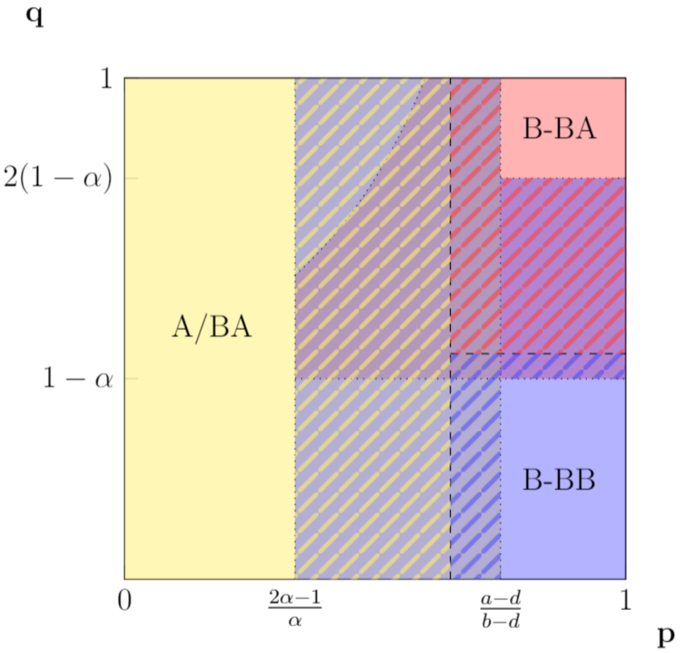


Figure 5: Stochastically stable states by combining Theorem 3 and simulation results. Types of stochastically stable states in the (p, q) parameter space: uniformly colored regions refer to both stochastic stability and simulation results, while striped regions refer to simulation results only. Case $a = 4$, $b = 5$, $c = 3$, $d = 1$.

2.5 Discussion

We have considered a finite population of agents who myopically best reply with a small probability of making a mistake, and we have studied how the mode of reasoning and the structure of interaction can affect the evolution of conventions in this setup.

We have found that if interactions are mostly global then selection favors conventions where agents are separated into different locations according to their mode of reasoning, with coarse reasoners playing the risk dominant action in all interactions and fine reasoners playing the

Qualitative summary of results

fraction q of fine reasoners	“high”	**semi-risk dominant** fine reasoners locally play payoff dominant	**semi-payoff dominant** fine reasoners globally play risk dominant
	“low”	**semi-risk dominant** fine reasoners locally play payoff dominant	**payoff dominant**
↑	→	“low”	“high”
		fraction p of local interactions	

Table 2: Summary of results. Qualitative summary of results in terms of conventions selected in the long run.

risk dominant action globally and the payoff dominant action locally. If instead, interactions are mostly local, then all agents stay in the same location and there are two cases: if coarse reasoners are sufficiently numerous then all agents play the payoff dominant action both globally and locally, while if coarse reasoners are not numerous enough then they play the payoff dominant action and fine reasoners play the risk dominant action globally and the payoff dominant action locally. This implies that the co-existence of coarse and fine reasoning may favor or hamper the adoption of the payoff dominant action depending on the structure of interactions. Table 2 summarizes such qualitative features of our results.

We have provided sufficient conditions for these results based on the fraction of fine reasoners and the frequency of local interactions. We have also provided numerical simulations of the model indicating that our findings hold in the entire parameter space, for small but non-negligible probabilities of mistakes, and for populations of size as small as one hundred agents.

From the comparison of the welfare associated with viable long-run outcomes, we can draw some conclusions. By looking at Table 2, we see that increasing the fraction of fine reasoners (moving vertically in the table) turns out to be welfare enhancing if interactions are mostly global because the long-run outcome remains the same (semi-risk dominant convention) and fine reasoners are more prosocial than coarse reasoners in this case. Instead, if interactions are mostly local then increasing the fraction of fine reasoners turns out to be welfare-reducing because it changes the long-run outcome from full adoption of Stag (payoff dominant convention) to a partial adoption of it (semi-payoff dominant convention). Furthermore, we notice that increasing the fraction of local interactions (moving horizontally in the table) is always welfare-enhancing.

A simplifying assumption in our model is that agents are always fine reasoners or coarse reasoners. In fact, we can expect real decision-makers to reason sometimes coarsely and sometimes finely depending on the choice problem faced, i.e., we can expect the mode of reasoning to be endogenous. Such endogeneity could disrupt the equilibria where fine reasoners and coarse reasoners co-exist, behave differently, and possibly live in different locations, as all agents might end up facing the same payoff structure and, therefore, following the same optimal behavior. However, if we allow for heterogeneity in cognitive abilities, i.e., different costs to reason finely, we may obtain that in equilibrium agents have heterogeneous probabilities to resort to fine reasoning, so that an equilibrium may arise where individuals are segregated in different locations and behave differently according to their mode of reasoning. Further research is needed to establish the precise conditions for this to happen, but we can already note that only for states where everybody plays the payoff dominant action both locally and globally (i.e., of the B-BB type) we can be confident that all agents will reason coarsely (since coordination is obtained without conditioning actions to the type of interaction), while for all other states which turned out to be stochastically stable in our model (i.e., of the A/BA and B-BA types) there are good reasons to expect that those with small costs of fine reasoning will reason finely (since conditioning one's action to the type of interaction allows better coordination) while those with high costs will not.

If, instead, we follow the literature on cognition and cooperation (Bear and David G Rand, 2016; Bear, Kagan, and David G Rand, 2017; Jagau and Veelen, 2017; Mosleh and David G Rand, 2018) in assuming that all agents draw their cost to reason finely from a common distribution, then our intuition is that the quality of our results would be preserved only under additional assumptions. One promising candidate in this regard is the assumption that the choice of location is conditional on the mode of reasoning, as in this case behaviors would potentially be ex-post heterogeneous across locations and modes of reasoning. Another promising candidate is to introduce a constraint to the maximum number of agents to interact with, either by introducing constraints at the local level (Dieckmann, 1999; Anwar, 2002; Shi, 2013; Pin, E. Weidenholzer, and S. Weidenholzer, 2017) or by imposing a maximum number of neighbors in a network structure (Staudigl and S. Weidenholzer, 2014; Bilancini and Boncinelli, 2018; Cui and S. Weidenholzer, 2021), which would work against having all agents in the same location while adopting the same mode of reasoning and action.

Further, keeping the mode of reasoning exogenous allows us to predict, in a simple and neat way, the effects of controlled manipulations of the extent of fine reasoning, as one would attempt to do in behavioral experiments. Actually, our findings can shed light on the existing experimental evidence on the effects of manipulating reasoning in the one-shot Stag Hung game (Belloc et al., 2019; Bilancini, Boncinelli, and Paolo, 2021) which suggests that greater reflection leads to a greater likelihood of playing the risk dominant action. In particular, if one interprets the interactions in experiments as global, then fine reasoning during the experiment would lead to playing the risk dominant action more or less often with respect to coarse reasoning depending on what coarse reasoners do in equilibrium, i.e., depending on the ruling social convention that coarse reasoners take with them in the experiment. So, the fact that greater reflection leads to a greater probability of playing the risk dominant action might be due to the fact that the ruling convention is the semi-payoff dominant (as indicated in Table 2).

Our results can also be interpreted in light of the recent lively debate on what mode of reasoning is more conducive to prosocial behavior (Zaki and Mitchell, 2013; Capraro, 2019). If we interpret Stag as a risky and collaborative action and Hare as a safe and individualistic action (Skyrms, 2004), then coordination on Stag can be seen as more prosocial than Hare. From this perspective, we can conclude that fine reasoning is more prosocial than coarse reasoning if interactions are mostly global; if, instead, interactions are mostly local, then coarse reasoning is as prosocial (when fine reasoners are few) or more prosocial (if fine reasoners are numerous) than fine reasoning.

A last remark stems from the observation that fine reasoners and coarse reasoners may end up in the long run living in different locations, which is what happens in A/BA: this outcome is reminiscent of real-world phenomena characterized by globalization and polarization, which is a widespread phenomenon at least for online interactions. We stress that when this cognitive segregation happens, the frequency of interaction with a fine reasoner is larger for a fine reasoner than for a coarse reasoner. This creates a novel form of assortativity – namely, assortativity in cognition – which deserves further investigation both theoretically, along the lines of Bilancini, Boncinelli, and Vicario (2022), and empirically.

Chapter 3

The Co-Evolution of Cooperation and Cognition Under Endogenous Population Structure

3.1 Introduction

In this paper, we study the co-evolution of cooperation and cognition in the presence of an endogenous interaction structure. More precisely, we analyze a model that builds on Bear and David G Rand (2016) while (i) addressing the limitations highlighted in Jagau and Veelen (2017), (ii) allowing for an endogenous interaction structure (Santos, Pacheco, and Lenaerts, 2006; Mosleh and David G Rand, 2018), and (iii) being analytically tractable.

The focus is on the evolutionary emergence and sustainability of individual cooperation (i.e., paying a cost to benefit someone else) in a prisoners' dilemma game (Axelrod and Hamilton, 1981; David G. Rand and Martin A. Nowak, 2013). More precisely, we study the evolution of cooperation in a large population of agents playing a prisoners' dilemma game which can be either one-shot and anonymous or infinitely repeated.

In both types of prisoners' dilemmas, agents can play one among the following two actions: Always Defect (AllD) and Tit-for-Tat (TFT). The payoff structures of the games are such that Always Defect is strictly dominant in the anonymous and one-shot game, while Tit-for-Tat is weakly dominant in the repeated game.

Evolution also shapes players' mode of reasoning which is modeled following dual process theories of cognition (Evans and Stanovich, 2013). Broadly speaking, dual process theories of cognition conjecture the existence of two possible modes or types of reasoning: one fast, cheap, and heuristic-based (intuition), the other slow, costly, and based on cost-benefit evaluations (deliberation). In our model, intuition gives rise to a generalized or unconditional response over average opponents' behavior. Instead, deliberation leads to actions that are conditional on whether the game played is one-shot and anonymous or repeated, but it comes at a cost. This generates a trade-off between error-prone intuition and costly deliberation.

An endogenous interaction structure is introduced via a set of locations. Topologically speaking, the presence of a set of locations gives rise to a series of disconnected fully-connected networks: within each location, all players are connected and may be randomly matched into couples to play a prisoners' dilemma game, while players staying in different locations are not connected and, thus, have no chance to interact. However, whenever given a revision opportunity a player may decide to migrate from one location to another in order to maximize its payoff. Within this setup, we analyze whether an endogenous interaction structure may foster the emergence and/or the evolutionary sustainability of cooperation.

We analyze the co-evolution of cooperation and cognition by means of stochastic stability techniques (Foster and Young, 1990; H. P. Young, 1993; Kandori, Mailath, and Rob, 1993) and, more precisely, we determine the stochastically stable states of the system via radius-coradius arguments (Ellison, 2000). Broadly speaking, the stochastically stable states of the system are those states that are relatively easier to reach in terms of mistakes if we assume that agents can make mistakes with small but positive probability.

Our main contribution is the analytical evolutionary account of the emergence of cooperation with players deciding to play TFT under intuition and deliberating with positive probability if repeated games are frequent enough. This result may also explain why in experimental settings individuals are observed to cooperate even in one-shot anonymous interactions.

The remaining part of the paper is structured as follows: Section 3.2 describes the model, Section 3.3 provides two examples of stochastic stability analysis in an alternative version of the model with a single location, Section 3.4 illustrates the main results, Section 3.5 checks the robustness of the results with respect to the choice of the revision protocol, and Section 3.6 concludes.

3.2 The model

Consider a set, $\mathcal{A}$, of A agents or players indexed by $a = 1, \ldots, A$ and a set, $\mathcal{L}$, of $L \geq 2$ locations denoted by $\ell = 1, \ldots, L$. Time is discrete and indexed by $t = 0, 1, ...$

In each period of time t, each agent randomly interacts with another agent staying in the same location. Randomly matched agents play a one-shot and anonymous prisoners' dilemma (PD) with probability $1-p$, while with probability p they play an infinitely repeated PD. In Figure 6 we report the normal-form representation of the two games.

		a'	
		TFT	AllD
a	TFT	$b-c$ $b-c$	$-c$ b
	AllD	b $-c$	0 0

One-shot PD

		a'	
		TFT	AllD
a	TFT	$b-c$ $b-c$	0 0
	AllD	0 0	0 0

Infinitely Repeated PD

Figure 6: One-shot and infinitely repeated Prisoners' Dilemmas. Normal-form representation of the one-shot and anonymous Prisoners' dilemma (to the left) and of the infinitely repeated Prisoners' dilemma (to the right).

In both games in every round of play, each agent must decide whether to pay a cost $c > 0$ to provide a benefit $b > c$ to the other agent. Following Bear and David G Rand (2016), agents can play either Always Defect (AllD) or Tit-For-Tat (TFT): by playing AllD an agent decides to never pay the cost of cooperation while by playing TFT the agent pays the cost of cooperation in the first round and in all the subsequent rounds of play it copies the choice made by the opponent in the previous round. In case of an infinitely repeated PD agents' payoff corresponds to the average payoff obtained in the individual rounds of play and, consequently, in such a game agents can obtain a strictly positive payoff, $b - c$, only if both play TFT otherwise they both receive a null payoff. Finally, we assume that if a player is alone in a given location, then it plays alone: it eventually pays the cost of cooperation c if – according to its strategy – it chooses TFT but it has no chance to receive a benefit b from another player.

Players also choose their cognitive style: they can play either in an intuitive or a deliberative way. Under intuition they play without knowing the type of prisoners' dilemma they are currently playing; instead, under deliberation, they have this information and can condition their action on the type of game played. However, deliberation comes at a cost, k, which is stochastically sampled decision-by-decision from a generic distribution $g(k) \in [0, K]$ with a cumulative distribution $G(k)$ in $[0, 1]$. We assume that $K \geq \min\{c, (b - c)\}$. This assumption guarantees that for any value of $p \in (0, 1)$ fully deliberating strategies - strategies according to which the agent decides to pay the cost of deliberation for any possible randomly sampled cost of deliberation - cannot give rise to absorbing states.

Figure 7 represents the entire game played by two randomly matched players staying in the same location.

Given this setup, we can fully describe agent a's strategy via a vector $\sigma_{a,t} = (\ell_{a,t}, i_{a,t}, o_{a,t}, r_{a,t}, k_{a,t})$ where:

- $\ell_{a,t} \in \{1, \ldots, L\}$ denotes the location chosen by agent a at time t;
- $i_{a,t} \in [0, 1]$ is the probability with which agent a plays TFT under

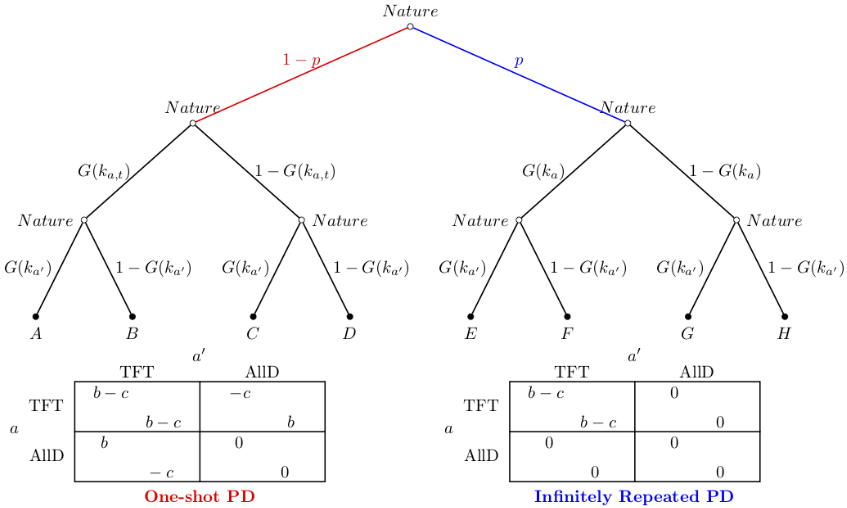


Figure 7: The entire game played. Extensive-form representation of the entire game played by two randomly matched agents. First, Nature randomly chooses whether the game played is anonymous and one-shot or infinitely repeated. Then, the actual costs of deliberation of player a and player a' are randomly sampled, and, consequently, the two agents deliberate or don't. Finally, the two agents play the prisoners' dilemma. Agent a's information sets are: {A, B}, {E, F}, and {C, D, G, H}. Agent a''s information sets are: {A, C}, {E, G}, and {B, D, F, H}.

intuition at time t;

- $o_{a,t} \in [0, 1]$ and $r_{a,t} \in [0, 1]$ denote the probabilities with which agent a plays TFT under deliberation at time t conditional on the game being, respectively, one-shot and anonymous or infinitely repeated;
- $k_{a,t} \in [0, K]$ is the maximum cost that agent a is willing to pay to deliberate at time t.

Following the terminology adopted in Bear and David G Rand (2016), we say that a strategy is intuitive if according to it an agent never deliberates, while we say that a strategy is dual process if it prescribes to deliberate with positive probability. Moreover, a strategy will be considered cooperative [defective] if it prescribes playing TFT [AllD] under intuition. Consequently, strategies of type (ℓ, 0, 0, 1, 0), (ℓ, 0, 0, 1, k), (ℓ, 1, 0, 1, k), and (ℓ, 1, o, r, 0) will be referred to as, respectively, intuitive defection strategy, dual process defection strategy, dual process cooperation strategy, and intuitive cooperation strategy.

In each period of time t, the state of the system can be described via a matrix

$$S_t = (\boldsymbol{\ell}_t, \mathbf{i}_t, \mathbf{o}_t, \mathbf{r}_t, \mathbf{k}_t)$$

where $\boldsymbol{\ell}_t = (\ell_{1,t}, \ldots, \ell_{A,t})^T$, $\mathbf{i}_t = (i_{1,t}, \ldots, i_{A,t})^T$, $\mathbf{o}_t = (o_{1,t}, \ldots, o_{A,t})^T$, $\mathbf{r}_t = (r_{1,t}, \ldots, r_{A,t})^T$, and $\mathbf{k}_t = (k_{1,t}, \ldots, k_{A,t})^T$.

The system is characterized by synchronous learning. More precisely, at the end of each period of time t each agent a has a fixed probability $\rho_a \in (0,1)$ to be given a revision opportunity. We assume $\rho_a = \rho_{a,\sigma} + \rho_{a,\gamma}$ where $\rho_{a,\sigma} \geq 0$ is the probability with which player a is given a full-strategy revision opportunity while $\rho_{a,\gamma} \geq 0$ is the probability with which player a is given a game-play revision opportunity with which the agent can revise all its strategy components except the location-choice. In each case, if given a revision opportunity the agent revises its strategy choice so as to maximize its payoff given the current state of the system S_t (myopic best reply).

In Figure 8 we report an example of a generic state of the system and strategy updating.

We assume that when revising their strategy choice agents can make a mistake with probability $\varepsilon \geq 0$. If an agent makes a mistake then it selects its new strategy completely at random (uniform mistakes). In the following, we will say that the system evolves according to unperturbed dynamics if $\varepsilon = 0$, while we will say that the system evolves according to perturbed dynamics if $\varepsilon > 0$.

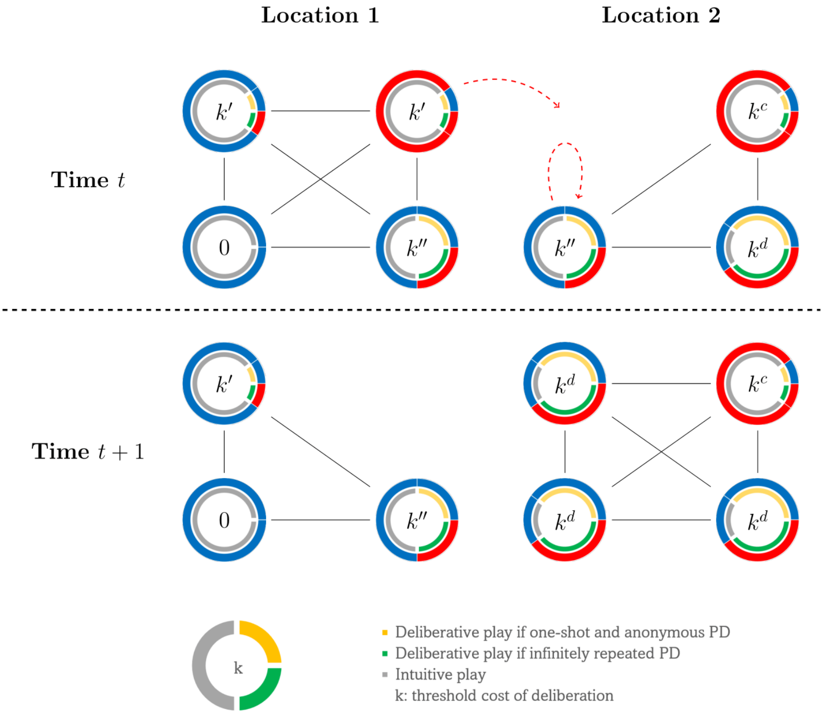


Figure 8: A generic state of the system. Graphical representation of a generic state of the system and its evolution over time (case $A = 4$, $L = 2$). Each agent is described via the strategy it adopts. For example, at time t in location 1 the agent on the bottom right position adopts a dual process defection strategy as it plays AllD both under intuition and under deliberation if the PD is one-shot, while it plays TFT under deliberation if the PD is infinitely repeated. Moreover, it deliberates with positive probability as its threshold cost of deliberation is $k'' > 0$. At the end of time t, two agents are given a revision opportunity: they both choose location 2 (as it is the one with more cooperative agents) and adopt a dual process defection strategy.

3.3 Stochastic Stability with a Single Location

Before providing the main results of the paper obtained with the setup presented in Section 3.2 and characterized by multiple locations ($L \geq 2$) we illustrate two examples of stochastic stability analysis under the assumption that there is a single location $L = 1$ which implies that all agents must stay in the same location. This version of the model coincides with the one in Jagau and Veelen (2017) with the additional assumption that K is such that fully deliberating absorbing states, i.e. absorbing states in which agents choose a threshold cost of deliberation such that they always deliberate, do not exist.

By employing the results in Jagau and Veelen (2017) we know that the single location setup is characterized by the following absorbing states:

- $(\mathbf{0}, \mathbf{0}, \mathbf{1}, \mathbf{0})$ if $p \in (0, 1)$;
- $(\mathbf{0}, \mathbf{0}, \mathbf{1}, \mathbf{k^d})$ if $p \in \left(0, \frac{c}{c+G(k^d)(b-c)}\right)$ with every k^d such that $k^d = p(b-c)G(k^d)$ if any;
- $(\mathbf{1}, \mathbf{0}, \mathbf{1}, \mathbf{k^c})$ if $p \in \left(0, \frac{c}{b}\right)$ with the unique $k^c = (1-p)c$.

The two examples of stochastic stability analysis provided in the following show that in the absence of an endogenous interaction structure (i.e, if $L = 1$) stochastic stability considerations crucially depend on the actual distribution of deliberation costs $G(\cdot)$; moreover, dual process cooperation strategies do not necessarily (and usually do not) prevail in the entire parameter space in which they give rise to absorbing states under unperturbed dynamics.

These findings are significantly different from the ones we provide in Section 3.4. This suggests that an endogenous interaction structure plays a crucial role in the evolution of dual process cooperation and provides an additional reason why our setup is of interest.

3.3.1 Example 1

Consider the case $L = 1$, $b = 2$, $c = 1$, $K = 1$ and assume that the distribution of deliberation costs is

$$G(k) = \begin{cases} 0 & \text{if } k < 0 \\ k & \text{if } 0 \leq k \leq K \\ 1 & \text{if } k > K \end{cases} \tag{3.1}$$

Then, the absorbing states of the system are:

- The intuitive defection absorbing state $(\mathbf{0}, \mathbf{0}, \mathbf{1}, \mathbf{0})$ if $p \in (0, 1)$;
- The dual process cooperation absorbing state $(\mathbf{1}, \mathbf{0}, \mathbf{1}, \mathbf{k^c})$ if $p \in \big(0.5, 1\big)$ with the unique $k^c = (1 - p)c$.

There are no other absorbing states. In Figure 9 we illustrate the optimal threshold costs of deliberation for strategies prescribing to play AllD under intuition and strategies prescribing to play TFT.

Consider the system under perturbed dynamics ($\varepsilon > 0$) with $p \in \big(0.5, 1\big)$ and consider the partition of the absorbing state of the system into $\Omega = \{(\mathbf{1}, \mathbf{0}, \mathbf{1}, \mathbf{k^c})\}$ and $\Omega^{-1} = \{(\mathbf{0}, \mathbf{0}, \mathbf{1}, \mathbf{0})\}$.

It can be showed[1] that if the population is large enough then one mistake can never lead with a positive probability the system into the basin of attraction of Ω^{-1} starting from the dual process cooperation absorbing state $(\mathbf{1}, \mathbf{0}, \mathbf{1}, \mathbf{k^c})$. But then, if the population is large enough the radius of the basin of attraction of Ω is strictly larger than one, formally $R(\Omega) > 1$.

Assume that at time t the system is in the intuitive defection absorbing state $(\mathbf{0}, \mathbf{0}, \mathbf{1}, \mathbf{0})$. In such a state, every agent plays AllD under intuition and never deliberates and so $(\tau_a^o(S_t), \tau_a^r(S_t)) = (0, 0)$ for every agent $a \in \mathcal{A}$. Assume now that at the end of time t one agent, say a', adopts by mistake the dual process cooperation strategy $(1, 0, 1, k^c)$ with $k^c = (1 - p)c$. Then, at the beginning of time $t + 1$ it will be $\big(\tau_a^o(S_{t+1}), \tau_a^r(S_{t+1})\big) = \big(\frac{1-G(k^c)}{A-1}, \frac{1}{A-1}\big)$ for every agent $a \neq a'$. But then,

---

[1] The argument is similar to the one employed in the proof of Theorem 5 and for this reason it is omitted here. In general, every step that is not explicitly proved in this section relies on one or more results reported in Appendix E.

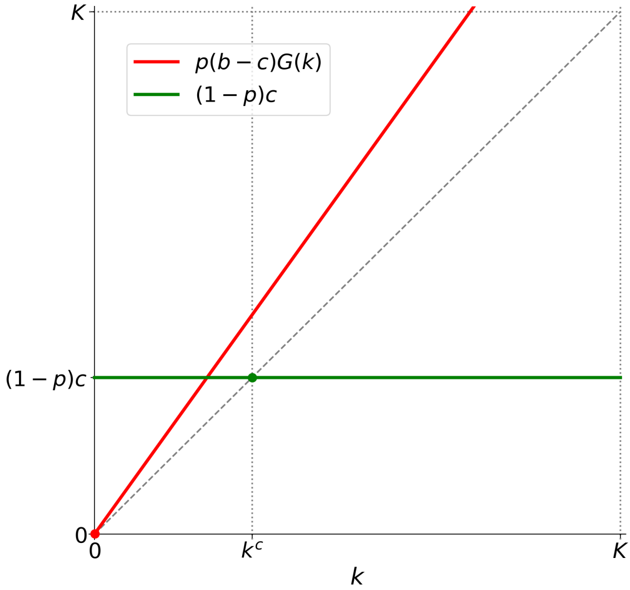


Figure 9: Optimal threshold costs of deliberation in Example 1. Myopic best reply thresholds for strategies prescribing to play AllD under intuition and for strategies prescribing to play TFT under intuition if the distribution of deliberation costs is the one in Equation (3.1)

it can be shown that for every agent $a \neq a'$ the best reply to the current state of the system belongs to the set

$$BR_a(S_{t+1}) \subseteq \big\{(0,0,1,k'),(1,0,1,k^c)\big\}$$

with $k' = p(b-c)\tau_a^r(S_{t+1}) > 0$ and $k^c = (1-p)c$ and, so, they will be willing to adopt either the optimal dual process cooperation strategy or they will be willing to adopt a dual process defection strategy:

- In the first scenario with positive probability at the end of time $t+1$ every agent $a \neq a'$ will be given a revision opportunity and will adopt strategy $(1,0,1,k^c)$ and consequently the system will have reached the dual process cooperation absorbing state.

- In the second scenario with positive probability at the end of time $t+1$ every agent $a \neq a'$ will be given a revision opportunity and will adopt strategy $(0,0,1,k')$. Then, at the beginning of time $t+2$ it will be $\big(\tau^o_a(S_{t+2}), \tau^r_a(S_{t+2})\big) = \big(\frac{1-G(k^c)}{A-1}, \frac{(A-2)G(k')+1}{A-1}\big)$ for every agent $a \neq a'$. But then, for every agent $a \neq a'$ the best reply to the current state of the system will belong to the set

  $$BR_a(S_{t+2}) \subseteq \big\{(0,0,1,k''),(1,0,1,k^c)\big\}$$

  with $k'' = p(b-c)\tau^r_a(S_{t+2}) > k'$ and $k^c = (1-p)c$ and, so, they will be willing to adopt either the optimal dual process cooperation strategy or they will be willing to adopt a dual process defection strategy with a higher threshold cost of deliberation than their current one. This process will continue with positive probability until all agents $a \neq a'$ will be given a revision opportunity and will find it optimal to adopt a dual process cooperation strategy and, so, until the system will reach the dual process cooperation absorbing state.

We have argued that for every $p \in (0.5,1)$ a single mistake can lead with positive probability the system into the basin of attraction of Ω starting from the intuitive defection state. But then the coradius of the basin of attraction of Ω is equal to one, formally $CR(\Omega) = 1$.

Hence, if the population is large enough then for every $p \in \big(0.5, 1\big)$ we have $R(\Omega) > 1 = CR(\Omega)$ and so the dual process cooperation absorbing state is stochastically stable. Therefore, in this case, the dual process cooperation absorbing state is stochastically stable in the entire parameter space in which it is an absorbing state.

3.3.2 Example 2

Consider now the case $L = 1$, $b = 4$, $c = 1$, $K = 3$ with a large population of agents. Assume that the distribution of deliberation costs is

$$G(k) = \begin{cases} 0 & \text{if } k < K \\ 1 & \text{if } k \geq K \end{cases} \tag{3.2}$$

Then, the absorbing states of the system are:

- The intuitive defection absorbing state $(\mathbf{0}, \mathbf{0}, \mathbf{1}, \mathbf{0})$ if $p \in (0, 1)$;
- The dual process cooperation absorbing state $(\mathbf{1}, \mathbf{0}, \mathbf{1}, \mathbf{k^c})$ if $p \in \big(0.25, 1\big)$ with the unique $k^c = (1-p)c$. Technically given the distribution of deliberation costs this is an intuitive cooperation absorbing state.

There are no other absorbing states. In Figure 10 we illustrate the optimal threshold costs of deliberation for strategies prescribing to play AllD under intuition and strategies prescribing to play TFT under intuition under the assumptions made on the distribution of deliberation costs. Note that for any threshold cost of deliberation strictly lower than K the probability of deliberation is null and, so, any strategy with a threshold cost of deliberation below K implies never deliberating and always playing under intuition.

Consider the system under perturbed dynamics ($\varepsilon > 0$) with $p \in \big(0.25, 1\big)$ and consider the partition of the absorbing state of the system into $\Omega = \{(\mathbf{1}, \mathbf{0}, \mathbf{1}, \mathbf{k^c})\}$ and $\Omega^{-1} = \{(\mathbf{0}, \mathbf{0}, \mathbf{1}, \mathbf{0})\}$.

In this scenario, we need to compute the radius and the coradius of Ω as it is not true that one mistake can lead with positive probability the system into the basin of attraction of Ω starting from the intuitive defection absorbing state.

Assume that at the end of time t the system is in the dual process cooperation absorbing state $(\mathbf{1}, \mathbf{0}, \mathbf{1}, \mathbf{k^c})$. Given the distribution of deliberation costs in Equation (3.2), in such state it must be $(\tau_a^o(S_t), \tau_a^r(S_t)) = (1, 1)$ for every agent $a \in \mathcal{A}$. Assume that μA agents adopt by mistake the intuitive defection strategy $(0, 0, 1, 0)$. Then, given that the population is large at time $t+1$ it will be $(\tau_a^o(S_{t+1}), \tau_a^r(S_{t+1})) \sim (1-\mu, 1-\mu)$ for every agent $a \in \mathcal{A}$. The minimum number of agents making a mistake required to lead with positive probability the system into the intuitive defection absorbing state is the one such that Equation (E.3) holds with equality and so

$$1 - \mu = \frac{(1-p)c}{p(b-c)} \Leftrightarrow \mu = \frac{pb - c}{p(b-c)}$$

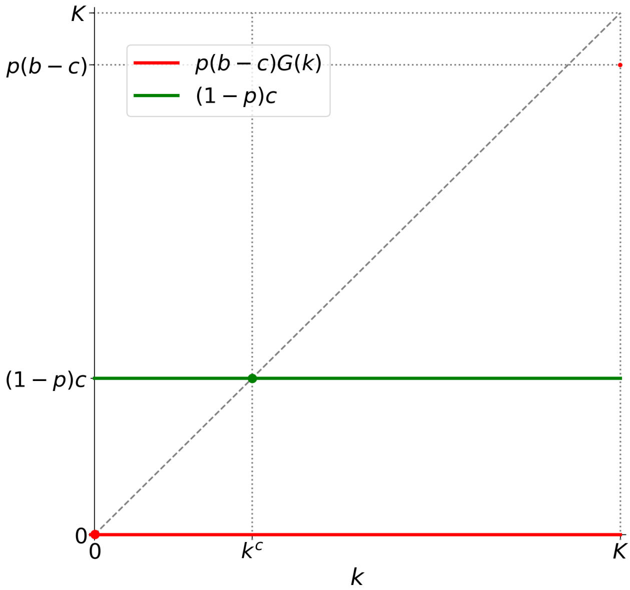


Figure 10: Optimal threshold costs of deliberation in Example 2. Myopic best reply thresholds for strategies prescribing to play AllD under intuition and for strategies prescribing to play TFT under intuition if the distribution of deliberation costs is the one in Equation (3.2)

But then, the radius of the basin of attraction of Ω is $R(\Omega) = \frac{pb-c}{p(b-c)}A$.

Assume that at the end of time t the system is in the intuitive defection absorbing state $(\mathbf{0}, \mathbf{0}, \mathbf{1}, \mathbf{0})$. In such state it must be $(\tau_a^o(S_t), \tau_a^r(S_t)) = (0, 0)$ for every agent $a \in \mathcal{A}$. Assume that μA agents adopt by mistake the dual process cooperation strategy $(1, 0, 1, k^c)$. Then, given that the population is large at time $t+1$ it will be $(\tau_a^o(S_{t+1}), \tau_a^r(S_{t+1})) \sim (\mu, \mu)$ for every agent $a \in \mathcal{A}$. The minimum number of agents making a mistake required to lead with positive probability the system into the dual process cooperation absorbing state is the one such that Equation (E.3) holds with equality and so

$$\mu = \frac{(1-p)c}{p(b-c)}$$

But then, the coradius of the basin of attraction of Ω is $CR(\Omega) = \frac{(1-p)c}{p(b-c)}A$.

The radius of Ω is larger than its coradius if

$$\frac{pb-c}{p(b-c)}A > \frac{(1-p)c}{p(b-c)}A \Leftrightarrow p > \frac{2c}{b+c}$$

which is a strictly more demanding condition than $p > \frac{c}{b}$.

3.4 Results

In this section, the two main results of the paper are reported: the first one lists the absorbing sets of the system under unperturbed dynamics, while the second one identifies the stochastically stable states of the system under perturbed dynamics.

The analysis of both absorbing and stochastically stable states of the system is performed focusing on pure strategies. Such simplification is without loss of generality given our focus on long-run outcomes and the fact that absorbing states in mixed strategies are never stochastically stable (H. P. Young, 1993).

3.4.1 Unperturbed Dynamics

We consider first the system under unperturbed dynamics and, so, we consider the case in which if an agent is given a revision opportunity then it myopically best replies to the current state of the system with unitary probability. Under such dynamics, the main interest lies in determining the absorbing sets of the system. We do this in Theorem 4.

Theorem 4. *Consider the system under unperturbed dynamics ($\varepsilon = 0$). The absorbing sets of the system are:*

- *($\mathcal{L}$, $\mathbf{0}$, $\mathbf{0}$, $\mathbf{1}$, $\mathbf{0}$) if $p \in (0,1)$;*
- *($\boldsymbol{\ell}^*$, $\mathbf{0}$, $\mathbf{0}$, $\mathbf{1}$, $\mathbf{k^d}$) if $p \in \left(0, \frac{c}{c+G(k^d)(b-c)}\right)$ with every k^d such that $k^d = p(b-c)G(k^d)$ if any;*
- *($\boldsymbol{\ell}^*$, $\mathbf{1}$, $\mathbf{0}$, $\mathbf{1}$, $\mathbf{k^c}$) if $p \in \left(\frac{c}{b}, 1\right)$ with the unique $k^c = (1-p)c$.*

There are no other absorbing sets.

In Figure 11 we provide a graphical representation of the absorbing sets of the system listed in Theorem 4, while the proof of Theorem 4 is reported in Appendix E and it is conveniently divided into four lemmas.

First, we show that the intuitive defection set ($\mathcal{L}$, $\mathbf{0}$, $\mathbf{0}$, $\mathbf{1}$, $\mathbf{0}$) is absorbing in the entire parameter space. This absorbing set contains all the states[2] in which all players adopt the intuitive defection strategy and select a location at random. More precisely, all players are indifferent between staying in their current location, moving to another inhabited location, and moving into an empty location as all these alternatives provide them a null payoff. Despite receiving a null payoff every agent myopically best replies to the current state of the system by keeping an intuitive defection strategy as every alternative strategy provides a negative expected payoff. This is the case because every other strategy implies paying with positive probability the cost of cooperation – and eventually a cost of deliberation – without the possibility of receiving a benefit of cooperation as in these states the rate of cooperation is null in both one-shot and anonymous and infinitely repeated prisoners' dilemmas.

This first result is consistent with the findings in Bear and David G Rand (2016) and Jagau and Veelen (2017) who also find that the intuitive defection strategy can be sustained as an equilibrium in the entire parameter space. In addition, according to our model, intuitive defection is associated with high players' mobility. More precisely, every distribution of agents ranging from full concentration in one location to complete dispersion over the various locations is feasible and this is the case because, in an intuitive defection environment, social interactions do not provide any benefit to agents.

Second, we show that dual process defection states ($\boldsymbol{\ell}^*$, $\mathbf{0}$, $\mathbf{0}$, $\mathbf{1}$, $\mathbf{k^d}$) with $\mathbf{k^d} = (k^d, \ldots, k^d)$ and $k^d = p(b-c)G(k^d)$ are absorbing states if $p \in \left(0, \frac{c}{c+G(k^d)(b-c)}\right)$. In a dual process defection absorbing state, all agents play AllD under intuition, deliberate with positive probability, and if they do so then they play optimally. Moreover, all players stay in the same location. In such states, the rate of cooperation is null in one-

---

[2] The intuitive defection set is composed by A^L states where A^L corresponds to the number of ways in which A agents can be distributed over L locations.

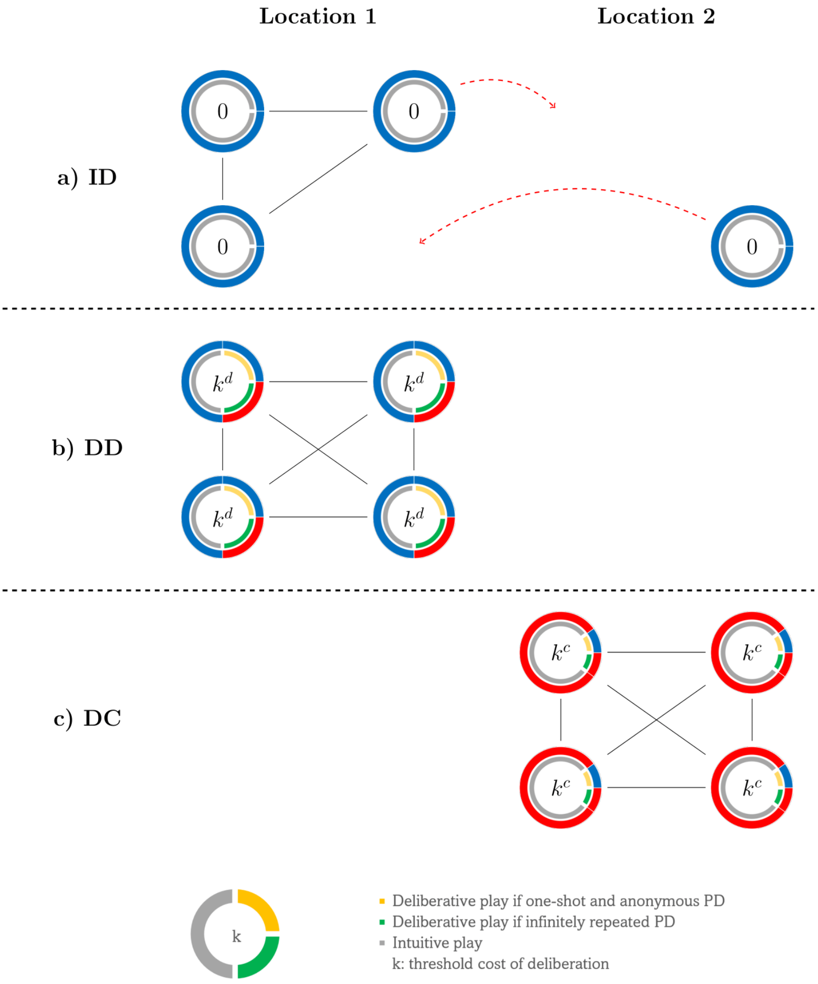


Figure 11: Absorbing states according to Theorem 4. Graphical representation of the absorbing sets of the system (case $A = 4$, $L = 2$). In the intuitive defection absorbing set, all agents adopt the intuitive defection strategy and are indifferent between staying in one location or another. In a dual process defection [cooperation] absorbing state all agents adopt a dual process defection [cooperation] strategy with the same optimal threshold cost of deliberation and they all stay in the same location.

shot and anonymous interactions while it is equal to $G(k^d)$ in the case of infinitely repeated prisoners' dilemmas.

It is worth underlying that dual process defection absorbing states do not necessarily exist and if they do there may be multiple types of dual process defection absorbing states differing not only in the equilibrium location choice but also in the equilibrium threshold cost of deliberation. This is the case because depending on the actual distribution of deliberation costs $G(\cdot)$ there may be none, one, or multiple $k^d : k^d = p(b-c)G(k^d)$. In general, there are as many dual process defection absorbing states as the number of locations times the number of threshold costs of deliberation such that $k^d = p(b-c)G(k^d)$. In case there exist multiple types of dual process defection absorbing states the various types can be ordered from the least to the most socially desirable by simply comparing the equilibrium threshold cost of deliberation: the higher k^d the higher the rate of cooperation in infinitely repeated interactions and the higher agents' expected payoff.

Third, we show that each dual process cooperation state ($\boldsymbol{\ell}^*$, $\mathbf{1}$, $\mathbf{0}$, $\mathbf{1}$, $\mathbf{k^c}$) with $\mathbf{k^c} = (k^c, \ldots, k^c)$ and $k^c = (1-p)c$ is an absorbing state if $p \in (\frac{c}{b}, 1)$. In each of the L dual process cooperation absorbing states all agents play TFT under intuition, deliberate with positive probability, and if they do so then they play the dominant strategy of the game they are currently facing. Moreover, all players stay in the same location. In such states, the rate of cooperation is equal to $1-G(k^c)$ in one-shot and anonymous interactions while it is equal to 1 in the case of infinitely repeated prisoners' dilemmas. These are the highest cooperation rates that may characterize an absorbing set of the system. Consequently, dual process cooperation absorbing states are the ones in which agents achieve the highest expected payoff and, thus, are socially optimal.

Also, this result is consistent with the findings in Bear and David G Rand (2016) and Jagau and Veelen (2017): both the parameter space in which such states are absorbing and the condition on the optimal threshold cost of deliberation coincide. Moreover, our finding that in dual process cooperation absorbing states all agents stay in the same location is a natural extension of the results in Bear and David G Rand (2016) and

Jagau and Veelen (2017) given our location choice setting.

Finally, we formally prove that the system under unperturbed dynamics does not have other absorbing sets than the intuitive defection set, (eventually) the set of dual process defection states, and the set of dual process cooperation states. We do this by showing that starting from a generic state the system will always reach – with a positive probability – one of the aforementioned absorbing sets. But then, we can conclude that there are no absorbing sets other than the ones listed in Theorem 4.

The results in Theorem 4 give rise to a problem of multiplicity of equilibria: for any possible probability that the prisoners' dilemma is infinitely repeated, p, at least two kinds of equilibria exist and, so, we do not have a clear prediction regarding the behavior of the system despite it being of particular interest given that the different kinds of absorbing sets can always be ranked in terms of social desirability.

3.4.2 Perturbed Dynamics

We consider now the system under perturbed dynamics and, so, we assume that if given a revision opportunity an agent can make a mistake with positive probability; if this happens then the agent selects a strategy uniformly at random rather than following myopic best reply. Under perturbed dynamics the system becomes ergodic and, thus, it is no more characterized by absorbing sets: with positive probability, the system may leave from any set of states and reach any other state.

Within this setting, we aim at solving the problem of multiplicity of equilibria arising from Theorem 4 by identifying the stochastically stable states of the system. These are the states in which the system is expected to spend a significant amount of time if the system evolves as $t \to \infty$ and $\varepsilon \to 0$.

Stochastic stability analysis can be performed only if the number of states of the system is finite. However, the model considered so far is characterized by an infinite state space as each agent's threshold cost of deliberation is a continuous variable. We then assume that agents choose their threshold cost of deliberation from the following discrete set:

$$\mathcal{K} = \{0, \delta, 2\delta, \dots, 1-2\delta, 1-\delta, 1\}$$

with $\delta \in \{\frac{1}{n} : n \in \mathbb{N}\}$. We require that the set of discrete thresholds cost of deliberation and, more precisely, δ satisfies the two following conditions. First, δ must be such that $k^c, k^d \in \mathcal{K}$ for $k^c = (1-p)c$ and for any $k^d : k^d = p(b-c)G(k^d)$ given $G(\cdot)$. This assumption guarantees that all the absorbing sets in Theorem 4 are still attainable in the discrete version of the model and, consequently, that the discretization performed does not drive the results reported in Theorem 5. Second, δ must be small enough so that if in the non-discrete version of the model an agent's mistake generates a profitable deviation for another player then this must hold in the discrete version of the model too; in addition, such profitable deviation must be attainable. This assumption guarantees that the discretization of the model does not affect the minimum number of mistakes required to lead with positive probability the system from a state S to state S' for any possible couple of states (S, S'). A δ satisfying the previous conditions always exists. More precisely, any

$$\delta = \frac{1}{\alpha \text{ L.C.M.}\big(k^c, \{k^d\}_{\forall k^d : k^d = p(b-c)G(k^d)}\big)}$$

where L.C.M. denotes the least common multiple and α is a large enough constant.

In Theorem 5 we report our main result under perturbed dynamics.

Theorem 5. *Consider the system under perturbed dynamics ($\varepsilon > 0$). If $p \in \left(\frac{c}{b}, 1\right)$, then all the stochastically stable states of the system are contained in the set of dual process cooperation states $\{(\boldsymbol{\ell}^*, \mathbf{1}, \mathbf{0}, \mathbf{1}, \mathbf{k^c})\}_{\boldsymbol{\ell}^*=\mathbf{1}}^{\mathbf{L}}$ with $\mathbf{k^c} = (k^c, \dots, k^c)$ and $k^c = (1-p)c$.*

We report the proof of Theorem 5 in Appendix F. In summary, we partition the set absorbing sets characterizing the system under unperturbed dynamics (Theorem 4) into two sets: the set of dual process cooperation states Ω, and its complementary set, Ω^{-1}, containing the intuitive defection set and dual process defection absorbing states. We then show that if $p \in \left(\frac{c}{b}, 1\right)$ one mistake may lead with positive probability the system into the basin of attraction of Ω starting from any state belonging to Ω^{-1} and, consequently, the coradius of the basin of attraction of Ω is equal to one, formally $CR(\Omega) = 1$. Moreover, we show that if the population is large enough then one mistake is never enough to make the system leave the basin of attraction of Ω starting from a dual process cooperation state and, therefore, the radius of the basin of attraction of Ω is strictly larger than one, formally $R(\Omega) > 1$. By combining these two results we can conclude that if $p \in \left(\frac{c}{b}, 1\right)$ then

$$R(\Omega) > 1 = CR(\Omega)$$

and, thus, all the stochastically states of the system are contained in Ω. But then, given that Ω contains only dual process cooperation states it must be that if $p \in \left(\frac{c}{b}, 1\right)$ dual process cooperation states are the only stochastically stable states.

According to Theorem 5 dual process cooperation states are stochastically stable in the entire parameter space in which they are absorbing states under unperturbed dynamics. This result holds for any distribution of deliberation costs $G(\cdot)$. Instead, in the case with a single location – as the two examples in Section 3.3 show – dual process cooperation states usually evolve in a subset of the parameter space in which they are absorbing states; moreover, the parameter space in which they evolve crucially depends on the actual distribution of deliberation costs. Therefore, the presence of an endogenous interaction structure does not only guarantee that dual process cooperation is favored by evolution for any value of p such that it gives rise to absorbing states but it also does this independently of the actual distribution of deliberation costs.

3.5 Strategy Revision with the Moran Process

To check whether our results crucially depend on the revision protocol adopted in this section we consider the model with $L \geq 1$ locations and we assume that the system evolves according to the Moran process with uniform mistakes. More precisely, at the end of each period of time one agent, say a', is randomly selected to revise its strategy. With probability $\varepsilon > 0$ agent a' adopts a new strategy uniformly at random while with probability $1 - \varepsilon > 0$ it copies the strategy of another agent $a \neq a'$. In particular, agent a' copies the strategy currently adopted by agent a with the probability

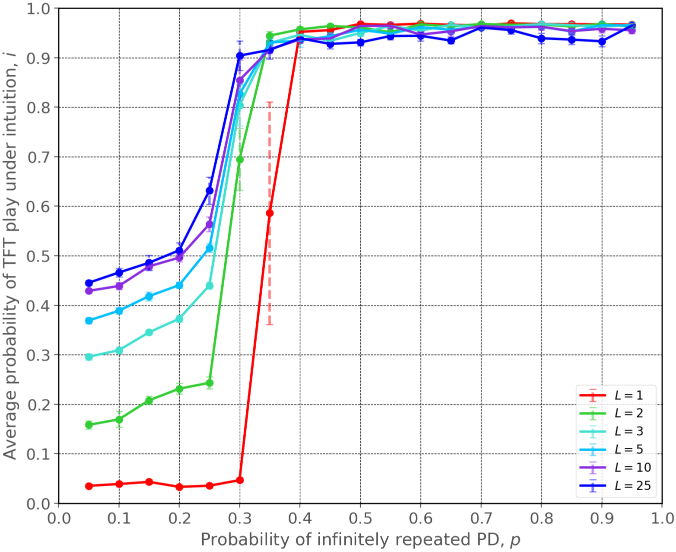
$$p_{a',a}(S_t) = \frac{e^{i\pi_a(S_t)}}{\sum_{\alpha \neq a'} e^{i\pi_\alpha(S_t)}}$$

where $i > 0$ is the intensity of selection and $\pi_a(S_t)$ is the expected payoff of agent a given the state of the system at time t.

We simulate the model setting the following parameter values: $b = 4$, $c = 1$, $A = 50$, $i = 10$, $\varepsilon = 0.05$, $g(\cdot) \sim \mathcal{U}(0, 3)$. We also assume that the probability that the game is an infinitely repeated prisoners' dilemma takes values in $p \in \{0.05, 0.10, 0.15, \ldots, 0.90, 0.95\}$ while the number of locations is such that $L \in \{1, 2, 3, 5, 10, 25\}$. For each (p, L) combination we perform 5 independent agent-based simulations of the model. Each simulation begins from a randomly selected initial state and consists of 10^5 iterations.

In Figure 12 we report the simulation results. More precisely, we plot the average probability to play Tit-For-Tat under intuition and the average threshold cost of deliberation as a function of p. Such average values are computed over all the simulations characterized by the same p and L excluding the first 10^4 iterations (the first 10% iterations) to neutralize initialization bias.

Overall the figure indicates that there is a qualitative difference in the behavior of the system as the number of locations changes from one to more than one. In particular, in the case with more than one location lower values of p sustain dual process cooperation as the main model with myopically best replying agents would predict given the assumptions made. In addition, we observe that in the case with multiple locations even when dual process cooperation cannot be sustained – as it does not give rise to absorbing states of the system – the average behavior of the system is inconsistent with an intuitive defection scenario. In fact, in cases with more than one location even for low values of p a significant



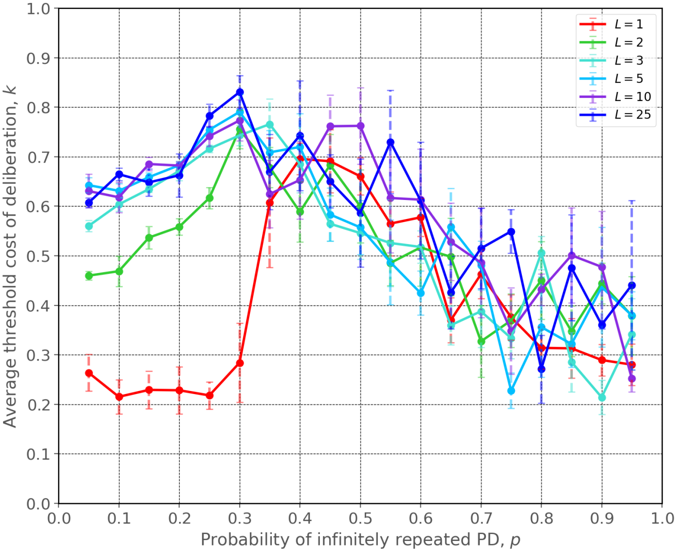


Figure 12: Simulations results. Average probability to play TFT under intuition (above) and average threshold cost of deliberation (below) and their standard errors by the number of locations L.

amount of agents play TFT under intuition and agents tend to deliberate frequently. Instead, in the case $L = 1$ for low values of p we observe most agents playing AllD under intuition and a limited use of deliberation which is in line with an intuitive defection scenario (given the presence of mistakes). We interpret this as the result of an indirect revision protocol combined with mistakes. More precisely, if there is more than one location the following series of events may occur: a location becomes empty, an agent makes a mistake and by doing so it moves into an empty location and it adopts a 'cooperative' strategy, the agent is then copied by other agents and a cooperative cluster emerges. Such a cooperative cluster may survive even for relatively long periods of time as the following considerations hold: if a cooperative cluster emerges then the agents belonging to the cluster obtain a higher payoff than agents staying in locations in which intuitive defection has developed and so such agents will be the ones with the highest probability to be copied by others; moreover, cooperative clusters can be disrupted only through mistakes and, in particular, only if an agent makes a mistake and adopts a 'defective' strategy in the location where the cooperative cluster has emerged. Note that the higher the number of locations the higher the probability that a cooperative cluster emerges and that such a cluster lasts for a significant amount of time. In fact, as the number of locations increases (i) the probability that at least one location is empty increases, (ii) the probability that an agent who has made a mistake and has adopted a cooperative strategy in a previously empty location is selected to be copied is unaffected by the number of locations ceteris paribus, and (iii) the probability that an agent adopts by mistake a defective strategy in the location where the cooperative cluster has emerged decreases. Note also that this different behavior of the system with multiple locations would not be observed if agents were myopically best responders. In fact, under myopic best reply, rather than having the chance of being copied by others, an agent adopting a cooperative strategy in an empty location would make it profitable for all the other agents to adopt a dual process defection strategy in the location where the mistake was made, and so no cooperative cluster would emerge for low values of p.

3.6 Conclusion

In this paper, we have analyzed a model of co-evolution of cooperation and cognition with an endogenous interaction structure.

Following Bear and David G Rand (2016) and Jagau and Veelen (2017), we have considered a model in which agents are randomly matched to play a prisoners' dilemma which can be either one-shot and anonymous or infinitely repeated. In addition, agents are uninformed about the type of interaction they are facing, and to retrieve this information they need to engage in costly deliberation. The presence of different types of games together with incomplete information generates a trade-off between intuitively defecting – which is the dominant strategy in case the interaction is one-shot and anonymous – and intuitively cooperating – which is optimal in infinitely repeated interactions. In the evaluation of such a trade-off, agents must also consider that different strategies imply different optimal levels of deliberation.

We have enriched this setting by introducing a set of locations and by allowing agents to choose the location in which they want to stay. Such a decision determines the set of agents with which an agent interacts as only agents staying in the same location can be matched to play a prisoners' dilemma. This gives rise to an endogenous interaction structure.

We have shown that the model presents at most three types of absorbing sets: (i) an intuitive defection set in which agents play Always Defect under intuition, never deliberate, and are indifferent with respect to their location choice, (ii) dual process cooperation states in which agents play Tit-For-Tat (cooperate) under intuition, deliberate with positive probability and all agents are concentrated in the same location, and (iii) depending on the actual distribution of deliberation costs none, one or more kinds of dual process defection absorbing states in which all agents defect under intuition, deliberate with positive probability and stay in the same location. This first result represents a natural extension of the findings in Bear and David G Rand (2016) and Jagau and Veelen (2017) in a location-choice setting.

We have then performed equilibrium selection via stochastic stability analysis and we have found that dual process cooperation states are the only stochastically stable states in the entire parameter space in which they are absorbing states. This in turn implies that in our model intuitive cooperation is favored by evolution in a larger parameter space than in Bear and David G Rand (2016). This result implies that an endogenous interaction structure promotes cooperation.

Chapter 4

A Dual Process Model for the Evolution of Fair Splits and Harsh Rejections in the Ultimatum Game

4.1 Introduction

Since the first experimental work dealing with the Ultimatum Game (Güth, Schmittberger, and Schwarze, 1982) there has been great interest in understanding why subjects' behavior differs so markedly from standard game theoretical arguments. In fact, while the latter predicts almost null offers and full acceptance of any positive offer, experimental evidence points to 'fair' offers – with the average offer being frequently between 40% and 50% of the amount to split – and frequent rejection of 'unfair' offers (Güth and M. G. Kocher, 2014).

However, the disagreement between theoretical predictions and experimental evidence might not be as strong as it might seem at a first glance. In fact, it is essential to remember that theoretical predictions are derived under many assumptions, and their validity is restricted to the situations in which such assumptions hold. In the case of an Ultimatum

Game, for example, players must believe that the game is anonymous and one-shot. Still, they must also be convinced that the receiver has no bargaining power other than being able to reject the offer. But what if this is not the case?

With this reasoning in mind, I develop a model of the evolution of Ultimatum Game bargaining in which proposers and receivers interact over time in either Ultimatum Games or simplified Bargaining Games. Hence, agents face different kinds of interactions. Following dual process theories of cognition (Evans and Stanovich, 2013) and their recent applications to the case of cooperation (Bear and David G Rand, 2016; Jagau and Veelen, 2017), I assume that although agents know the probability with which they are playing one game or another, they do not know the actual game they are playing in a given interaction. To obtain this information, agents need to pay a deliberation cost which is randomly sampled from a known and fixed distribution. Both proposers and receivers can choose their cognitive effort, and they do so by selecting a threshold cost of deliberation, which represents the maximum cost they are willing to pay to deliberate. In a given interaction, if an agent pays the deliberation cost, then it deliberates, and it gets to know the type of game it is currently facing and, consequently, it can condition its behavior to the game played; while if the agent does not deliberate, then it must play without knowing the actual game. Both proposers and receivers update their strategies according to reinforcement learning. More precisely, each agent is characterized by sets of propensities, with each propensity expressing the agent's attitude to choose a given action, such as selecting a given threshold cost of deliberation, making a given offer, and accepting or rejecting an offer. After every interaction, each agent revises its propensities according to Roth-Erev reinforcement (Erev and Roth, 1998).

To my knowledge, this is the first model analyzing the evolution of Ultimatum Game bargaining within a dual process cognition framework. This is of great interest because – in the attempt to understand whether differences between theoretical predictions and experimental findings in the Ultimatum Game are driven by intuitive behaviors – many experimental studies have applied cognitive manipulations to subjects playing the Ultimatum Game (see Capraro (2019) for a review). Some papers find intuition increasing (or equivalently deliberation decreasing) both the offers made by proposers (Cappelletti, Güth, and Ploner, 2011; Achtziger, Alós-Ferrer, and Wagner, 2016; Halali, Bereby-Meyer, and Ockenfels, 2013) and receivers' rejection rates (Sutter, M. Kocher, and Strauß, 2003; Knoch et al., 2006; Cappelletti, Güth, and Ploner, 2011; Grimm and Mengel, 2011; Neo et al., 2013; Achtziger, Alós-Ferrer, and Wagner, 2016). However, some other works suggest that cognitive manipulations generate either no effects (Clare Anderson, 2010; Cappelletti, Güth, and Ploner, 2011) or even opposite effects to the ones just mentioned (Clare Anderson, 2010; Halali, Bereby-Meyer, and Ockenfels, 2013). However, despite these unclear results in the experimental literature, no theoretical work has directly tackled the issue.

Some theoretical papers have suggested other possible explanations for the experimental evidence on Ultimatum Game bargaining. For example, reputation (Martin A Nowak, Page, and Sigmund, 2000; André and Baumard, 2011), structured interactions (Page, Martin A Nowak, and Sigmund, 2000; Alexander, 2007), and mistakes (David G. Rand, Tarnita, et al., 2013) have been argued to give rise to fair offers. A more comprehensive review and classifications of models providing explanations for the evolution of fairness in the Ultimatum Game can be found in Debove, Baumard, and André (2016). These models are certainly of great interest. However, they seem to provide only partial explanations of the phenomenon. In fact, these contributions either cannot fully explain why experimental subjects deviate from theoretical predictions in one-shot and anonymous Ultimatum Game bargaining, which do not feature elements such as reputation and structured interactions; or they rely on relatively high mistake probabilities.

On the one side, I find that receivers consistently end up adopting the same strategy: they do not accept an offer unless it is as large as their outside option, which depends on their cognitive contingency (intuition, deliberation if the game is an Ultimatum Game, and deliberation if the game is a simplified Bargaining Game). Moreover, receivers' deliberation patterns are significantly affected by the offer made by the proposer. In fact, while null offers or generous offers, i.e., offers equal or above 50% of the value to split, always make receivers deliberate relatively infrequently; medium to low offers push receivers to deliberate more, and this especially holds for given offer-specific probabilities that the game is a simplified Bargaining Game. This implies that proposers' behavior can significantly impact receivers' reliance on deliberation (endogenous receivers' deliberation).

On the other side, I find that proposers follow different strategies depending on the width of the cost distribution considered. More precisely, if the cost distribution has large support (and, so, on average high deliberation costs must be paid to guarantee frequent deliberation), proposers adopt a pooling strategy according to which they always make the same offer, which is equal to receivers' intuitive outside option. Instead, if the cost distribution is narrow, proposers switch to the following differentiated strategy: if they deliberate and find out that the game is a simplified Bargaining Game, they make an offer equal to the receivers' outside option under deliberation if the game is a simplified Bargaining Game; otherwise, they offer to receivers their intuitive outside option. Further, I find that proposers' deliberation patterns mainly depend on the proposers' own strategy. More precisely, proposers tend to deliberate more frequently the more their strategy is differentiated, which happens for intermediate values of the probability that the game is a simplified Bargaining Game.

Finally, the model predicts huge variability in rejection rates conditional on a given offer being made by proposers. Such heterogeneity can be found both across offers made and for a given offer as different probabilities of playing a simplified Bargaining Game are considered. Moreover, even in the absence of mistakes, the model generates rejections rates of about $5\% - 10\%$.

The paper provides a new theoretical explanation for many findings in the experimental literature employing the Ultimatum Game, such as proposers making fair offers, receivers frequently rejecting strictly positive offers, and possibly the existence of cross-country differences in behavior. Further, the model provides interesting predictions about the effects of cognitive manipulations in the Ultimatum Game and further suggests new research questions regarding proposers' and receivers' behavior and their strategic interaction in the Ultimatum Game.

The remaining part of the paper is structured as follows: in Section 4.2 I present the theoretical framework considered, in Section 4.3 I illustrate and discuss the main results, and Section 4.4 concludes and provides insights for future research.

4.2 The Model

Consider a population composed of two types of agents: a set, $\mathcal{I}$, of I proposers indexed by $i = 1, \ldots, I$ and a set, $\mathcal{J}$, of J receivers indexed by $j = 1, \ldots, J$. Time is discrete and indexed by $t = 0, 1, \ldots$.

In each period of time t, proposer-receiver couples are randomly generated. Randomly matched agents play either a simplified Bargaining Game (BG) – with probability $\beta \in (0, 1)$ – or an Ultimatum Game (UG) – with probability $1 - \beta$. In Figure 13, I report the extensive form representation of the two games. As the figure shows, the two games have a similar structure. An amount $\alpha > 0$ must be split between the proposer and the receiver. The proposer moves first and chooses how much to offer to the receiver given a discrete set of possible offers $p \in \{0, \ldots, P\}$. Once received an offer, the receiver decides whether to accept, Acc, or to reject, Rej, it. If the receiver accepts the offer, then the suggested split is implemented. Instead, if the receiver rejects the offer, then agents' payoffs depend on the actual game played. If the agents are playing an UG, then they both get a null payoff. Instead, if they are playing a simplified BG, then the receiver gets a payoff of $\pi = \frac{\delta\alpha}{1+\delta}$ while the proposer gets a payoff of $\delta\pi$ where $\delta \in (0, 1)$ corresponds to the discounting factor with which agents would discount the amount α were they playing an infinitely repeated Bargaining Game (Rubinstein, 1982). I interpret these last payoffs as the ones characterizing the sub-game perfect Nash Equilibrium of the infinitely repeated Bargaining Game in case the amount to split is $\delta\alpha$ – the amount α is discounted because such payoffs are achieved after a first rejection from the receiver.

Agents do not know the type of game they are playing; to obtain this information, they need to engage in costly deliberation. A deliberation cost, c_{it} [c_{jt}], is randomly sampled decision-by-decision and independently for each agent from a known distribution $\mathcal{C}(\cdot) \in [0, C]$. Each agent chooses a threshold cost of deliberation, k_{it} [$k_{jt}(p)$], which corresponds to the maximum deliberation cost the agent is willing to pay to deliberate – note that a receiver's threshold cost of deliberation may depend on the actual offer p made by the proposer. I assume that agents choose their threshold cost of deliberation from a discrete set $\mathcal{K} = \{0, \ldots, k, \ldots, K\}$. If an agent's threshold cost of deliberation is larger than its sampled deliberation cost, i.e. if, for example, $k_{it} \geq c_{it}$, then the agent deliberates and, consequently, it becomes informed about which game it is currently facing and, thus, it can condition its offer [decision to accept or reject the offer] on whether the game is an UG or a simplified BG.

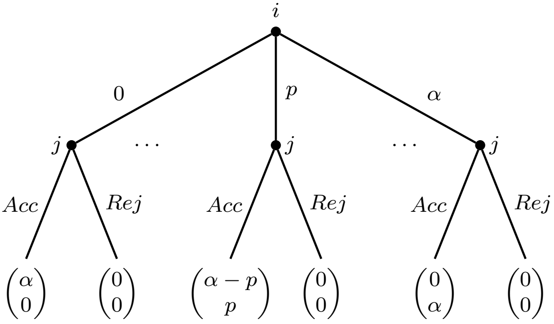
I consider reinforcement learning agents. Every agent has weights or propensities associated with each possible behavior it may adopt. In particular, every proposer's strategy is characterized by four vectors of propensities:

1. $W_{it}^{k} = (w_{it,0}, \ldots, w_{it,k}, \ldots, w_{it,K})$ which contains one propensity for each available threshold cost of deliberation k;

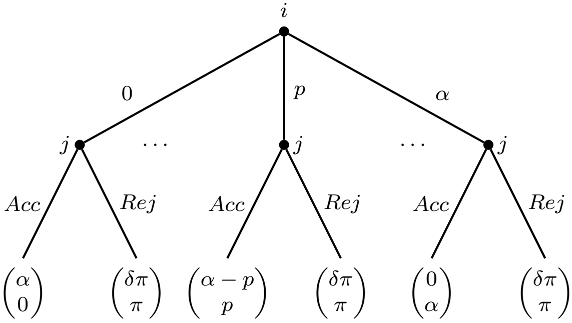
2. $W_{it}^{Int} = (w_{it,0}^{Int}, \ldots, w_{it,p}^{Int}, \ldots, w_{it,P}^{Int})$ containing one propensity for each available offer p; these propensities are used if the proposer plays under intuition;

3. $W_{it}^{UG} = (w_{it,0}^{UG}, \ldots, w_{it,p}^{UG}, \ldots, w_{it,P}^{UG})$ containing one propensity for each available offer p; these propensities are used if the proposer plays under deliberation and it finds out that the game is an UG;

4. $W_{it}^{BG} = (w_{it,0}^{BG}, \ldots, w_{it,p}^{BG}, \ldots, w_{it,P}^{BG})$ containing one propensity for each available offer p; these propensities are used if the proposer plays under deliberation and it finds out that the game is a simplified BG.



(a) Ultimatum Game



(b) Simplified Bargaining Game

Figure 13: Ultimatum Game and simplified Bargaining Game. Extensive form representations of the Ultimatum Game (a) and of the simplified Bargaining Game (b) under the assumption $P = \alpha$, i.e., the maximum offer the proposer can make is equal to the entire amount to split.

Instead, every receiver's strategy can be described by the following four matrices:

1. $W_{jt}^{k} = \begin{pmatrix} w_{jt,0}^{0} & \cdots & w_{jt,k}^{0} & \cdots & w_{jt,K}^{0} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ w_{jt,0}^{p} & \cdots & w_{jt,k}^{p} & \cdots & w_{jt,K}^{p} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ w_{jt,0}^{P} & \cdots & w_{jt,k}^{P} & \cdots & w_{jt,K}^{P} \end{pmatrix}$
   This (P, K) matrix contains all the receiver's propensities associated with threshold costs of deliberation. The (p, k)-th element is the propensity associated to threshold k given that the proposer has offered an amount p;

2. $W_{jt}^{Int} = \begin{pmatrix} w_{jt,Acc}^{0,Int} & w_{jt,Rej}^{0,Int} \\ \vdots & \vdots \\ w_{jt,Acc}^{p,Int} & w_{jt,Rej}^{p,Int} \\ \vdots & \vdots \\ w_{jt,Acc}^{P,Int} & w_{jt,Rej}^{P,Int} \end{pmatrix}$
   This $(P, 2)$ matrix contains the receiver's propensities to accept or reject the offer p made by the proposer given that the receiver is playing under intuition;

3. $W_{jt}^{UG} = \begin{pmatrix} w_{jt,Acc}^{0,UG} & w_{jt,Rej}^{0,UG} \\ \vdots & \vdots \\ w_{jt,Acc}^{p,UG} & w_{jt,Rej}^{p,UG} \\ \vdots & \vdots \\ w_{jt,Acc}^{P,UG} & w_{jt,Rej}^{P,UG} \end{pmatrix}$
   This $(P, 2)$ matrix contains the receiver's propensities to accept or reject the offer p made by the proposer given that the receiver is playing under deliberation and it has found out that the game is an UG;

4. $W_{jt}^{BG} = \begin{pmatrix} w_{jt,Acc}^{0,BG} & w_{jt,Rej}^{0,BG} \\ \vdots & \vdots \\ w_{jt,Acc}^{p,BG} & w_{jt,Rej}^{p,BG} \\ \vdots & \vdots \\ w_{jt,Acc}^{P,BG} & w_{jt,Rej}^{P,BG} \end{pmatrix}$

   This $(P, 2)$ matrix contains the receiver's propensities to accept or reject the offer p made by the proposer given that the receiver is playing under deliberation and it has found out that the game is a simplified BG.

In every period of time t, each agent behaves according to its current propensities unless it makes a mistake. More precisely, when making a choice in a given information set an agent chooses behavior s' among the available behaviors in the set $\mathcal{S}$ with probability proportional to its current weight:

$$Pr(s') = (1 - \varepsilon)\frac{w_{s'}}{\sum_{s \in \mathcal{S}} w_s} + \varepsilon \frac{1}{|\mathcal{S}|}$$

where $\varepsilon \in [0, 1)$ is the error rate with which an agent selects a behavior at random rather than following its propensities.

After each interaction, every agent updates its relevant weights, i.e., the ones among which it has selected its behavior. It does this by discounting each weight by a factor $\varphi \in (0, 1)$ and by adding to the chosen weights the payoff obtained. In particular, if in a given information set an agent has chosen option s' among the ones in the set $\mathcal{S}$ then the propensity associated with each choice $s \in \mathcal{S}$ is updated as follows:

$$w_{t+1,s} = \varphi w_{t,s} + \text{payoff } I_{s=s'}$$

where $I_{s=s'}$ is an indication function taking value 1 if an only if $s = s'$. Therefore, propensities associated with unused behaviors are simply discounted while propensities associated with used behaviors are first discounted and then increased by the payoff they have contributed to generate. Discounting allows an agent to gradually forget information provided by past play and, consequently, it makes an agent weigh more current payoff information. To initialize the model one must choose a baseline value of propensities, say w, which determines agents' behavior in the first rounds of play.

To summarize, in each period of time:

1. Randomness is resolved: (i) proposers and receivers are randomly matched into couples, (ii) the game played by the two agents is randomly selected, and (iii) each agent's actual cost of deliberation is sampled from the distribution $\mathcal{C}(\cdot)$;
2. Agents play the game according to their propensities and payoffs are realized;
3. Each agent updates its propensities using payoff information.

4.2.1 Model Example

To better illustrate the model, I provide a simple example. Consider the case in which at time t proposer i and receiver j are randomly matched to play an Ultimatum Game.

Assume that the amount to split is $\alpha = 10$ and that proposers can make any offer $p \in \{0,\ 5,\ 10\}$. Assume also that the cost distribution is $\mathcal{C}(\cdot) \sim \mathcal{U}(0,1)$, that at time t proposer i has been assigned a cost of deliberation $c_{it} = 0.13$, while receiver j faces a cost $c_{jt} = 0.3$. Further, assume that agents must choose their threshold cost from the set $k \in \{0,\ 0.5,\ 1\}$. Finally assume that $\delta = 0.6$ so that $\pi = 3.75$, agents make a mistake in implementing their strategy with probability $\varepsilon = 0.1$ and they discount propensities with a discounting factor $\varphi = 0.9$.

Assume that proposer i's strategy at time t is:

1. $W_{it}^{k} = (w_{it,0},\ w_{it,0.5},\ w_{it,1}) = (5,\ 10,\ 0)$;
2. $W_{it}^{Int} = (w_{it,0}^{Int},\ w_{it,5}^{Int},\ w_{it,10}^{Int}) = (8,\ 8,\ 8)$;
3. $W_{it}^{UG} = (w_{it,0}^{UG},\ w_{it,5}^{UG},\ w_{it,10}^{UG}) = (12,\ 4,\ 0)$;
4. $W_{it}^{BG} = (w_{it,0}^{BG},\ w_{it,5}^{BG},\ w_{it,10}^{BG}) = (1,\ 95,\ 4)$.

This means that, for example, proposer i selects threshold $k_{it} = 0.5$ with probability:

$$\begin{aligned} Prob(k_{it} = 0.5) &= (1-\varepsilon)\frac{w_{it,0.5}}{\sum_{k\in\mathcal{K}} w_{it,k}} + \varepsilon\frac{1}{|\mathcal{K}|} \\ &= (1-0.1)\frac{10}{15} + 0.1\frac{1}{3} \\ &= 0.6\overline{3} \end{aligned}$$

Moreover, $Prob(k_{it} = 0) = 0.\overline{3}$ and $Prob(k_{it} = 1) = 0.0\overline{3}$.

Assume now that proposer i chooses $k_{it} = 0.5$. But then, given that $c_{it} = 0.13 < 0.5 = k_{it}$, proposer i deliberates. Consequently, the vector $W_{it}^{UG} = (12,\ 4,\ 0)$ will determine proposer i's offer choice – as by assumption the game is an UG.

Assume that according to its propensities in W_{it}^{UG} proposer i offers $p_{it} = 5$ and that receiver j accepts the offer. Then proposer i will obtain a payoff of $u_{it} = \alpha - p_{it} = 10 - 5 = 5$. Then, proposer i will update its propensities in W_{it}^{k} and W_{it}^{UG}, while the ones contained in W_{it}^{Int} and W_{it}^{BG} will remain the same. For example, given its play ($p_{it} = 5$) and given the payoff it has obtained ($u_{it} = 5$) proposer i will update its propensities associated with making offers under deliberation if the game is an UG as follows:

$$\begin{aligned} W_{it+1}^{UG} &= (\varphi \times w_{it,0}^{UG},\ \ \varphi \times w_{it,5}^{UG} + u_{it},\ \ \varphi \times w_{it,10}^{UG}) \\ &= (0.9 \times 12,\ \ 0.9 \times 4 + 5,\ \ 0.9 \times 0) \\ &= (10.8,\ 8.6,\ 0) \end{aligned}$$

Similar reasoning applies to receiver j's strategy implementation and strategy revision. The main difference is that the receiver will implement its strategy conditionally on the offer it has received from the proposer. For example, given that the proposer has offered $p_{it} = 5$, the receiver will choose its threshold cost of deliberation and – given its cognitive contingency – whether to accept or reject the offer by considering only the propensities associated with an offer $p = 5$ (in the example made only the second row of each matrix would matter). Moreover, only a specific row of each matrix used will be updated after the receiver has obtained its payoff (again, the row associated with offer $p = 5$).

4.3 Results

I simulate the model by setting the parameters to the following values. First, $I = J = 50$ and so both proposer′ and receivers′ populations are made up of fifty agents. This in turn implies that in every period of time, every proposer is matched with a unique receiver, and the other way around: no agent is left unmatched. I let the system evolve for 10^6 iterations and, so, $t = 0, \dots, 10^6$. With respect to the UG and simplified BG parameters, I set the amount to split to $\alpha = 10$ and allow each proposer to choose its offer p from the set $\{0, 1, \dots, 10\}$. This together with setting agents′ patience factor to $\delta = 0.8$ implies $\pi = \frac{\delta\alpha}{1+\delta} = 4.4$. I assume that agents make a mistake in implementing their strategy with probability $\varepsilon = 0.05$ and that while revising their strategy they discount old propensities by $\varphi = 0.99$. All propensities are initially set to $w = 100$, thus, generating an initial random behavior of the system. Regarding the cognitive part of the model, I assume that time-agent idiosyncratic costs of deliberation are randomly sampled from a distribution $\mathcal{C}(\cdot) \sim \mathcal{U}(0, C)$ with $C \in \{0.25, 0.5, 1, 2\}$. Moreover, agents choose a threshold cost of deliberation from the set $\mathcal{K} = \{0, \frac{1}{8}C, \frac{2}{8}C, \dots, C\}$ and, consequently, the maximum threshold cost an agent may choose always corresponds to the maximum deliberation cost that may be sampled from the distribution $\mathcal{C}(\cdot)$. Moreover, independently of the maximum threshold cost available to agents, they can always choose from a fixed amount of threshold costs of deliberation. Finally, I focus the analysis on the case in which the probability that the game is a simplified Bargaining Game β belongs to the set $\{0.1, 0.2, \dots, 0.9\}$. For each possible combination of β and C values, I perform five independent simulations of the model.

In the following, I present first the results concerning receivers′ behavior, then the ones regarding proposers′ behavior, and finally, I combine these findings to analyze implied rejection rates.

4.3.1 Results: Proposers' Equilibrium Play

I begin by analyzing the proposers' equilibrium play. In Figure 14 I plot the average offer made by proposers in the three possible cognitive contingencies (intuition, deliberation if the game is an UG, and deliberation if the game is a simplified BG) as a function of β by considered cost distribution.

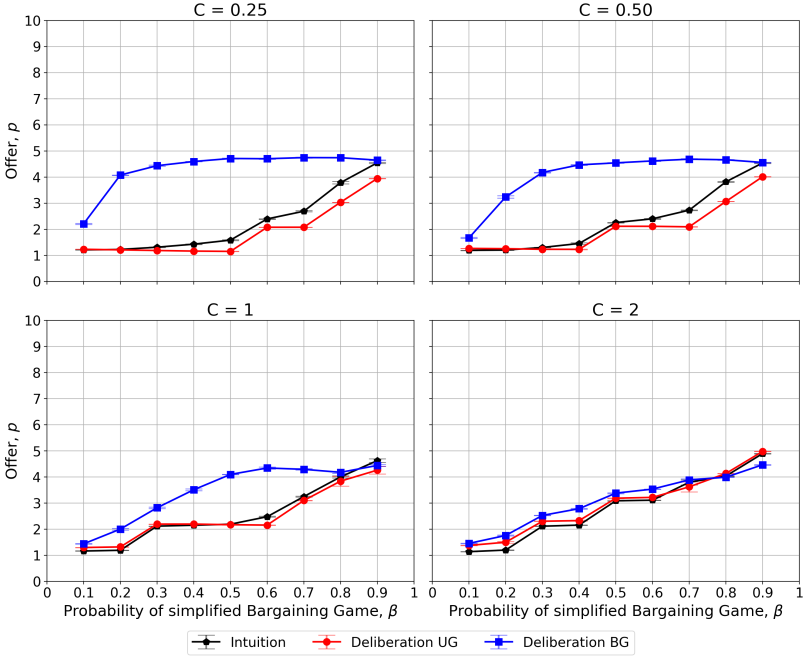


Figure 14: Proposers' behavior. Proposers' average offer under intuition, under deliberation if the game is an UG, and under deliberation if the game is a simplified BG and their standard errors as a function of β by cost distribution.

The first interesting result is that the proposers' equilibrium play is not unique: depending on the distribution of deliberation costs considered different proposers' strategies may be favored by evolution. More precisely, if the cost distribution is wide (case $C = 2$) proposers evolve to offer roughly the same amount independently of their specific cognitive contingency. Such an equilibrium offer is increasing in the probability that the game is a simplified BG and it is quite close to $\beta\pi$. As one moves to intermediate cost distributions (cases $C = 1$ and for low values of β case $C = 0.5$) proposers evolve to make differentiated offers. In particular, under intuition and under deliberation if the game is an UG they still make offers close to $\beta\pi$; while they make fair offers (close to π) under deliberation if the game is a simplified BG. Finally, in the case of narrow cost distributions (cases $C = 0.5$ for low values of β and $C = 0.25$), proposers further differentiate their strategy: in addition to making offers close to π if they deliberate and find out that the game is a simplified BG they also make consistently higher offers under intuition than under deliberation if the game is an UG even though the difference between the two is small. However, it is reasonable to believe that as the maximum cost of deliberation goes to zero proposers evolve to make differentiated offers in all cognitive contingencies with offers made under deliberation if the game is an UG approaching zero.

These findings suggest that experimental cognitive manipulations in the UG should have little to no effect on proposers' behavior unless subjects have developed their strategies in a context characterized by both negligible deliberation costs (small C) and (eventually) frequent bargaining interactions (high β). Moreover, if such conditions are met, then one should find increased reliance on intuitive behaviors associated with (weakly) increasing proposers' offers in the UG, but never the other way around.

Another feature that is worth discussing is the shape of the average offers made under intuition and under deliberation if the game is an UG as a function of β for different cost distributions. In fact, even though independently of C (i) these measures are increasing in β, (ii) start from a minimum of one, and (iii) reach a maximum of about four, different patterns are observed. In fact, if the cost distribution is wide (case $C = 2$) as β increases the average offers made in the UG game rise at an approximately constant rate and, consequently, offers are almost linear in β. Instead, in the case of narrow cost distributions the average offers made under intuition and the ones made under deliberation if the game is an UG tend to stay at their minimum level for increasingly more values of β; only for medium-high values of β offers begin to increase and they do so at a fast rate.

The previous considerations imply that in most of the values of β considered proposers tend to offer less in the UG if the maximum cost of deliberation is low. For example, if $\beta = 0.5$ then in the case $C = 0.25$ proposers' average offers in the UG are close to one, while if $C = 2$ then such offers are more than three times as large. I interpret this as the benefit proposers obtain by adopting a credible differentiated strategy. However, this becomes clearer if one considers the receivers' behavior. For this reason, I defer the explanation to Section 4.3.2.

Figure 15 reports proposers' probability to deliberate – which corresponds to the average threshold cost of deliberation chosen over the maximum deliberation cost C – as a function of β for the different cost distributions considered.

As the figure shows proposers' probability to deliberate decreases less than proportionally as the cost distribution widens. However, the probability that the game is a simplified BG does not have a clear and consistent effect on proposers' propensity to deliberate across the cost distributions considered. Despite this, in the case of most cost distributions (with the exception of the case $C = 2$) one can find either one or two peaks of deliberation levels. This might be explained by looking at Figure 14. In fact, peaks of proposers' propensity to deliberate occur for values of β such that the difference between the average offer made under deliberation if the game is a simplified BG game and the average offers made in the other cognitive contingencies is locally maximal. Therefore, it seems that proposers' propensity to deliberate is mostly determined by proposers' own strategy: they will have higher incentives to deliberate as they adopt more differentiated strategies.

This finding can also provide an interesting interpretation of why proposers adopt less differentiated strategies in the case of cost distributions with high C. In fact, given that wider cost distributions make deliberation more costly for a fixed probability of deliberation (or equiv-

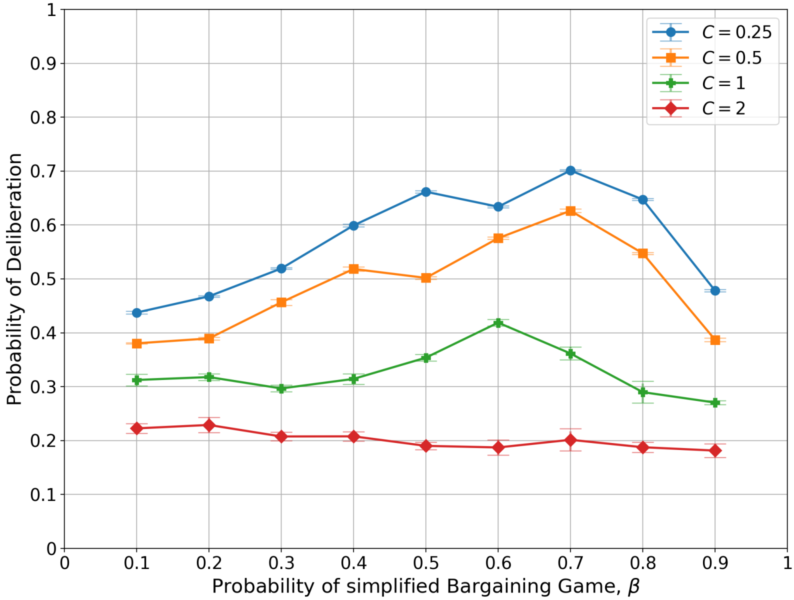


Figure 15: Proposers' deliberation patterns. Average proposers' probability to deliberate and their standard errors as a function of β by cost distribution.

alently they make deliberation less frequent for a fixed threshold cost of deliberation), by adopting less differentiated strategies in the case of wide cost distribution proposers' decrease their own incentive to deliberate and, thus, avoid incurring into costly deliberation.

4.3.2 Results: Receivers' Equilibrium Play

Figure 16 provides a synthetic description of receivers' game-play equilibrium behavior. More precisely, the figure reports receivers' average minimum accepted offer (MAO) under intuition, receivers' MAO under deliberation if the game is an UG, and receivers' MAO under deliberation if the game is a simplified BG as a function of β separately for the different cost distributions considered. I define the MAO of a receiver in a given cognitive contingency as the minimum offer p such that the receiver is willing to accept the offer p with probability 0.5 or higher conditional on being in that cognitive contingency.

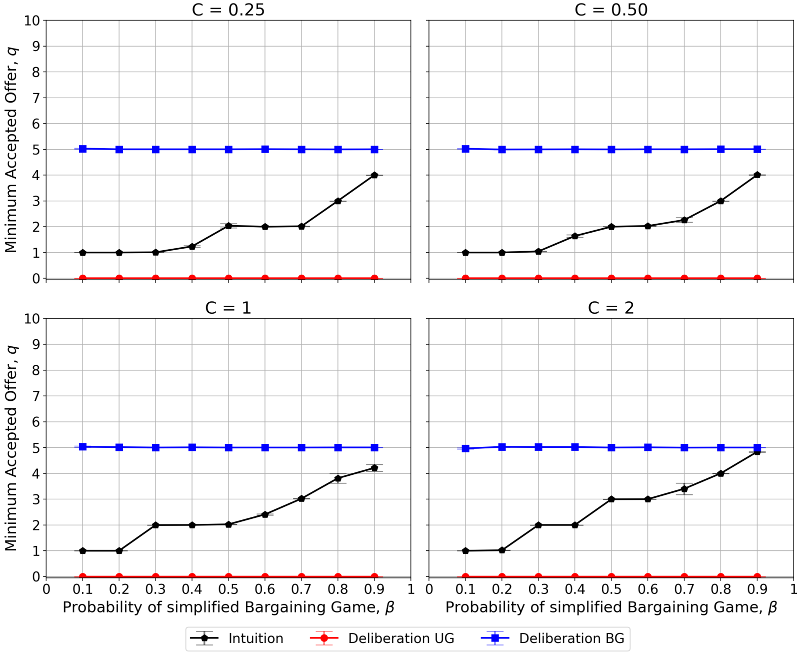


Figure 16: Receivers' behavior. Receivers' average MAO under intuition, under deliberation if the game is an UG, and under deliberation if the game is a simplified BG and their standard errors as a function of β by assumed cost distribution.

As the figure shows, on the one hand, receivers' behavior under deliberation is independent of the probability that the game played is a simplified BG. In fact, if receivers deliberate then independently of the actual value of β their MAO is equal to five under deliberation if the game is a simplified BG, while it is consistently equal to zero if they deliberate and find out that the game is an UG. On the other hand, receivers' average MAO under intuition is (i) increasing in β, (ii) bounded below by the average MAO under deliberation if the game is an UG, and (iii) bounded above by the average MAO under deliberation if the game is a simplified BG. Overall, receivers' MAO under intuition is often close to the value $\beta\pi$.

These findings can be interpreted as follows: a receiver does not accept an offer unless such offer is at least as large as the receiver's expected outside option given its current cognitive contingency where a receiver's outside option corresponds to the payoff it can secure itself by rejecting the proposer's offer. In fact, a receiver's outside option is null if it deliberates and it finds out that the game is an UG, while it corresponds to $\pi = 4.4$ (given $\alpha = 10$ and $\delta = 0.8$) if the receiver deliberates and it finds out that the game played is a simplified BG. Moreover, given that under intuition a receiver does not know the type of game it is currently playing, but it knows the probability it is playing one game or another, then a receiver's expected outside option under intuition is $\beta\pi$ as with probability β the game is a simplified BG, and consequently, its outside option is $\pi = 4.4$, while with probability $1 - \beta$ the game is an UG, and its outside option is null.

It is also interesting to underline that while receivers' MAO under intuition is often close to $\beta\pi$, in the case of narrow cost distributions, receivers are less demanding and this is more evident for low and medium values of β. This is clearly illustrated by the different shapes assumed by the average MAO under intuition for different cost distributions: for $C = 2$ the average MAO under intuition is almost linear in β, while in the case $C = 0.25$ it is flat for low values of β.

This finding can be accounted for by the previous interpretation (i.e., a receiver does not accept an offer unless it is as large as its expected outside option) by recognizing that a receiver's expected outside option does not depend solely on receiver's cognitive contingency and the type of game played, but it may also be affected by the proposer's behavior. More precisely, in the case of narrow cost distributions proposers adopt a differentiated strategy such that they make a low offer under intuition and under deliberation if they find out that the game is an UG, while they make a high offer under deliberation if the game is a simplified BG. Given that proposers adopt this kind of differentiated strategy, receivers' expected outside option conditional on observing a low offer is lower than $\beta\pi$. In fact, given proposers' behavior, conditional on observing a low offer from the proposer a receiver should expect to be in a simplified BG with probability $\beta\big(1-\mathbb{E}[\mathcal{C}(k_p)]\big)$ where $\mathbb{E}[\mathcal{C}(k_p)]$ is the expected probability that a proposer deliberates. But then, receiver's expected outside option is $\beta\big(1-\mathbb{E}[\mathcal{C}(k_p)]\big)\pi < \beta\pi$. Interestingly, this outside option is decreasing in proposers' expected probability to deliberate. But then, when adopting a differentiated strategy of this kind, proposers face a trade-off between their level of deliberation and the offer they must make to ensure acceptance from receivers: deliberating more allows to reduce the offer made but increases deliberation costs and vice versa. If this is the case, one can explain the flattening of average MAO and offers in the case of narrow cost distributions. In fact, by looking at Figure 15, one can observe that in the cases $C = 0.25$ and $C = 0.5$ (in which proposers adopt differentiated strategies) for low values of β proposers significantly increase their level of deliberation as β increases. This implies a significant decrease of $\big(1-\mathbb{E}[\mathcal{C}(k_p)]\big)$ which counteracts the increase of β and, so, keeps MAO and offers at low levels. Moreover, MAO and offers become increasing in β as proposers stabilize their level of deliberation (because it becomes too costly to deliberate more).

In any case, the results reported in Figure 16 provide clear predictions regarding the effects of cognitive manipulations on the behavior of experimental subjects playing as receivers in the UG. In fact, these results imply that increased reliance on deliberation always decreases receivers' rejection rates in the UG. Moreover, differences between the rejection rates under intuition and the ones under deliberation become more evident as β increases.

I move now to the analysis of the receivers' decision to deliberate. In Figure 17 I report the average probability that a receiver deliberates conditional on being offered an amount p. These measures are reported separately by assumed cost distribution.

The most interesting cases to analyze are receivers' probabilities to deliberate conditional on offers $p \in \{1, 2, 3, 4\}$. Therefore, I will describe the main findings associated to Figure 17 by focusing only on these of-

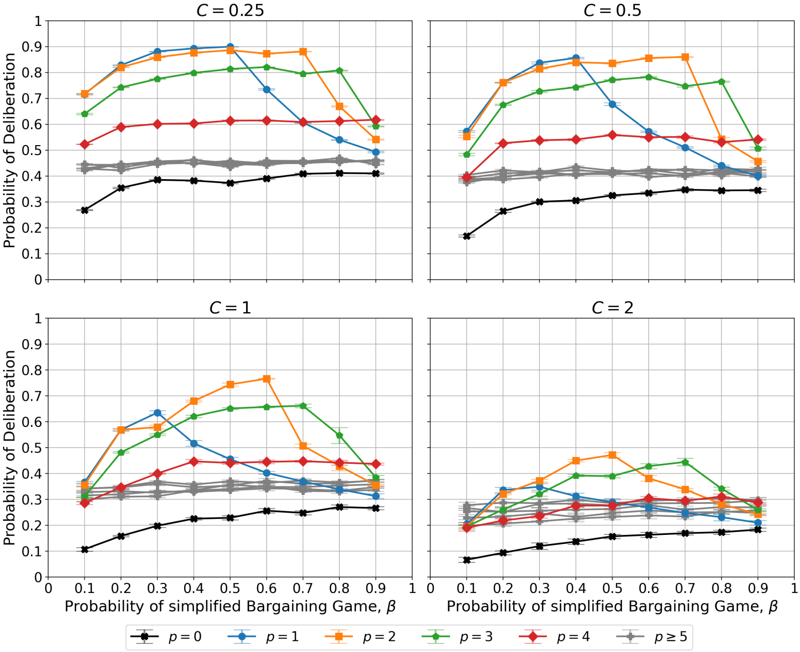


Figure 17: Receivers' deliberation patterns. Average receivers' probability to deliberate given that the proposer has offered an amount p and their standard errors as a function of β by cost distribution.

fers. Afterward, I will comment on the cases $p \in \{0, 5, 6, 7, 8, 9, 10\}$ which are of lower interest for the following reasons: (i) they do not present interesting patterns and (ii) in equilibrium proposers never make this kind of offers (as seen in Section 4.3.1).

As the cost distribution widens, i.e., as C increases, receivers tend to rely less frequently on deliberation. However, they do so less than proportionally as in the case of proposers. Another interesting result regarding changes in the cost distribution is that receivers' deliberation patterns change and, more precisely, they specialize as the cost distribution widens. This specialization pattern is twofold.

The first specialization in deliberation patterns refers to changes in deliberation across different values of β as the cost distribution changes conditional on a given offer p. More precisely, if the maximum cost of deliberation, C, is low then receivers tend to deliberate with the same (high) probability for most values of β. However, as the cost distribution widens even though receivers tend to deliberate less frequently ceteris paribus they decrease by different extents their probability to deliberate depending on the specific value of β. Consider, for example, the case $p = 2$. As C increases receivers deliberate less frequently but such a decrease in deliberation is significantly more pronounced for extreme values of β. Therefore, as deliberation costs increase – conditional on being offered an amount p – receivers give up deliberating more or less depending on the actual value of β.

The second form of receivers' specialization in deliberation patterns refers to changes in deliberation patterns for different offers made by proposers. More precisely, as the costs of deliberation increase depending on the specific value of β receivers tend to concentrate their cognitive efforts on the cases in which they are offered some specific amount. In particular, if β is low receivers deliberate relatively more if they get offered $p = 1$. However, as β increases receivers switch to deliberate more in case they receive an offer $p = 2$ (middle values of β), $p = 3$ (middle-high values of β), or $p = 4$ (high β). Moreover, this tendency is more pronounced for high values of C.

I interpret this second form of specialization in receivers' deliberation patterns as endogenous receivers' deliberation. This gives rise to interesting strategic considerations as in principle a proposer could affect the likelihood that a receiver incurs into deliberation to its own benefit. But, more generally, even if the proposer does not make strategic offers it will still usually affect the receivers' propensity to deliberate via the offer it makes. If this is the case, proposers' behavior can in principle alter the intended effects of cognitive manipulations in experimental settings. Consider, for example, a time delay treatment on proposers and receivers. If proposers under time delay make really low offers, say zero or one, then such low offers may push receivers to play intuitively (contrary to the intended effect of the time delay treatment) and to reject the offer.

Finally, by comparing Figure 17 with Figure 15 it is clear that receivers tend to deliberate more frequently than proposers. This is likely due to the differences in strategic complexity across types of agents which in turn affect receivers' and proposers' expected benefits from deliberation. Consider a receiver's point of view. Given an offer p, either incurring into deliberation is useless (cases $p = 0$ or $p \geq 5$ which however almost never occur given proposers' equilibrium play) or deliberating has a pivotal role (cases $p \in \{1, 2, 3, 4\}$) as it tells the receiver whether it is optimal to accept or reject the offer and, so, it allows a receiver to achieve the highest payoff conditional on proposer's offer and the actual game being played. Consider now a proposer's point of view. Even though deliberating informs the proposer about its optimal play, the actual outcome a proposer will get will crucially depend on the receiver's play and, in particular, on whether it will deliberate or not.

I move now to the analysis of receivers' deliberation patterns if they are offered an amount $p \geq 5$. Given receivers' game-play behavior, they accept this kind of offer in any cognitive contingency and, consequently, deliberation does not provide any benefit. Despite this, receivers' probabilities to deliberate conditional on being offered amounts $p \geq 5$ are consistently bounded away from zero, they are higher in the case of narrow cost distributions, and they are close to 0.5 if $C = 0.25$. I interpret these findings as evidence of low evolutionary pressure on receivers' deliberation patterns in case they are offered generous amounts. In fact, given the cost distributions considered and offers $p \geq 5$, even if a receiver pays the maximum possible deliberation cost it will still obtain a high payoff by accepting the offer and such payoff will not be significantly lower than the one obtainable by paying zero deliberation costs. For example, if $p = 6$ and $C = 0.25$, choosing the maximum threshold cost available, i.e., $k = 0.25$, implies (i) always deliberating, (ii) paying on average a deliberation cost of 0.125 and so to obtain an expected payoff of 5.875 while choosing the minimum threshold cost available, i.e., $k = 0$, implies never deliberating and, thus, obtaining an expected payoff of 6 which is not significantly higher than the one obtained by always paying the deliberation cost. Of course, as the cost distribution widens evolutionary selection of receivers' deliberation patterns conditional on generous offers becomes stronger and it does so especially for relatively less generous offers. This explains why as one considers wider cost distributions receivers' probabilities to deliberate conditional on receiving an offer $p \geq 5$ become more heterogeneous.

Finally, consider the case in which a receiver is offered an amount $p = 0$. As Figure 17 shows this is the scenario in which receivers deliberate less frequently. This is likely the combination of no incentive to deliberate (as in the case of offers $p \geq 5$) together with stronger selection pressure. The low incentive to deliberate arises because independently of the actual game played receivers are in expectation better off by rejecting a null offer. In fact, accepting the offer implies a null payoff, while rejecting the offer implies an expected payoff of $\beta\pi$ – or $\beta\big(1 - \mathbb{E}[\mathcal{C}(k_p)]\big)\pi$ if proposers adopt a differentiated strategy. Instead, the higher selection pressure comes from the fact that $\beta\pi$ is always smaller than any generous offer considered before. This can also explain why receivers' probability to deliberate conditional on being offered an amount $p = 0$ is increasing in β: for low values of β receivers expected payoff is small and so selection pressure on deliberation patterns is strong while for high values of β receivers' expected payoff converges to $\pi = 4.44$ which can be considered as a 'generous offer' and, consequently, (i) the strength of selection is low and (ii) receivers probability to deliberate gets close to the ones of generous offers.

4.3.3 Results: Play in the Ultimatum Game

In this section I combine the findings related to receivers' and proposers' equilibrium behavior to derive predictions regarding expected rejection rates in the Ultimatum Game; that is the rejection rates predicted by the model conditional on the game being an UG. I compute two types of rejection rates. The first ones are conditional rejection rates, i.e, measures of the average probability that a receiver rejects a given offer p conditional on the game being an UG. These quantities can be derived by simply looking at receivers' equilibrium propensities. The second type of rejection rates are expected rejection rates computed by taking into account simultaneously proposers′ and receivers′ expected behaviors. In other words, these are the rejections rates one should observe in an UG according to the model.

I begin by presenting the results related to conditional rejection rates. Conditional rejection rates are computed as follows. For any given (β, C) combination I compute the average receivers′ rejection rate conditional on receiving an offer p both under intuition and under deliberation if the game is an UG. Then, for any possible offer p, I derive the probability that a receiver plays intuitively or according to deliberation given that it has been offered an amount p. Finally, I use this information regarding receivers′ probability to deliberate as weights for intuitive and deliberative rejection rates, thus obtaining conditional rejection rates.

In Figure 18 I report receivers′ conditional rejection rates by offer p as a function of β for each cost distribution considered.

Overall conditional rejection rates are decreasing in the offer p: by making a higher offer a proposer usually guarantees himself a higher chance of having its offer accepted. However, there are some exceptions to this general rule. For example, in the case $C = 1$ for many intermediate values of β if a proposer offers four rather than three then it faces a higher rejection rate. I explain this in the light of Figure 17 which describes receivers′ deliberation patterns conditional on the offer received. In fact, by offering four rather than three a proposer triggers lower deliberation on the receivers′ side which in turn increases receivers′ rejection rates as rejection rates under intuition tend to be higher than the ones under deliberation if the game is an UG.

As in the case of receivers′ probability to deliberate given an offer p I now distinguish rejection rates conditional on offers $p \in \{1, 2, 3, 4\}$ from the ones conditional on null offers and the ones conditional on generous offers $p \geq 5$.

As the figure shows, receivers′ rejection rates conditional on being offered an amount $p \in \{1, 2, 3, 4\}$ are weakly increasing (non-decreasing) in β for a given cost distribution. More precisely, they tend to be stable and close to zero for low values of β but once β reaches a critical level (which depends both on p and C) conditional rejection rates increase dramati-

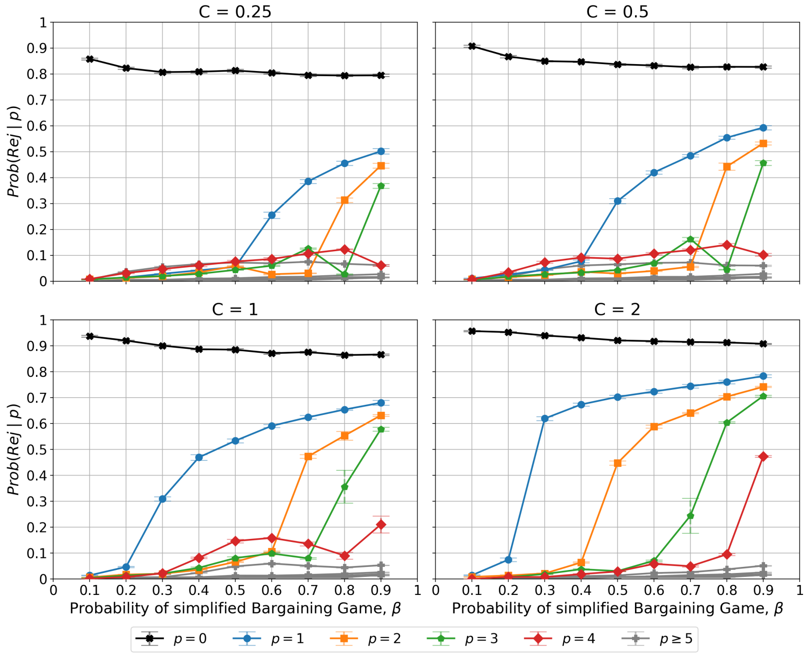


Figure 18: Conditional rejection rates. Average receivers' conditional rejection rates in the Ultimatum Game and their standard errors as a function of β by cost distribution.

cally until they converge to a maximum level. I explain this pattern as follows. For low values of β receivers tend to accept any offer both under intuition and under deliberation if the game is an UG and, consequently, conditional rejection rates are close to zero for low values of β. However, as β increases receivers begin to reject low offers under intuition (in order, offers equal to one, two, three). After receivers start to reject offer p under intuition, rejection rates conditional on offer p remain low as long as receivers deliberate frequently enough conditional on being offered an amount p; when they stop doing so rejection rates rise steadily. This explains why conditional rejection rates begin to increase after β reaches a critical value that depends both on p and C.

It is interesting to underline that conditional rejections rates reach a maximum level that is increasing in C and also depends on p. This is because the maximum level reached by conditional rejection rates is determined by the probability that receivers incur into deliberation given that they have been offered an amount p. More precisely, if receivers deliberate a fraction of times x then conditional rejection rates can be up to $1-x$ as under deliberation receivers always accept the offer if the game is an UG. This together with the findings that deliberation levels are lower in the case of wide cost distributions and they also depend on p – as seen in Section 4.3.2 – explains why the maximum level of conditional rejection rates is higher in the case of wide cost distributions and also depends on the offer p.

Conditional rejection rates of fair offers $p \geq 5$ are always close to zero, as expected, given that both under intuition and under deliberation if the game is an UG receivers always accept these kinds of offers. Instead, rejection rates conditional on null offers are always high, but they are never equal to one. This holds because it is always the case that under intuition receivers reject null offers while under deliberation they accept zero offers with high probability. Despite these opposing effects the first one (intuitive rejection) dominates as receivers deliberate with low probability if they are offered nothing. Interestingly, the shape of the conditional rejection rates associated with null offers closely remembers receivers' probability to deliberate conditional on $p = 0$ inverted.

It is worth underlining that the model can account for high conditional rejection rates in the UG for offers up to 40% of the value to split.

I move now to the analysis of overall rejection rates implied by the model. These rejection rates have been computed as follows. For any (β, C) combination I have derived proposers' expected probability to make any possible offer under intuition and under deliberation if the game is an UG; moreover, I have computed proposers' probability to incur into deliberation. With these measures at hand, I have derived the probabilities that proposers make any given offer in the UG. Then, I combined these measures with the corresponding conditional rejection rates to derive a measure of overall rejection rates. The rejection rates thus obtained are illustrated in Figure 19.

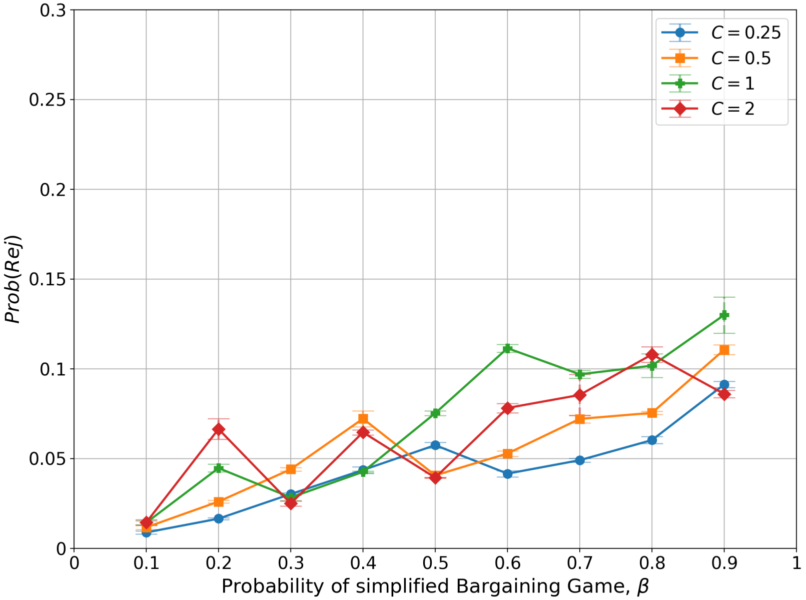


Figure 19: Expected rejection rates. Overall expected rejection rates in the Ultimatum Game and their standard errors as a function of β by cost distribution. These rejection rates are computed by taking into account both proposers' behavior and receivers' conditional rejection rates.

As the figure shows, expected rejection rates in the UG are overall increasing in the probability that the game is a simplified BG. This might be expected as the higher β the more receivers are demanding and, thus, the more they should reject offers. However, interestingly rejection rates in the UG are not monotonically increasing in β. In fact, expected rejection rates present multiple peaks and troughs. By comparing this finding with Figure 14 it seems that implied rejection rates decrease at values of β in correspondence of which proposers' average offer switches from one value (say p') to the next available one ($p' + 1$).

There is no clear relationship between expected rejection rates in the UG and the cost distribution considered. However, rejection rates are slightly lower in the case of narrow cost distributions. Moreover, under narrower distributions expected rejection rates are more stable and monotone.

Overall, in the absence of players' mistakes in implementing their strategy, the model predicts rejection rates between 5% and 10%. These figures are significantly lower than the rejection rates observed in the laboratory but still can explain a relevant part of the phenomenon.

Appendix G and Appendix H report some robustness checks. The former analyses the role played by the shape of the cost distribution, while the latter studies the effects associated with changes in agents' patience factor.

4.4 Conclusion

I have considered an evolutionary model featuring proposers and receivers interacting over time in Ultimatum Games and simplified Bargaining Games. Following recent modelings of dual process theories of cognition these agents cannot recognize the type of interaction they are facing unless they incur costly deliberation. I study how agents' behavior both in terms of game-play (offers and minimum accepted offers) and in terms of deliberation patterns is affected by changes in the probability that the game is a simplified BG and as deliberation costs are drawn from different uniform cost distributions.

In terms of game-play behavior, I found that depending on the cost distribution considered proposers end up following different strategies. More precisely, if the cost distribution is wide they adopt a pooling strategy according to which they always make offers close to the receivers' outside option under intuition. Instead, if the cost distribution is narrow, proposers switch to a differentiated strategy: under intuition and under deliberation if the game is an UG they still offer an amount equal to the receivers' intuitive outside option, while under deliberation if the game is a simplified BG they make an offer equal to the receivers' outside option in the simplified BG. Instead, receivers consistently evolve to play the same strategy which requires them to accept only offers higher than or equal to their outside option conditional on their cognitive contingency. Interestingly enough receivers' outside option does not depend solely on their cognitive contingency and the probability of a simplified BG, but it may also be affected by proposers' behavior.

According to these results, cognitive manipulations fostering subjects' reliance on intuition in experimental settings should have either no effect or slightly increase proposers' offers in the UG; moreover, these manipulations should increase receivers' rejection rates.

In terms of deliberation patterns, I found that overall receivers tend to deliberate more frequently than proposers. I interpret this as the effect of differences in strategic complexity faced by the two types of agents which in turn affects agents' expected benefits deriving from incurring into deliberation. Moreover, I found evidence of endogenous deliberation patterns: receivers' reliance on deliberation is crucially affected by proposers' behavior. This implies that in theory proposers may strategically induce receivers to deliberate more or less to their own benefit. Moreover, even if this was not the case the presence of endogenous deliberation patterns on the receivers' side has important implications in terms of cognitive manipulations in the lab. In fact, proposers' behavior could in principle exacerbate or neutralize the intended effects of cognitive manipulations on receivers.

I have also analyzed rejection rates in the UG. First, I have computed conditional rejection rates, i.e., the probability with which a receiver rejects would reject any given offer in the UG. I found that the model can generate high rejection rates conditional on receiving a given offer. Moreover, while rejection rates are in general decreasing in the offer made there are circumstances in which offering more can increase rejection rates. This happens when offering more significantly reduces receivers' probability to deliberate. Then, I have also computed overall rejection rates generated by proposers' and receivers' combined behaviors. By doing this I found rejection rates between 5% and 10%. These rejection rates are significantly below the ones observed in experimental settings. However, it is worth stressing that these rejection rates are generated without taking into account possible mistakes on the proposers' and receivers' sides. By including mistakes the differences between model predictions and experimental findings would be much smaller.

The model considered in this paper can account for several recurrent findings in experiments dealing with the UG such as generous offers made by proposers, frequent rejections of positive offers, and potentially cross-country differences in behavior in the UG. It also provides interesting predictions regarding the effects of cognitive manipulations on proposers' and receivers' behaviors. Further, the results obtained hint at potential issues associated with the implementation of cognitive manipulations in the UG given that receivers' deliberation patterns may be significantly affected by proposers' behavior. It is certainly of interest to study whether this phenomenon is actually present in experimental studies and whether it can significantly impact or even neutralize the intended effects of cognitive manipulations. If this is the case, it would be of primary importance to develop experimental setups that can take into account or sterilize the effects of receivers' endogenous deliberation patterns.

Other interesting avenues for future research are, for example, the following. First, it could be interesting to analyze whether the results are robust to different revision protocols such as imitation and myopic best reply. Second, it could be of interest to develop a similar model within a signaling setting to study more carefully how much information receivers can extract from proposers' offers and on the other side whether and to what extent can proposers strategically exploit receivers' information acquisition from the offer received. Finally, it might also be of use to replicate the setup considered in this paper within an experimental setting. To this end a preliminary step could be the definition of measures of the probability subjects expect to incur into bargaining interactions rather than ultimatum interactions in their daily life.

Chapter 5

Conclusion

In this thesis, I have presented three papers studying the evolution of prosocial behaviors with a dual process perspective.

The first paper, reported in Chapter 2, is a joint work with Ennio Bilancini and Leonardo Boncinelli. In the paper, we have studied the evolution of collaboration conceived as playing the payoff dominant convention in the Stag Hunt game. More precisely, we have analyzed how agents' cognition and the structure of interaction together affect the emergence of the payoff dominant convention.

We have considered a finite population of myopically best replying agents and a set of locations in which agents locate themselves. Agents are randomly matched to play a Stag Hunt game. Sometimes only agents staying in the same location are matched (local interactions) while other times agents are matched independently of their location choice (global interactions). Finally, there are two types of agents: fine reasoners who can distinguish local from global interactions and coarse reasoners who cannot.

By considering this setup we have found that if interactions are mostly global then selection favors conventions where agents are separated into different locations according to their mode of reasoning, with coarse reasoners playing the risk dominant action in all interactions and fine reasoners playing the risk dominant action globally and the payoff dominant action locally. If instead, interactions are mostly local, then all agents stay in the same location and there are two cases: if coarse reasoners are sufficiently numerous then all agents play the payoff dominant action both globally and locally, while if coarse reasoners are not numerous enough then they play the payoff dominant action and fine reasoners play the risk dominant action globally and the payoff dominant action locally. This implies that the co-existence of coarse and fine reasoning may favor or hamper the adoption of the payoff dominant action depending on the structure of interactions.

These results hint at an interesting interplay between cognition and the structure of interaction. In particular, depending on whether interactions are mostly local or global increased reliance on fine reasoning may be either beneficial or detrimental to the diffusion of collaboration. This in turn has interesting welfare implications. Moreover, the setting analyzed allows to derive several testable predictions regarding the effects of cognitive manipulations on the tendency to coordinate on the payoff dominant convention.

The second paper, reported in Chapter 3, is also a joint work with Ennio Bilancini and Leonardo Boncinelli. In the paper, we have analyzed a model of co-evolution of cooperation and cognition with an endogenous interaction structure.

The model builds on the one in Bear and David G Rand (2016) and on the findings in Jagau and Veelen (2017) and introduces an endogenous interaction structure. More precisely, We have considered a population of agents and a set of locations in which agents locate themselves. Over time agents staying in the same location are randomly matched to play a prisoners' dilemma which can be either one-shot and anonymous or infinitely repeated. However, agents need to engage in costly deliberation to recognize the type of interaction they are facing. The presence of different types of games together with incomplete information generates a trade-off between intuitively defecting – which is the dominant strategy in case the interaction is one-shot and anonymous – and intuitively cooperating – which is optimal in infinitely repeated interactions. In the evaluation of such a trade-off, agents must also consider that different strategies imply different optimal levels of deliberation.

We have shown that at most three types of absorbing sets exist: (i) an intuitive defection set in which agents defect under intuition, never deliberate, and are indifferent with respect to their location choice, (ii) dual process cooperation states in which agents cooperate under intuition, deliberate with positive probability and all agents stay in the same location, and (iii) depending on the actual distribution of deliberation costs none, one, or more kinds of dual process defection absorbing states in which all agents defect under intuition, deliberate with positive probability and stay in the same location. Afterward, we have shown that dual process cooperation states are the only stochastically stable states in the entire parameter space in which they are absorbing states. This in turn implies that in our model dual process cooperation is favored by evolution in a larger parameter space than in Bear and David G Rand (2016) and, so, that an endogenous interaction structure promotes cooperation.

Finally, in Chapter 4 I have presented a model for the evolution of fair splits and harsh rejections in the Ultimatum Game.

In the model, reinforcement learning proposers and receivers interact over time in Ultimatum Games and simplified Bargaining Games, but they cannot recognize which one they are playing unless they incur costly deliberation.

On the one hand, I found that receivers significantly change their reliance on deliberation depending on the amount they get offered by proposers, and in the case of specific offers, they may be willing to incur relatively high costs of deliberation. Moreover, receivers consistently evolve to play the same strategy which requires them to accept only offers higher than or equal to their outside option conditional on their cognitive contingency. On the other hand, I found that proposers deliberate less frequently than receivers and they tend to keep fixed their deliberation patterns. However, depending on the cost distribution considered proposers end up following different strategies. More precisely, if the cost distribution is wide they adopt a pooling strategy according to which they always make offers close to the outside option of receivers under intuition. Instead, if the cost distribution is narrow they switch to a differentiated strategy: under intuition and under deliberation if the game is an UG they still offer an amount equal to the receivers' intuitive outside option, while under deliberation if the game is a simplified BG they make an offer equal to receivers' outside option in the simplified BG.

By focusing on agents' behaviors conditional on playing an UG, I found that the model can account for fair offers and harsh rejections. Moreover, these results provide several predictions regarding the effects of cognitive manipulations on proposers' and receivers' behavior. In addition, the finding that receivers' deliberation patterns can be significantly influenced by proposers' behavior suggests interesting strategic considerations but also hints that proposers' behavior might counteract the intended effects of cognitive manipulations which in turn could explain why the experimental literature has frequently found mixed results.

Even though each of the three papers presented focuses on a different prosocial behavior, they all contribute to the relevant literature by suggesting possible mechanisms affecting the emergence of the prosocial behavior considered. Moreover, they all do this with a dual process approach. Finally, each paper opens up interesting avenues for future research.

Appendix A

Chapter 2: Proof of Theorem 1

Before providing the proof of Theorem 1, we introduce some useful concepts. This is necessary because of the location choice setting we analyse. More precisely, even if we allow for a large enough population so that in global interactions each agent becomes negligible there always exist states of the system in which really few agents stay in a given location. In such states the evaluation of which is the best location – and, thus, the best reply – to choose can be different from one agent to another. For this reason we need to consider location- and agent-specific measures in order to describe the behavior of each agent and, consequently, of the system.

For convenience, in the following we will set $A = 0$ and $B = 1$ so that, for example, $l_{nt} = 0$ means that at time t agent n plays action A if the interaction is local. Moreover, let $\langle \ell \rangle_t$ denote the number of agents staying in location ℓ at time t and let $I_{n\ell t}$ be an indicator function that takes value one if agent n stays in location ℓ at time t, i.e., if $\ell_{nt} = \ell$, and zero otherwise.

Let $\Lambda^A_{\ell t}$ be the probability that at time t a randomly selected agent in location ℓ plays action A if the interaction is local. Then,

$$\Lambda^A_{\ell t} = \begin{cases} \dfrac{1}{\langle \ell \rangle_t} \displaystyle\sum_{n:\ell_{nt}=\ell} (1 - l_{nt}) & \text{if } \langle \ell \rangle_t > 0 \\ 0 & \text{if } \langle \ell \rangle_t = 0 \end{cases}$$

and let $\Lambda^B_{\ell t}$ be the probability that at time t a randomly selected agent in location ℓ plays action B if the interaction is local. Then,

$$\Lambda^B_{\ell t} = \begin{cases} \dfrac{1}{\langle \ell \rangle_t} \displaystyle\sum_{n:\ell_{nt}=\ell} l_{nt} & \text{if } \langle \ell \rangle_t > 0 \\ 0 & \text{if } \langle \ell \rangle_t = 0 \end{cases}$$

Note that:

- If $\langle \ell \rangle_t > 0$ (location ℓ is non-empty at time t), then $\Lambda^A_{\ell t} = 1 - \Lambda^B_{\ell t}$;
- If $\langle \ell \rangle_t = 0$ (location ℓ is empty at time t), then $\Lambda^A_{\ell t} = \Lambda^B_{\ell t} = 0$.

Let $\lambda^A_{n\ell t}$ denote agent n's probability at time t to be matched with a randomly selected agent $n' \neq n$ staying in location ℓ and playing action A if the interaction is local (if $\ell_{nt} \neq \ell$, this is the probability in case agent n moved into location ℓ). Then,

$$\lambda^A_{n\ell t} = \begin{cases} \dfrac{\langle \ell \rangle_t \Lambda^A_{\ell t} - (1 - l_{nt}) I_{n\ell t}}{\langle \ell \rangle_t - I_{n\ell t}} & \text{if } \langle \ell \rangle_t - I_{n\ell t} > 0 \\ 0 & \text{if } \langle \ell \rangle_t - I_{n\ell t} = 0 \end{cases}$$

Let $\lambda^B_{n\ell t}$ denote agent n's probability at time t to be matched with a randomly selected agent $n' \neq n$ staying in location ℓ and playing action B if the interaction is local (again if $\ell_{nt} \neq \ell$, this is the probability if agent n moved into location ℓ). Then,

$$\lambda^B_{n\ell t} = \begin{cases} \dfrac{\langle \ell \rangle_t \Lambda^B_{\ell t} - l_{nt} I_{n\ell t}}{\langle \ell \rangle_t - I_{n\ell t}} & \text{if } \langle \ell \rangle_t - I_{n\ell t} > 0 \\ 0 & \text{if } \langle \ell \rangle_t - I_{n\ell t} = 0 \end{cases}$$

Note that the following relations holds:

- if $\ell_{nt} \neq \ell$ (agent n is not staying in location ℓ), then $\lambda^A_{n\ell t} = \Lambda^A_{\ell t}$ and $\lambda^B_{n\ell t} = \Lambda^B_{\ell t}$;
- if $\langle \ell \rangle_t - I_{n\ell t} > 0$ and so (i) location ℓ is non-empty and (ii) if agent n stays in location ℓ then agent n is not alone in location ℓ, then $\lambda^A_{n\ell t} = 1 - \lambda^B_{n\ell t}$ (and a fortiori $\Lambda^A_{\ell t} = 1 - \Lambda^B_{\ell t}$);
- if $\ell_{nt} = \ell$ (agent n stays in location ℓ), $\langle \ell \rangle_t - I_{n\ell t} > 0$ (agent n is not alone in location ℓ), and $l_{nt} = 0$ (agent n plays action A in local interactions), then $\lambda^B_{n\ell t} \geq \Lambda^B_{\ell t}$ with strict inequality if $\Lambda^B_{\ell t} > 0$. Moreover, $\lambda^B_{n\ell t} \geq \lambda^B_{n'\ell t}$ for every $n' \neq n$;
- if $\ell_{nt} = \ell$ (agent n stays in location ℓ), $\langle \ell \rangle_t - I_{n\ell t} > 0$ (agent n is not alone in location ℓ), and $l_{nt} = 1$ (agent n plays action B in local interactions), then $\lambda^B_{n\ell t} \leq \Lambda^B_{\ell t}$ with strict inequality if $\Lambda^B_{\ell t} < 1$. Moreover, $\lambda^B_{n\ell t} \leq \lambda^B_{n'\ell t}$ for every $n' \neq n$.

Let Γ_t be the probability at time t that an agent plays action B if the interaction is global. Then,

$$\Gamma_t = \frac{1}{N} \sum_{n=1}^{N} g_{nt}$$

where $g_{nt} = 1$ if agent n at time t plays the payoff dominant action B under global interactions, while $g_{nt} = 0$ if agent n at time t plays the risk dominant action A under global interactions.

Moreover, let γ_{nt} denote the probability at time t that agent n is matched with another agent $n' \neq n$ playing action B if the interaction is global. Then,

$$\gamma_{nt} = \frac{N\Gamma_t - g_{nt}}{N-1}$$

If $N \geq 2$, γ_{nt} is always well defined. Note also that $1 - \gamma_{nt}$ denotes the probability at time t that agent n is matched with another agent playing action A if the interaction is global. Moreover, by definition:

- if $g_{nt} = 1$ (agent n plays action B in global interactions), then $\gamma_{nt} \leq \Gamma_t$ with strict inequality if $\Gamma_t < 1$. Moreover, $\gamma_{nt} \leq \gamma_{n't}$ for every $n' \neq n$;

- if $g_{nt} = 0$ (agent n plays action A in global interactions), then $\gamma_{nt} \geq \Gamma_t$ with strict inequality if $\Gamma_t > 0$. Moreover, $\gamma_{nt} \geq \gamma_{n't}$ for every $n' \neq n$.

These definitions allow us to define a player's expected payoff associated to a given strategy and to determine an agent's best reply in a generic state of the system. More precisely, we can express agent n's expected payoff associated to strategy (ℓ, x, y) at time t as follows:

$$\begin{aligned}\pi_{nt}(\ell, x, y) = \; & p\big[\lambda^B_{n\ell t}\big(xb + (1-x)c\big) + \lambda^A_{n\ell t}\big(xd + (1-x)a\big)\big] + \dots \\ & \dots + (1-p)\big[\gamma_{nt}\big(yb + (1-y)c\big) + (1-\gamma_{nt})\big(yd + (1-y)a\big)\big]\end{aligned}$$

This generic expression for the expected payoff of an agent allows us to easily derive the following conclusion: if $N \geq 2$, then a strategy, say (ℓ', x', y'), prescribing to move into an empty location or to stay alone in a location is never a best reply to the current state of the system. In fact, there always exists at least one strategy providing a strictly higher payoff which is a strategy (ℓ'', x', y') such that $\langle \ell'' \rangle_t - I_{n\ell t} > 0$: such alternative strategy implies the same expected payoff in global interactions, but a strictly higher expected payoff in local interactions. Note also that if strategy (ℓ, x, y) does not imply being alone in a location, then we can substitute $\lambda^A_{n\ell t}$ with $1 - \lambda^B_{n\ell t}$.

A fine reasoner $j \in \mathcal{F}$ will adopt with positive probability a strategy prescribing to play action B in global interactions if $\pi_{jt}(\ell, x, B) \geq \pi_{jt}(\ell, x, A)$ holds for some $\ell \in \mathcal{L}$, $x \in \{A, B\}$ and, consequently, if

$$\gamma_{jt} \geq \alpha \tag{A.1}$$

otherwise it will adopt a strategy prescribing to play action A in global interactions.

Note that the optimal location choice crucially depends on the action played in local interactions. More precisely, if an agent decides to play action B in local interactions, then it will find it optimal to locate itself in the non-empty location with maximal $\lambda^B_{n\ell t}$. In fact, $\pi_{jt}(\ell', B, y) \geq \pi_{jt}(\ell'', B, y)$ if and only if $\lambda^B_{j\ell' t} \geq \lambda^B_{j\ell'' t}$. Instead, if an agent decides to play action A in local interactions, then it will find it optimal to locate itself in the non-empty location with maximal $\lambda^A_{j\ell t}$ or equivalently in the non-empty location with minimal $\lambda^B_{j\ell t}$.

Let ℓ^B be the non-empty location with maximal $\lambda^B_{j\ell t}$ and let ℓ^A be the non-empty location with minimal $\lambda^B_{j\ell t}$. Then, fine reasoner j will adopt with positive probability a strategy prescribing to play action B in local interactions if $\pi_{jt}(\ell^B, B, y) \geq \pi_{jt}(\ell^A, A, y)$ and so if:

$$\lambda^B_{j\ell^B t}\frac{b-d}{a-d+b-c} + \lambda^B_{j\ell^A t}\frac{a-c}{a-d+b-c} \geq \alpha \tag{A.2}$$

Note that the LHS of Equation (A.2) is a weighted average of $\lambda^B_{j\ell^B t}$ and $\lambda^B_{j\ell^A t}$ and fine reasoner j will be willing adopt with positive probability a strategy prescribing to play action B in local interactions if and only if such weighted average is larger than or equal to α. In general, this is a more demanding condition than $\lambda^B_{j\ell^B t} \geq \alpha$. In fact, it must not only hold $\lambda^B_{j\ell^B t} \geq \alpha$ but also there must be no non-empty location with $\lambda^B_{j\ell^A t}$ sufficiently low to make it more convenient to play action A in such location. In other words, we may say that Equation (A.2) is an optimality condition for playing action B in local interactions discounted by the opportunity-cost of playing locally action A. In fact, by exploiting $\lambda^B_{j\ell^A t} = 1 - \lambda^A_{j\ell^A t}$ we can rewrite Equation (A.2) as:

$$\lambda^B_{j\ell^B t}b + (1 - \lambda^B_{j\ell^B t})d \geq \lambda^A_{j\ell^A t}a + (1 - \lambda^A_{j\ell^A t})c$$

which requires that the expected payoff of playing action B in location ℓ^B is larger than the opportunity-cost of playing action A in location ℓ^A.

Finally, note that if there is only a single non-empty location ℓ' then $\ell^B = \ell^A = \ell'$ and, consequently, Equation (A.2) simplifies to $\lambda^B_{j\ell' t} \geq \alpha$.

The location choice by coarse reasoners follows the same logic as the one by fine reasoners: agents locate themselves in the location that allows them to maximize local coordination given the action they decide to play in local (and global) interactions. Let ℓ^B be the non-empty location with maximal $\lambda^B_{i\ell t}$ and let ℓ^A be the non-empty location with minimal $\lambda^B_{i\ell t}$. Then, coarse reasoner i will adopt with positive probability strategy

(ℓ^B, B, B) if $\pi_{it}(\ell^B, B, B) \geq \pi_{it}(\ell^A, A, A)$ which requires:

$$p\Big(\lambda^B_{i\ell^B t}\frac{b-d}{a-d+b-c} + \lambda^B_{i\ell^A t}\frac{a-c}{a-d+b-c}\Big) + (1-p)\gamma_{it} \geq \alpha \tag{A.3}$$

otherwise coarse reasoner i will adopt strategy (ℓ^A, A, A). Note that the term on the LHS is a weighted average with weights $(p, 1-p)$ of fine reasoners' local and global optimality conditions. This implies that if both Equation (A.1) and Equation (A.2) are satisfied, then also Equation (A.3) must hold.

With these tools in hand we can now prove Theorem 1. The proof of Theorem 1 consists in proving eight distinct lemmas (Lemma A.1 to Lemma A.8) referring to points (1.1)-(1.8) of the statement of the theorem.

Lemma A.1. *All the states of the type A-AA are absorbing if* $p, q \in (0, 1)$

Proof. Assume that the system is in a state of the type A-AA. Then, every $n \in \mathcal{N}$ adopts strategy (ℓ^*, A, A). Therefore, for every agent $n \in \mathcal{N}$ it must be $\lambda^B_{n\ell^*} = 0$ and $\gamma_n = 0$. Moreover, given that location ℓ^* is the only non-empty location, (i) we can express Equation (A.2) as $\lambda_{j\ell^*} \geq \alpha$ and (ii) no agent has an incentive to change location choice as this would imply moving into an empty location.

States of the type A-AA are absorbing states if neither Equation (A.1) nor Equation (A.2) nor Equation (A.3) are satisfied. This holds for every $p, q \in (0, 1)$ because $\lambda^B_{n\ell^*} = 0$ and $\gamma_n = 0$ for every $n \in \mathcal{N}$.

Hence, all the states of the type A-AA are absorbing if $p, q \in (0, 1)$. $\square$

Lemma A.2. *If the population is large enough, then all the states of the type A-AB are absorbing if* $p \in (0, 1)$ *and* $q \in \big(\alpha, \min\big\{\frac{\alpha}{1-p}, 1\big\}\big)$.

Proof. Assume that the system is in a state of the type A-AB. Then, every $i \in \mathcal{C}$ adopts strategy (ℓ^*, A, A), while every $j \in \mathcal{F}$ adopts strategy (ℓ^*, A, B). Therefore, if the population is large enough for every agent $n \in \mathcal{N}$ it must be $\lambda^B_{n\ell^*} = 0$ and $\gamma_n = q$. Moreover, given that location ℓ^* is the only non-empty location, (i) we can express Equation (A.2) as $\lambda_{j\ell^*} \geq \alpha$ and (ii) no agent has an incentive to change location choice as this would imply moving into an empty location.

States of type A-AB are absorbing states if:

- Equation (A.1) holds strictly. This is the case if $\gamma_j = q > \alpha$;

- Equation (A.2) does not hold. This is the case as $\lambda^B_{j\ell^*} = 0 < \alpha$;
- Equation (A.3) does not hold. This is the case if $q < \frac{\alpha}{1-p}$.

Hence, if the population is large enough then all the states of the type A-AB are absorbing if $p \in (0,1)$ and $q \in \left(\alpha, \min\left\{\frac{\alpha}{1-p}, 1\right\}\right)$. $\square$

Lemma A.3. *If the population is large enough, then all the states of the type A-BA are absorbing if* $p \in (0,1)$ *and* $q \in \left(\alpha, \min\left\{\frac{\alpha}{p}, 1\right\}\right)$.

Proof. Assume that the system is in a state of the type A-BA. Then, every $i \in \mathcal{C}$ adopts strategy (ℓ^*, A, A), while every $j \in \mathcal{F}$ adopts strategy (ℓ^*, B, A). Therefore, if the population is large enough for every agent $n \in \mathcal{N}$ it must be $\lambda^B_{n\ell^*} = q$ and $\gamma_n = 0$. Moreover, given that location ℓ^* is the only non-empty location, (i) we can express Equation (A.2) as $\lambda_{j\ell^*} \geq \alpha$ and (ii) no agent has an incentive to change location choice as this would imply moving into an empty location.

States of type A-BA are absorbing if:

- Equation (A.1) does not hold. This is the case as $\gamma_j = 0 < \alpha$;
- Equation (A.2) holds strictly. This is the case if $\lambda^B_{j\ell^*} = q > \alpha$;
- Equation (A.3) does not hold. This is the case if $q < \frac{\alpha}{p}$.

Hence, if the population is large enough then all the states of the type A-BA are absorbing if $p \in (0,1)$ and $q \in \left(\alpha, \min\left\{\frac{\alpha}{p}, 1\right\}\right)$. $\square$

Lemma A.4. *If the population is large enough, then all the states of the type A/BA are absorbing if* $p \in \left(0, \frac{a-d}{b-d}\right)$ *and* $q \in (0,1)$.

Proof. Assume that the system is in a state of the type A/BA. In such a state, every $i \in \mathcal{C}$ adopts strategy (ℓ^*, A, A), while every $j \in \mathcal{F}$ adopts strategy (ℓ^{**}, B, A) with $\ell^{**} \neq \ell^*$. Therefore, if the population is large enough for every agent $n \in \mathcal{N}$ it must be $\lambda^B_{n\ell^*} = \lambda^B_{n\ell^A} = 0$, $\lambda^B_{n\ell^{**}} = \lambda^B_{n\ell^B} = 1$ and $\gamma_n = 0$. Moreover, no agent has an incentive to move into a location $\ell' \neq \ell^*$, $\ell' \neq \ell^{**}$ as this would imply moving into an empty location.

States of the type A/BA are absorbing states if:

- Equation (A.1) does not hold. This is the case as $\gamma_j = 0 < \alpha$;
- Equation (A.2) holds strictly. This is always the case because $\frac{b-d}{a-d+b-c} > \alpha = \frac{a-d}{a-d+b-c}$;

- Equation (A.3) does not hold. This is the case if $p\frac{b-d}{a-d+b-c} < \alpha$ and, so, if $p < \frac{a-d}{b-d}$.

Hence, if the population is large enough then all the states of the type A/BA are absorbing if $p \in \left(0, \frac{a-d}{b-d}\right)$ and $q \in (0, 1)$. ☐

Lemma A.5. *If the population is large enough, then all the states of the type B-AB are absorbing if* $p \in (0, 1)$ *and* $q \in \left(1-\alpha, \min\left\{\frac{1-\alpha}{p}, 1\right\}\right)$.

Proof. Assume that the system is in a state of the type B-AB. Then, every $i \in \mathcal{C}$ adopts strategy (ℓ^*, B, B), while every $j \in \mathcal{F}$ adopts strategy (ℓ^*, A, B). Therefore, if the population is large enough for every agent $n \in \mathcal{N}$ it must be $\lambda^B_{n\ell^*} = 1-q$ and $\gamma_n = 1$. Moreover, given that location ℓ^* is the only non-empty location, (i) we can express Equation (A.2) as $\lambda_{j\ell^*} \geq \alpha$ and (ii) no agent has an incentive to change location choice as this would imply moving into an empty location.

States of the type B-AB are absorbing if:

- Equation (A.1) holds strictly. This is always the case because $\gamma_j = 1 > \alpha$;

- Equation (A.2) does not hold. This is the case if $\lambda^B_{j\ell^*} = 1-q < \alpha$ and, so, $q > 1-\alpha$;

- Equation (A.3) holds strictly. This is the case if $p(1-q)+(1-p) > \alpha$ and, so, $q < \frac{1-\alpha}{p}$.

Hence, if the population is large enough then all the states of the type B-AB are absorbing if $p \in (0, 1)$ and $q \in \left(1-\alpha, \min\left\{\frac{1-\alpha}{p}, 1\right\}\right)$. ☐

Lemma A.6. *If the population is large enough, then all the states of the type B-BA are absorbing if* $p \in (0, 1)$ *and* $q \in \left(1-\alpha, \min\left\{\frac{1-\alpha}{1-p}, 1\right\}\right)$.

Proof. Assume that the system is in a state of type B-BA. Then, every $i \in \mathcal{C}$ adopts strategy (ℓ^*, B, B), while every $j \in \mathcal{F}$ adopts strategy (ℓ^*, B, A). Therefore, if the population is large enough for every agent $n \in \mathcal{N}$ it must be $\lambda^B_{n\ell^*} = 1$ and $\gamma_n = 1-q$. Moreover, given that location ℓ^* is the only non-empty location, (i) we can express Equation (A.2) as $\lambda_{j\ell^*} \geq \alpha$ and (ii) no agent has an incentive to change location choice as this would imply moving into an empty location.

States of the type B-BA are absorbing if:

- Equation (A.1) does not hold. This is the case if $\gamma_j = 1 - q < \alpha$ and, so, $q > 1 - \alpha$;
- Equation (A.2) holds strictly. This is always the case because $\lambda^B_{j\ell^*} = 1 > \alpha$;
- Equation (A.3) holds strictly. This is the case if $p + (1-p)(1-q) > \alpha$ and, so, $q < \frac{1-\alpha}{1-p}$.

Hence, if the population is large enough then all the states of the type B-BA are absorbing if $p \in (0,1)$ and $q \in \left(1-\alpha, \min\left\{\frac{1-\alpha}{1-p}, 1\right\}\right)$. □

Lemma A.7. *All the states of the type B-BB are absorbing if* $p, q \in (0,1)$.

Proof. Assume that the system is in a state of the type B-BB. Then, every $n \in \mathcal{N}$ adopts strategy (ℓ^*, B, B). Therefore, for every agent $n \in \mathcal{N}$ it must be $\lambda^B_{n\ell^*} = 1$ and $\gamma_n = 1$. Moreover, given that location ℓ^* is the only non-empty location, (i) we can express Equation (A.2) as $\lambda_{j\ell^*} \geq \alpha$ and (ii) no agent has an incentive to change location choice as this would imply moving into an empty location.

States of the type B-BB are absorbing states if both Equation (A.1), Equation (A.2), and Equation (A.3) are strictly satisfied and this hold for every $p, q \in (0,1)$ because $\lambda^B_{n\ell^*} = 1$ and $\gamma_n = 1$ for every $n \in \mathcal{N}$.

Hence, all the states of the type B-BB are absorbing if $p, q \in (0,1)$. □

Lemma A.8. *There are no absorbing sets other than the states of types A-AA, A-AB, A-BA, A/BA, B-AB, B-BA, and B-BB.*

Proof. We provide an algorithm showing that starting from any possible state the system converges with positive probability to an absorbing state among the states of types A-AA, A-AB, A-BA, A/BA, B-AB, B-BA, and B-BB. In Figure A.1 we provide a sketch of such algorithm.

Assume that the system is in a generic state and set $t = 0$. Then, two scenarios are possible: either there is at least one fine reasoner j' willing to adopt or to keep a strategy prescribing to play action B in global interactions (*Scenario 1*) or there is none (Scenario 2).

Consider first *Scenario 1*. Given that there is at least one fine reasoner, say j', willing to adopt or to keep a strategy prescribing to play action B in global interactions, then with positive probability at the end of time $t = 0$ only agent j' will be given a revision opportunity and will adopt such a strategy.

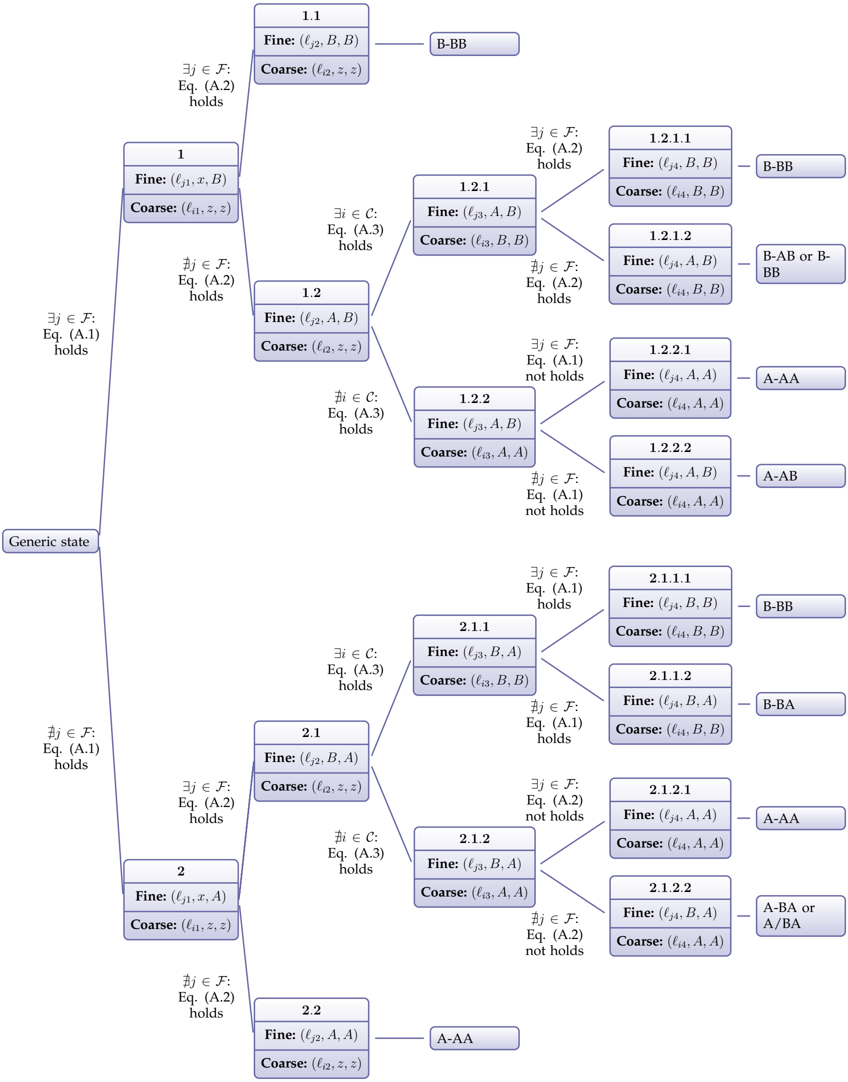


Figure A.1: Summary of the arguments used in the proof of Lemma A.8.

At time $t = 1$ it must be $\gamma_{j'1} = \gamma_{j'0} \geq \alpha$ as at time $t = 0$ Equation (A.1) was satisfied. Moreover, given that agent j' now plays action B in global interactions it must also be $\gamma_{j'1} \leq \gamma_{j1}$ for every $j \neq j'$. But then, for every other fine reasoner Equation (A.1) must hold and so every fine reasoner must be willing to adopt or to keep a strategy prescribing to play action B in global interactions. But then with positive probability at the end of time $t = 1$ every fine reasoner $j \neq j'$ will be given a revision opportunity and will adopt a strategy prescribing to play action B in global interactions.

At the beginning of time $t = 2$ the system will be in a state in which every fine reasoner adopts a strategy prescribing to play action B in global interactions. Then, two scenarios are possible: either there is at least one fine reasoner j' willing to adopt or to keep strategy $(\ell_{j'3}, B, B)$ for some $\ell_{j'3} \in \mathcal{L}$ (*Scenario 1.1*) or there is none (*Scenario 1.2*).

Consider first *Scenario 1.1*. Given that there is at least one fine reasoner j' willing to adopt or to keep strategy $(\ell_{j'3}, B, B)$ for some $\ell_{j'3} \in \mathcal{L}$, then with positive probability at the end of time $t = 2$ only agent j' will be given a revision opportunity and will adopt strategy $(\ell_{j'3}, B, B)$.

At the beginning of time $t = 3$ $(\lambda_{j'13}, \dots, \lambda_{j'L3}) = (\lambda_{j'12}, \dots, \lambda_{j'L2})$ must be such that Equation (A.2) is satisfied as otherwise adopting strategy $(\ell_{j'3}, B, B)$ would have not been a best reply for agent j'. Moreover, given that agent j' now adopts strategy $(\ell_{j'3}, B, B)$ it must also be $\lambda^B_{j'\ell_{j'3}3} \leq \Lambda^B_{\ell_{j'3}3}$ and $\lambda^B_{j'\ell 3} = \Lambda^B_{\ell 3}$ for any $\ell \neq \ell_{j'3}$. If at time $t = 3$ there is at least one fine reasoner $j \neq j'$ adopting a strategy prescribing to play action A in local interactions, then it must be

$$(\lambda^B_{j13}, \dots, \lambda^B_{j\ell 3}, \dots, \lambda^B_{jL3}) \geq (\lambda^B_{j'13}, \dots, \lambda^B_{j'\ell 3}, \dots, \lambda^B_{j'L3})$$

But then, Equation (A.2) must hold for every fine reasoner j adopting a strategy prescribing to play action A in local interactions and, consequently, every such fine reasoner must be willing to adopt with positive probability strategy (ℓ_{j4}, B, B). But then with positive probability at the end of time $t = 3$ every fine reasoner adopting a strategy prescribing to play action A in local interactions will be given a revision opportunity and will adopt strategy (ℓ_{j4}, B, B).

At the beginning of time $t = 4$ the system will be in a state in which every fine reasoner j adopts a strategy of type (ℓ_{j4}, B, B). In such a state every coarse reasoner i adopting a strategy of the type (ℓ_{i4}, A, A) for some $\ell_{i4} \in \mathcal{L}$ must be willing to adopt strategy (ℓ_{i5}, B, B) for some

$\ell_{i5} \in \mathcal{L}$. In fact for every such coarse reasoner it must be

$$(\lambda^B_{i14}, \ldots, \lambda^B_{i\ell 4}, \ldots, \lambda^B_{iL4}, \gamma_{i4}) \geq (\lambda^B_{j14}, \ldots, \lambda^B_{j\ell 4}, \ldots, \lambda^B_{jL4}, \gamma_{j4})$$

where j is a generic fine reasoner. But given that $(\lambda^B_{j14}, \ldots, \lambda^B_{jL4}, \gamma_{j4})$ is such that both Equation (A.1) and Equation (A.2) are satisfied, then a fortiori $(\lambda^B_{i14}, \ldots, \lambda^B_{iL4}, \gamma_{i4})$ must be such that Equation (A.3) holds. But then, with positive probability at the end of time $t = 4$ every coarse reasoner adopting a strategy of type (ℓ_{i4}, A, A) will be given a revision opportunity and will adopt strategy (ℓ_{i5}, B, B).

At the beginning of time $t = 5$ the system will be in a state in which all agents play action B in both local and global interactions. But then, the system will be in an absorbing state of the type B-BB.

Consider now *Scenario 1.2* in which there is no fine reasoner j' willing to adopt or to keep strategy $(\ell_{j'3}, B, B)$ for some $\ell_{j'3} \in \mathcal{L}$. But then every fine reasoner j must be willing to adopt or to keep strategy (ℓ_{j3}, A, B) for some $\ell_{j3} \in \mathcal{L}$. But then, with positive probability at the end of time $t = 2$ every fine reasoner will be given a revision opportunity and will adopt such strategy.

At time $t = 3$ the system will be in a state in which each fine reasoner j adopts a strategy of type (ℓ_{j3}, A, B). Then two scenarios are possible: either there is at least one coarse reasoner, say i', willing to adopt or to keep strategy $(\ell_{i'4}, B, B)$ for some $\ell_{i'4} \in \mathcal{L}$ (*Scenario 1.2.1*) or there is none (*Scenario 1.2.2*).

Consider first *Scenario 1.2.1*. Given that there is at least one coarse reasoner i' willing to adopt or to keep strategy $(\ell_{i'4}, B, B)$ for some $\ell_{i'4} \in \mathcal{L}$, then with positive probability at the end of time $t = 3$ only agent i' will be given a revision opportunity and will adopt strategy $(\ell_{i'4}, B, B)$.

At the beginning of time $t = 4$ for every coarse reasoner i adopting a strategy prescribing to play action A in both local and global interactions it must be

$$(\lambda^B_{i14}, \ldots, \lambda^B_{i\ell 4}, \ldots, \lambda^B_{iL4}, \gamma_{i4}) \geq (\lambda^B_{i'14}, \ldots, \lambda^B_{i'\ell 4}, \ldots, \lambda^B_{i'L4}, \gamma_{i'4})$$

But, given that $(\lambda^B_{i'14}, \ldots, \lambda^B_{i'\ell 4}, \ldots, \lambda^B_{i'L4}, \gamma_{i'4})$ is such that Equation (A.3) is satisfied, then the same must be true also for $(\lambda^B_{i14}, \ldots, \lambda^B_{i\ell 4}, \ldots, \lambda^B_{iL4}, \gamma_{i4})$. Consequently, each coarse reasoner i adopting a strategy of type (ℓ_{i4}, A, A) must be willing to adopt with positive probability strategy (ℓ_{i5}, B, B) for some $\ell_{i5} \in \mathcal{L}$. But then, with positive probability at the end of time $t = 4$ every coarse reasoner i adopting a strategy prescribing to play A in both local interactions will be given a revision opportunity and will adopt strategy (ℓ_{i5}, B, B).

At the beginning of time $t = 5$ the system will be in a state in which every coarse reasoner i adopts a strategy of type (ℓ_{i5}, B, B) while every fine reasoner j adopts a strategy of type (ℓ_{j5}, A, B). Then two scenarios are possible. If there is at least one fine reasoner j' willing to adopt strategy $(\ell_{j'6}, B, B)$ for some $\ell_{j'6} \in \mathcal{L}$ (*Scenario 1.2.1.1*), then we are back to *Scenario 1.1* and the system will move with positive probability into an absorbing state of type B-BB. If, instead, there is no fine reasoner j' willing to adopt strategy $(\ell_{j'6}, B, B)$ for some $\ell_{j'6} \in \mathcal{L}$ (*Scenario 1.2.1.2*), then every fine reasoner j must be willing to keep a strategy of type (ℓ_{j6}, A, B) for some $\ell_{j6} \in \mathcal{L}$. Let ℓ^* be the location with minimum $\Lambda_{\ell 5}$ and ℓ^{**} be the location with maximum $\Lambda_{\ell 5}$, then with positive probability at the end of time $t = 5$ all coarse reasoners i such that $\ell_{i5} \neq \ell^*$ and all fine reasoners j such that $\ell_{j5} \neq \ell^{**}$ will be given a revision opportunity and each coarse reasoner given a revision opportunity will adopt strategy (ℓ^*, B, B) while every fine reasoner given a revision opportunity will adopt strategy (ℓ^{**}, A, B).

At the beginning of time $t = 6$ the system will be a state in which every coarse reasoner i adopts strategy (ℓ^*, B, B) while every fine reasoner j adopts strategy (ℓ^{**}, A, B). Then, two scenarios are possible: if $\ell^* \neq \ell^{**}$ then every fine reasoner j must now be willing to adopt strategy (ℓ^*, B, B) and, consequently, the system will reach an absorbing state of type B-BB; if, instead, $\ell^* = \ell^{**}$ the system will have reached an absorbing state of type B-AB.

Consider now *Scenario 1.2.2* in which there is no coarse reasoner i' willing to adopt or to keep strategy $(\ell_{i'4}, B, B)$ for some $\ell_{i'4} \in \mathcal{L}$. But then every coarse reasoner i must be willing to adopt a strategy of type (ℓ_{i4}, A, A) for some $\ell_{i4} \in \mathcal{L}$. But then, with positive probability at the end of time $t = 3$ every coarse reasoner will be given a revision opportunity and will adopt such strategy.

At the beginning of time $t = 4$ the system will be in a state in which every coarse reasoner i adopts strategy (ℓ_{i4}, A, A) while every fine reasoner j adopts strategy (ℓ_{j4}, A, B). Then two scenarios are possible: either there is at least one fine reasoner, say j', willing to adopt strategy $(\ell_{j'5}, A, A)$ for some $\ell_{j'5} \in \mathcal{L}$ (*Scenario 1.2.2.1*) or there is none (*Scenario 1.2.2.2*).

Consider first *Scenario 1.2.2.1*. Given that there is at least one fine reasoner j' willing to adopt strategy $(\ell_{j'5}, A, A)$ for some $\ell_{j'5} \in \mathcal{L}$, then with positive probability at the end of time $t = 4$ only agent j' will be given a revision opportunity and will adopt strategy $(\ell_{j'5}, A, A)$.

At time $t = 5$ every fine reasoner $j \neq j'$ must be willing to adopt strategy (ℓ_{j6}, A, A) as

$$(\lambda^B_{j15}, \ldots, \lambda^B_{j\ell 5}, \ldots, \lambda^B_{jL5}) \leq (\lambda^B_{j'15}, \ldots, \lambda^B_{j'\ell 5}, \ldots, \lambda^B_{j'L5})$$

and $(\lambda^B_{j'15}, \ldots, \lambda^B_{j'\ell 5}, \ldots, \lambda^B_{j'L5})$ is such that Equation (A.2) holds. Then, with positive probability at the end of time $t = 5$ every fine reasoner j will be given a revision opportunity and will adopt strategy (ℓ_{j5}, A, A).

At the beginning of time $t = 6$ the system will be in a state in which every agent adopts a strategy prescribing to play action A in both local and global interactions. But then the system will reach an absorbing state of type A-AA.

Consider now *Scenario 1.2.2.2* in which there is no fine reasoner j' willing to adopt strategy $(\ell_{j'5}, A, A)$ for some $\ell_{j'5} \in \mathcal{L}$. But then every fine reasoner j must be willing to keep a strategy of type (ℓ_{j5}, A, B) for some $\ell_{j5} \in \mathcal{L}$. Given that all agents play action A in local interactions, every agent must be indifferent between staying in his current location and moving into a non-empty location. But then, with positive probability at the end of time $t = 4$ every agent will be given a revision opportunity and every coarse reasoner will adopt strategy (ℓ^*, A, A) while every fine reasoner will adopt strategy (ℓ^*, A, B). Therefore, the system will have reached an absorbing state of type A-AB.

Consider now *Scenario 2* in which there is no fine reasoner willing to adopt or to keep a strategy prescribing to play action B in global interactions, then every fine reasoner must be willing to adopt a strategy prescribing to play action A in global interactions. But then with positive probability at the end of time $t = 0$ every fine reasoner will be given a revision opportunity and will adopt such a strategy.

At time $t = 1$ the system will be in a state in which every fine reasoner adopts a strategy prescribing to play action A in global interactions. Then, two scenarios are possible: either there is at least one fine reasoner, say j', willing to adopt or to keep strategy $(\ell_{j'2}, B, A)$ for some $\ell_{j'2} \in \mathcal{L}$ (*Scenario 2.1*) or there is none (*Scenario 2.2*).

Consider first *Scenario 2.1*. Given that there is at least one fine reasoner j' willing to adopt or to keep strategy $(\ell_{j'2}, B, A)$ for some $\ell_{j'2} \in \mathcal{L}$, with positive probability at the end of time $t = 1$ only agent j' will be given a revision opportunity and will adopt strategy $(\ell_{j'2}, B, A)$.

At the beginning of time $t = 2$ it must be $\lambda^B_{j'\ell_{j'2}2} = \lambda^B_{j'\ell_{j'2}1} \geq \alpha$ as otherwise $(\ell_{j'2}, B, A)$ would have not been a best reply; moreover, given that agent j' now adopts strategy $(\ell_{j'2}, B, A)$ it must also be $\lambda^B_{j'\ell_{j'2}2} \leq \Lambda^B_{\ell_{j'2}2}$ and $\lambda^B_{j'\ell 2} = \Lambda^B_{\ell 2}$ for any $\ell \neq \ell_{j'2}$. If at time $t = 2$ there is at least one fine reasoner j adopting strategy (ℓ_{j2}, A, A), then such fine reasoners must be willing to adopt with positive probability strategy (ℓ_{j3}, B, A) for some $\ell_{j3} \in \mathcal{L}$ as

$$(\lambda^B_{j12}, \dots, \lambda^B_{j\ell 2}, \dots, \lambda^B_{jL2}) \geq (\lambda^B_{j'12}, \dots, \lambda^B_{j'\ell 2}, \dots, \lambda^B_{j'L2})$$

and $(\lambda^B_{j'12}, \dots, \lambda^B_{j'\ell 2}, \dots, \lambda^B_{j'L2})$ is such that Equation (A.2) holds. But then with positive probability at the end of time $t = 2$ every fine reasoner j adopting strategy (ℓ_{j2}, A, A) will be given a revision opportunity and will adopt strategy (ℓ_{j3}, B, A) with $\ell_{j3} \in \mathcal{L}$.

At the beginning of time $t = 3$ the system will be in a state in which every fine reasoner j adopts a strategy of type (ℓ_{j3}, B, A). Then, two scenarios are possible: either there is at least one coarse reasoner, say i', willing to adopt or to keep strategy $(\ell_{i'4}, B, B)$ with $\ell_{i'4} \in \mathcal{L}$ (*Scenario 2.1.1*) or there is none (*Scenario 2.1.2*).

Consider first *Scenario 2.1.1*. Given that there is at least one coarse reasoner i' willing to adopt or to keep strategy $(\ell_{i'4}, B, B)$ with $\ell_{i'4} \in \mathcal{L}$, then with positive probability at the end of time $t = 3$ only agent i' will be given a revision opportunity and will adopt strategy $(\ell_{i'3}, B, B)$.

At time $t = 4$ every coarse reasoner i adopting strategy (ℓ_{i4}, A, A) must be willing to adopt a strategy of type (ℓ_{i4}, B, B) for some $\ell_{i4} \in \mathcal{L}$ as

$$(\lambda^B_{i14}, \dots, \lambda^B_{i\ell 4}, \dots, \lambda^B_{iL4}, \gamma_{i4}) \geq (\lambda^B_{i'14}, \dots, \lambda^B_{i'\ell 4}, \dots, \lambda^B_{i'L4}, \gamma_{i'4})$$

and $(\lambda^B_{i'14}, \dots, \lambda^B_{i'\ell 4}, \dots, \lambda^B_{i'L4}, \gamma_{i'4})$ is such that Equation (A.3) holds. But then, with positive probability at the end of time $t = 4$ every coarse reasoner i adopting strategy (ℓ_{i4}, A, A) will be given a revision opportunity and will adopt strategy (ℓ_{i5}, B, B).

At the beginning of time $t = 5$ the system will be in a state in which every coarse reasoner i adopts strategy (ℓ_{i5}, B, B) while every fine reasoner j adopts strategy (ℓ_{j5}, B, A). Then, if there is at least one fine reasoner j' willing to adopt strategy $(\ell_{j'6}, B, B)$ (*Scenario 2.1.1.1*), then we are back to *Scenario 1.1* and the system will reach an absorbing state of type B-BB. Instead, If there is no fine reasoner j' willing to adopt strategy $(\ell_{j'6}, B, B)$ (*Scenario 2.1.1.2*), then given that every agent is adopting a strategy prescribing to play action B in local interactions each agent is indifferent between every non-empty location choice. But then, with positive probability at the end of time $t = 5$ every agent will be given a revision opportunity and all coarse reasoners will adopt strategy (ℓ^*, B, B) while every fine reasoner will adopt strategy (ℓ^*, B, A). Consequently the system will reach an absorbing state of type B-BA.

Consider now *Scenario 2.1.2* in which there is no coarse reasoner i' willing to adopt or to keep strategy $(\ell_{i'4}, B, B)$ for some $\ell_{i'4} \in \mathcal{L}$, then every coarse reasoner i must be willing to adopt strategy (ℓ_{i4}, A, A) for some $\ell_{i4} \in \mathcal{L}$. But then, with positive probability at the end of time $t = 3$ every coarse reasoner i will be given a revision opportunity and will adopt strategy (ℓ_{i4}, A, A).

At the beginning of time $t = 4$ the system will be in a state in which every coarse reasoner i adopts strategy (ℓ_{i4}, A, A) while every fine reasoner j adopts strategy (ℓ_{j4}, B, A). Then, if there is at least one fine reasoner j' willing to adopt strategy $(\ell_{j'5}, A, A)$ (*Scenario 2.1.2.1*), then we are back to *Scenario 1.2.2.1* and the system will reach an absorbing state of type A-AA. Instead, If there is no fine reasoner j' willing to adopt strategy $(\ell_{j'5}, A, A)$ (*Scenario 2.1.2.2*), then every fine reasoner j must be willing to keep strategy (ℓ_{j5}, B, A). Let ℓ^* be the location with minimum $\Lambda_{\ell 5}$ and ℓ^{**} be the location with maximum $\Lambda_{\ell 5}$ (in case of multiple maxima [minima] the maximal [minimal] location is randomly selected among them), then with positive probability at the end of time $t = 4$ all coarse reasoners i such that $\ell_{i5} \neq \ell^*$ and all fine reasoners j such that $\ell_{j5} \neq \ell^{**}$ will be given a revision opportunity and each coarse reasoner given a revision opportunity will adopt strategy (ℓ^*, A, A) while every fine reasoner given a revision opportunity will adopt strategy (ℓ^{**}, B, A).

At the beginning of time $t = 6$ the system will be in an absorbing state either of type A/BA (if $\ell^* \neq \ell^{**}$) or of type A-BA.

Consider now *Scenario 2.2* in which there is no fine reasoner j' willing to adopt or to keep strategy $(\ell_{j'2}, B, A)$ for some $\ell_{j'2} \in \mathcal{L}$. But then every fine reasoner j must be willing to adopt a strategy of type (ℓ_{j2}, A, A) for some $\ell_{j2} \in \mathcal{L}$. But then, with positive probability at the end of time $t = 1$ every fine reasoner j will be given a revision opportunity and will adopt strategy (ℓ_{j2}, A, A).

At the beginning of time $t = 2$ the system will be in a state in which every fine reasoner j adopts strategy (ℓ_{j2}, A, A). In such a state every coarse reasoner i adopting a strategy of type (ℓ_{i2}, B, B) for some $\ell_{i2} \in \mathcal{L}$ must be willing to adopt strategy (ℓ_{i3}, A, A) for some $\ell_{i3} \in \mathcal{L}$. In fact for every such coarse reasoner it must be

$$(\lambda^B_{i12}, \ldots, \lambda^B_{i\ell 2}, \ldots, \lambda^B_{iL2}, \gamma_{i2}) \leq (\lambda^B_{j12}, \ldots, \lambda^B_{j\ell 2}, \ldots, \lambda^B_{j'L2}, \gamma_{j2})$$

where j is a generic fine reasoner. But given that $(\lambda^B_{j12}, \dots, \lambda^B_{j'L2}, \gamma_{j2})$ is such that neither Equation (A.1) nor Equation (A.2) are satisfied, then we can conclude that $(\lambda^B_{i12}, \dots, \lambda^B_{iL2}, \gamma_{i2})$ must be such that Equation (A.3) does not hold. But then, with positive probability at the end of time $t = 2$ every coarse reasoner i adopting a strategy of type (ℓ_{i2}, B, B) will be given a revision opportunity and will adopt strategy (ℓ_{i3}, A, A).

At the beginning of time $t = 3$ the system will be in a state in which every agent adopts a strategy prescribing to play action A in local interactions. But then, every agent must be indifferent between staying in his current location and moving into a non-empty location. But then, with positive probability at the end of time $t = 3$ every agent will be given a revision opportunity and will adopt strategy (ℓ^*, A, A). Therefore, the system will have reached an absorbing state of type A-AA.

Hence, there are no absorbing sets other than the states of types A-AA, A-AB, A-BA, A/BA, B-AB, B-BA, and B-BB.

$\square$

Theorem 1. *If the population is large enough, then all the states of the following types are absorbing:*

(1.1) A-AA, if $p, q \in (0, 1)$;

(1.2) A-AB, if $p \in (0, 1)$ *and* $q \in \left(\alpha, \min\left\{\frac{\alpha}{1-p}, 1\right\}\right)$;

(1.3) A-BA, if $p \in (0, 1)$ *and* $q \in \left(\alpha, \min\left\{\frac{\alpha}{p}, 1\right\}\right)$;

(1.4) A/BA, if $p \in \left(0, \frac{a-d}{b-d}\right)$ *and* $q \in (0, 1)$;

(1.5) B-AB, if $p \in (0, 1)$ *and* $q \in \left(1-\alpha, \min\left\{\frac{1-\alpha}{p}, 1\right\}\right)$;

(1.6) B-BA, if $p \in (0, 1)$ *and* $q \in \left(1-\alpha, \min\left\{\frac{1-\alpha}{1-p}, 1\right\}\right)$;

(1.7) B-BB, if $p, q \in (0, 1)$;

Further:

(1.8) there are no absorbing sets other than the states of types A-AA, A-AB, A-BA, A/BA, B-AB, B-BA, and B-BB.

Proof. The theorem follows as a direct consequence of the joint consideration of results in Lemma A.1 to Lemma A.8, and in case the population size is sufficiently large to satisfy the most demanding condition on population sizes among Lemma A.2, Lemma A.3, Lemma A.4, Lemma A.5, and Lemma A.6. $\square$

Appendix B

Chapter 2: Proof of Theorem 2

Let S, S' be two sets of absorbing states. In the following, we will denote with $M(S, S')$ the minimum number of mistakes required to lead with positive probability the system into the basin of attraction of S' starting from an absorbing state belonging to S. Moreover, for ease of exposition we will refer to the set of states of a given type, say X-YZ, by simply writing X-YZ so that, for example, A/BA $\in S$ means that all the states of the type A/BA belong to the set S.

We now provide some preliminary results that are needed in order to prove Theorem 2.

Lemma B.1. *Let $S' = \{A\text{-}AA, A\text{-}AB, A\text{-}BA, B\text{-}AB\}$. If the population is large enough and states of the type A/BA are absorbing, then one mistake is never enough to lead with positive probability the system into the basin of attraction of S' starting from a state of the type A/BA, i.e., $M(A/BA, S') > 1$.*

Proof. Let $S' = \{$A-AA, A-AB, A-BA, B-AB$\}$ and consider an absorbing state of the type A/BA where agents are agglomerated according to their type. Note that, if the population is large enough, then a single mistake by any agent will have a negligible effect on the expected payoff of other agents in case they keep their current strategy and, moreover, the other agents can possibly increase their expected payoff only if the mistaken agent has moved into an empty location and coordinating locally with it provides a better payoff than their current local interactions.

Assume also that the system is in an absorbing state of the type A/BA[1]. Then, every $i \in \mathcal{C}$ adopts strategy (ℓ^*, A, A) while every $j \in \mathcal{F}$ adopts strategy (ℓ^{**}, B, A).

Table B.1 reports the expected payoff that an agent $n \in \mathcal{N}$ would obtain by adopting strategy σ_{nt+1} at time $t+1$ given that at time t the system was in an absorbing state of the type A/BA and that at the end of time t an agent $n' \in \mathcal{N}$, $n' \neq n$ has adopted by mistake strategy $\sigma_{n't}$. The table reports such expected payoff for each possible combination of $(\sigma_{nt+1}, \sigma_{n't})$ under the assumption that the population is large enough.

σ_{nt+1}	$\sigma_{n't}$			
	(ℓ', A, A)	(ℓ', A, B)	(ℓ', B, A)	(ℓ', B, B)
$\boldsymbol{(\ell^*, A, A)}$	$\boldsymbol{a}$	$\boldsymbol{a^-}$	$\boldsymbol{a}$	$\boldsymbol{a^-}$
(ℓ^{**}, A, A)	$pc + (1-p)a$	$pc + (1-p)a^-$	$pc + (1-p)a$	$pc + (1-p)a^-$
(ℓ', A, A)	a	a^-	$pc + (1-p)a$	$pc + (1-p)a^-$
(ℓ^*, A, B)	$pa + (1-p)d$	$pa + (1-p)d^+$	$pa + (1-p)d$	$pa + (1-p)d^+$
(ℓ^{**}, A, B)	$pc + (1-p)d$	$pc + (1-p)d^+$	$pc + (1-p)d$	$pc + (1-p)d^+$
(ℓ', A, B)	$pa + (1-p)d$	$pa + (1-p)d^+$	$pc + (1-p)d$	$pc + (1-p)d^+$
(ℓ^*, B, A)	$pd + (1-p)a$	$pd + (1-p)a^-$	$pd + (1-p)a$	$pd + (1-p)a^-$
$\boldsymbol{(\ell^{**}, B, A)}$	$\boldsymbol{pb + (1-p)a}$	$\boldsymbol{pb + (1-p)a^-}$	$\boldsymbol{pb + (1-p)a}$	$\boldsymbol{pb + (1-p)a^-}$
(ℓ', B, A)	$pc + (1-p)a$	$pc + (1-p)a^-$	$pb + (1-p)a$	$pb + (1-p)a^-$
(ℓ^*, B, B)	d	d^+	d	d^+
(ℓ^{**}, B, B)	$pb + (1-p)d$	$pb + (1-p)d^+$	$pb + (1-p)d$	$pb + (1-p)d^+$
(ℓ', B, B)	d	d^+	$pb + (1-p)d$	$pb + (1-p)d^+$

Table B.1: Effects of one mistake in an absorbing state of the type A/BA. Expected payoff of agent n at time $t+1$ if agent n decides to adopt strategy σ_{nt+1} given that at time t the system was in an absorbing state of the type A/BA and that at the end of time t agent $n' \neq n$ has adopted by mistake strategy $\sigma_{n't} = (\ell', x, y)$ with $\ell' \neq \ell^*, \ell^{**}$. Strategies available to coarse reasoners in gray, current strategies of coarse and fine reasoners in bold. $+$ $[-]$ indicates that the expected payoff is approximated from above [below] under the assumption of a large enough population.

---

[1] By Theorem 1 a state of the type A/BA is an absorbing state if $p \in \big(0, \frac{a-d}{b-d}\big)$. If, instead, $p \in (\frac{a-d}{b-d}, 1)$, then for every coarse reasoner $i \in \mathcal{C}$ $\pi_{it}(\ell^*, A, A) < \pi_{it}(\ell^{**}, B, B)$. But then, with positive probability at the end of the next period of time every $i \in \mathcal{C}$ will be given a revision opportunity and will adopt strategy (ℓ^{**}, B, B). Consequently, without any mistake the system will reach a state of the type B-BA and, so, the analysis in Lemma B.2 is the one of interest.

As the table shows, in most cases the expected payoff provided by the current strategy adopted by agent n, $\sigma_i = (\ell^*, A, A)$ if $i \in \mathcal{C}$ and $\sigma_j = (\ell^{**}, B, A)$ if $j \in \mathcal{F}$, is strictly higher than the one associated to alternative strategies after agent n' has adopted by mistake strategy $\sigma_{n't}$. However, some cases are worth discussing.

The population must be large enough in order to guarantee the following inequalities:

- $\pi_{it+1}(\ell^*, A, A) > \pi_{it+1}(\ell^{**}, B, B)$ for every $i \in \mathcal{C}$, $i \neq n'$ if $\sigma_{n't} \in \{(\ell', A, B), (\ell', B, B)\}$;
- $\pi_{it+1}(\ell^*, A, A) > \pi_{it+1}(\ell', B, B)$ for every $i \in \mathcal{C}$, $i \neq n'$ if $\sigma_{n't} = (\ell', B, B)$;
- $\pi_{jt+1}(\ell^{**}, B, A) > \pi_{jt+1}(\ell^{**}, B, B)$ for every $j \in \mathcal{F}$, $j \neq n'$ if $\sigma_{n't} \in \{(\ell', A, B), (\ell', B, B)\}$;
- $\pi_{jt+1}(\ell^{**}, B, A) > \pi_{jt+1}(\ell', B, B)$ for every $j \in \mathcal{F}$, $j \neq n'$ if $\sigma_{n't} = (\ell', B, B)$.

Moreover, if $\sigma_{n't} = (\ell', A, A)$ or $\sigma_{n't} = (\ell', A, B)$, then $\pi_{it+1}(\ell^*, A, A) = \pi_{it+1}(\ell', A, A)$ and, so, every coarse reasoner $i \neq n'$ is indifferent between keeping strategy (ℓ^*, A, A) and adopting strategy (ℓ', A, A). But then with positive probability at then end of the next period of time every agent $i \in \mathcal{C}$, $i \neq n'$ will be given a revision opportunity and will adopt strategy (ℓ', A, A). Consequently, the system will reach a different absorbing state of the type A/BA which does not belong to the basin of attraction of S'.

If either $\sigma_{n't} = (\ell', B, A)$, or $\sigma_{n't} = (\ell', B, B)$, then $\pi_{jt+1}(\ell^{**}, B, A) = \pi_{jt+1}(\ell', B, A)$ and, so, every fine reasoner $j \neq n'$ is indifferent between keeping strategy (ℓ^{**}, B, A) and adopting strategy (ℓ', B, A). But then with positive probability at then end of the next period of time every fine reasoner $j \in \mathcal{F}$, $j \neq n'$ will be given a revision opportunity and will adopt strategy (ℓ', B, A). Consequently, the system will reach a different absorbing state of the type A/BA which does not belong to the basin of attraction of S'.

Hence, if $S' = \{$A-AA, A-AB, A-BA, B-AB$\}$, the population is large enough, and states of the type A/BA are absorbing, then $M(\text{A/BA}, S') > 1$.

$\square$

Lemma B.2. *Let $S' = \{A\text{-}AA, A\text{-}AB, A\text{-}BA, B\text{-}AB\}$. If the population is large enough and states of the type B-BA are absorbing, then one mistake is never enough to lead with positive probability the system into the basin of attraction of S' starting from a state of the type B-BA, i.e., $M(B\text{-}BA, S') > 1$.*

Proof. Let $S' = \{$A-AA, A-AB, A-BA, B-AB$\}$ and consider an absorbing state of the type B-BA where all agents stay in the same location. Note that, if the population is large enough, then a single mistake by any agent will have a negligible effect on the expected payoff of other agents in case they keep their current strategy and, moreover, the other agents can possibly increase their expected payoff only if the mistaken agent has moved into an empty location and coordinating locally with it provides a better payoff than their current local interactions.

Assume that the system is in an absorbing state of the type B-BA[2]. Then, every $i \in \mathcal{C}$ adopts strategy (ℓ^*, B, B) while every $j \in \mathcal{F}$ adopts strategy (ℓ^*, B, A).

Table B.2 reports the expected payoff that an agent $n \in \mathcal{N}$ would obtain by adopting strategy σ_{nt+1} at time $t + 1$ given that at time t the system was in an absorbing state of the type B-BA and that at the end of time t an agent $n' \in \mathcal{N}$, $n' \neq n$ has adopted by mistake strategy $\sigma_{n't}$. The table reports such expected payoff for each possible combination of $(\sigma_{nt+1}, \sigma_{n't})$ under the assumption that the population is large enough.

As the table shows, in most cases the expected payoff provided by the current strategy adopted by agent n, $\sigma_i = (\ell^*, B, B)$ if $i \in \mathcal{C}$ and $\sigma_j = (\ell^*, B, A)$ if $j \in \mathcal{F}$, is strictly higher than the one associated to alternative strategies after agent $n' \neq n$ has adopted by mistake strategy $\sigma_{n't}$. However, some cases are worth discussing.

---

[2] By Theorem 1 a state of the type B-BA is an absorbing state if $q \in \big(1 - \alpha, \min\big\{\frac{1-\alpha}{1-p}, 1\big\}\big)$.
If $q \in (0, 1 - \alpha)$, then for every fine reasoner $j \in \mathcal{F}$ $\pi_{jt}(\ell^*, B, A) < \pi_{jt}(\ell^*, B, B)$. But then, with positive probability at the end of the next period of time every $j \in \mathcal{F}$ will be given a revision opportunity and will adopt strategy (ℓ^*, B, B). Consequently, without any mistake the system will reach an absorbing state of the type B-BB. Therefore, if $q \in (0, 1 - \alpha)$, the analysis in Lemma B.3 is the one of interest.
If, instead, $q \in \big(\min\big\{\frac{1-\alpha}{1-p}, 1\big\}, 1\big)$, then for every coarse reasoner $i \in \mathcal{C}$ $\pi_{it}(\ell^*, B, B) < \pi_{it}(\ell^*, A, A)$. But then, with positive probability at the end of the next period of time every $i \in \mathcal{C}$ will be given a revision opportunity and will adopt strategy (ℓ^*, A, A). Consequently, without any mistake the system will reach a state of type A-BA which belongs to the basin of attraction of S'. In other words, if $q \in \big(\min\big\{\frac{1-\alpha}{1-p}, 1\big\}, 1\big)$ then with positive probability – without any mistake – the system can reach the basin of attraction of S' starting from a state of type B-BA.

Case $\sigma_{n't} = (\ell', A, y)$

σ_{nt+1}	$\sigma_{n't}$	
	(ℓ', A, A)	(ℓ', A, B)
(ℓ^*, A, A)	$pc + (1-p)[qa + (1-q)c]^{(+)}$	$pc + (1-p)[qa + (1-q)c]^{-}$
(ℓ', A, A)	$pa + (1-p)[qa + (1-q)c]^{(+)}$	$pa + (1-p)[qa + (1-q)c]^{-}$
(ℓ^*, A, B)	$pc + (1-p)[qd + (1-q)b]^{(-)}$	$pc + (1-p)[qd + (1-q)b]^{+}$
(ℓ', A, B)	$pa + (1-p)[qd + (1-q)b]^{(-)}$	$pa + (1-p)[qd + (1-q)b]^{+}$
$\boldsymbol{(\ell^*, B, A)}$	$\boldsymbol{pb + (1-p)[qa + (1-q)c]^{(+)}}$	$\boldsymbol{pb + (1-p)[qa + (1-q)c]^{-}}$
(ℓ', B, A)	$pd + (1-p)[qa + (1-q)c]^{(+)}$	$pd + (1-p)[qa + (1-q)c]^{-}$
$\boldsymbol{(\ell^*, B, B)}$	$\boldsymbol{pb + (1-p)[qd + (1-q)b]^{(-)}}$	$\boldsymbol{pb + (1-p)[qd + (1-q)b]^{+}}$
(ℓ', B, B)	$pd + (1-p)[qd + (1-q)b]^{(-)}$	$pd + (1-p)[qd + (1-q)b]^{+}$

Case $\sigma_{n't} = (\ell', B, y)$ with $\ell' \neq \ell^*$

σ_{nt+1}	$\sigma_{n't}$	
	(ℓ', B, A)	(ℓ', B, B)
(ℓ^*, A, A)	$pc + (1-p)[qa + (1-q)c]$	$pc + (1-p)[qa + (1-q)c]^{(-)}$
(ℓ', A, A)	$pc + (1-p)[qa + (1-q)c]$	$pc + (1-p)[qa + (1-q)c]^{(-)}$
(ℓ^*, A, B)	$pc + (1-p)[qd + (1-q)b]$	$pc + (1-p)[qd + (1-q)b]^{(+)}$
(ℓ', A, B)	$pc + (1-p)[qd + (1-q)b]$	$pc + (1-p)[qd + (1-q)b]^{(+)}$
$\boldsymbol{(\ell^*, B, A)}$	$\boldsymbol{pb + (1-p)[qa + (1-q)c]}$	$\boldsymbol{pb + (1-p)[qa + (1-q)c]^{(-)}}$
(ℓ', B, A)	$pb + (1-p)[qa + (1-q)c]$	$pb + (1-p)[qa + (1-q)c]^{(-)}$
$\boldsymbol{(\ell^*, B, B)}$	$\boldsymbol{pb + (1-p)[qd + (1-q)b]}$	$\boldsymbol{pb + (1-p)[qd + (1-q)b]^{(+)}}$
(ℓ', B, B)	$pb + (1-p)[qd + (1-q)b]$	$pb + (1-p)[qd + (1-q)b]^{(+)}$

Table B.2: Effects of one mistake in an absorbing state of the type B-BA. Expected payoff of agent n at time $t+1$ if agent n decides to adopt strategy σ_{nt+1} given that at time t the system was in an absorbing state of the type B-BA and that at the end of time t agent $n' \neq n$ has adopted by mistake strategy $\sigma_{n't} = (\ell', x, y)$ with $\ell' \neq \ell^*, \ell^{**}$. Strategies available to coarse reasoners in gray, current strategies of coarse and fine reasoners in bold. $+$ [$-$] indicates that the expected payoff is approximated from above [below] under the assumption of a large enough population.

The population must be large enough in order to guarantee the following inequalities:

- $\pi_{it+1}(\ell^*, B, B) > \pi_{it+1}(\ell^*, A, A)$ for every $i \in \mathcal{C}$, $i \neq n'$ if $\sigma_{n't} =$

(ℓ', A, A);

- $\pi_{jt+1}(\ell^*, B, A) > \pi_{jt+1}(\ell^*, B, B)$ for every $j \in \mathcal{F}$, $j \neq n'$ if $\sigma_{n't} \in \{(\ell', A, B), (\ell', B, B)\}$;
- $\pi_{jt+1}(\ell^*, B, A) > \pi_{jt+1}(\ell', B, B)$ for every $j \in \mathcal{F}$, $j \neq n'$ if $\sigma_{n't} = (\ell', B, B)$.

Moreover, if $\sigma_{n't} = (\ell', B, A)$ or $\sigma_{n't} = (\ell', B, B)$, then $\pi_{it+1}(\ell^*, B, B) = \pi_{it+1}(\ell', B, B)$ for every coarse reasoner $i \neq n'$ and $\pi_{jt+1}(\ell^*, B, A) = \pi_{jt+1}(\ell', B, A)$ for every fine reasoner $j \neq n'$. Then, every coarse reasoner $i \neq n'$ is indifferent between keeping strategy (ℓ^*, B, B) and adopting strategy (ℓ', B, B) while every fine reasoner $j \neq n'$ is indifferent between keeping strategy (ℓ^*, B, A) and adopting strategy (ℓ', B, A). But then with positive probability at then end of the next period of time every agent $n \in \mathcal{N}$, $n \neq n'$ will be given a revision opportunity and will adopt strategy (ℓ', B, B) if $n \in \mathcal{C}$ or strategy (ℓ', B, A) if $n \in \mathcal{F}$. Consequently, the system will reach a different absorbing state of the type B-BA which does not belong to the basin of attraction of S'.

Finally, if $\sigma_{n't} \in \{(\ell', A, A), (\ell', A, B)\}$, then for every coarse reasoner $i \in \mathcal{C}$

$$\begin{gathered}\pi_{it+1}(\ell^*, B, B) > \pi_{it+1}(\ell', A, A) \Leftrightarrow \\ \Leftrightarrow pb + (1-p)[qd + (1-q)b] > pa + (1-p)[qa + (1-q)c] \Leftrightarrow \\ \Leftrightarrow (b-c) - p(a-c) > (1-p)q(a-d+b-c) \Leftrightarrow \\ \Leftrightarrow q < \frac{1-\alpha}{1-p} - \frac{p}{1-p}\frac{a-c}{a-d+b-c}\end{gathered}$$

But then, if $q \in \left(\min\{\frac{1-\alpha}{1-p} - \frac{p}{1-p}\frac{a-c}{a-d+b-c}, 1\}, \min\{\frac{1-\alpha}{1-p}, 1\}\right)$, coarse reasoners prefer strategy (ℓ', A, A) to strategy (ℓ^*, B, B). Therefore, with positive probability at the end of the next period of time every agent $i \in \mathcal{C}$ will be given a revision opportunity and will adopt strategy (ℓ', A, A). Consequently, the system will reach an absorbing state of the type A/BA, which does not belong to the basin of attraction of S'.

Hence, if $S' = \{$A-AA, A-AB, A-BA, B-AB$\}$, the population is large enough, and states of the type B-BA are absorbing, then $M(\text{B-BA}, S') > 1$. □

Lemma B.3. *Let $S' = \{$A-AA, A-AB, A-BA, B-AB$\}$. If the population is large enough, then one mistake is never enough to lead with positive probability the* *system into the basin of attraction of S' starting from an absorbing state of the type B-BB, i.e., $M(\text{B-BB}, S') > 1$.*

Proof. Let $S' = \{\text{A-AA, A-AB, A-BA, B-AB}\}$ and consider an absorbing state of the type B-BB where all agents stay in the same location. Note that, if the population is large enough, then a single mistake by any agent will have a negligible effect on the expected payoff of other agents in case they keep their current strategy and, moreover, the other agents can possibly increase their expected payoff only if the mistaken agent has moved into an empty location and coordinating locally with it provides a better payoff than their current local interactions.

Assume that the system is in a state of type B-BB. Then, every $n \in \mathcal{N}$ adopts strategy (ℓ^*, B, B). Moreover, by Theorem 1 states of type B-BB are absorbing states if $p, q \in (0, 1)$.

Table B.3 reports the expected payoff that an agent $n \in \mathcal{N}$ would obtain by adopting strategy σ_{nt+1} at time $t+1$ given that at time t the system was in an absorbing state of the type B-BB and that at the end of time t an agent $n' \in \mathcal{N}$, $n' \neq n$ has adopted by mistake strategy $\sigma_{n't}$. The table reports such expected payoff for each possible combination of $(\sigma_{nt+1}, \sigma_{n't})$ under the assumption that the population is large enough.

σ_{nt+1}	$\sigma_{n't}$			
	(ℓ', A, A)	(ℓ', A, B)	(ℓ', B, A)	(ℓ', B, B)
(ℓ^*, A, A)	c^+	c	c^+	c
(ℓ', A, A)	$pa + (1-p)c^+$	$pa + (1-p)c$	c^+	c
(ℓ^*, A, B)	$pc + (1-p)b^-$	$pc + (1-p)b$	$pc + (1-p)b^-$	$pc + (1-p)b$
(ℓ', A, B)	$pa + (1-p)b^-$	$pa + (1-p)b$	$pc + (1-p)b^-$	$pc + (1-p)b$
(ℓ^*, B, A)	$pb + (1-p)c^+$	$pb + (1-p)c$	$pb + (1-p)c^+$	$pb + (1-p)c$
(ℓ', B, A)	$pd + (1-p)c^+$	$pd + (1-p)c$	$pb + (1-p)c^+$	$pb + (1-p)c$
$\boldsymbol{(\ell^*, B, B)}$	$\boldsymbol{b^-}$	$\boldsymbol{b}$	$\boldsymbol{b^-}$	$\boldsymbol{b}$
(ℓ', B, B)	$pd + (1-p)b^-$	$pd + (1-p)b$	b^-	b

Table B.3: Effects of one mistake in an absorbing state of the type B-BB. Expected payoff of agent n at time $t+1$ if agent decides to adopt strategy σ_{nt+1} given that at time t the system was in an absorbing state of the type B-BB and that at the end of time t agent $n' \neq n$ has adopted by mistake strategy $\sigma_{n't} = (\ell', x, y)$ with $\ell' \neq \ell^*, \ell^{**}$. Strategies available to coarse reasoners in gray, current strategy of coarse and fine reasoners in bold. $+$ $[-]$ indicates that the expected payoff is approximated from above [below] under the assumption of a large enough population.

As the table shows, in most cases the expected payoff provided by the current strategy adopted by agent n, $\sigma_n = (\ell^*, B, B)$, is strictly higher than the one associated to alternative strategies after agent $n' \neq n$ has adopted by mistake strategy $\sigma_{n't}$. However, some cases are worth discussing.

The population must be large enough in order to guarantee the following inequalities:

- $\pi_{nt+1}(\ell^*, B, B) > \pi_{nt+1}(\ell^*, A, A)$ for every $n \in \mathcal{N}$, $n \neq n'$ if $\sigma_{n't} \in \{(\ell', A, A), (\ell', B, A)\}$;
- $\pi_{nt+1}(\ell^*, B, B) > \pi_{nt+1}(\ell', A, A)$ for every $n \in \mathcal{N}$, $n \neq n'$ if $\sigma_{n't} \in \{(\ell', A, A), (\ell', B, A)\}$;
- $\pi_{jt+1}(\ell^*, B, B) > \pi_{jt+1}(\ell^*, B, A)$ for every $j \in \mathcal{F}$, $j \neq n'$ if $\sigma_{n't} \in \{(\ell', A, A), (\ell', B, A)\}$;
- $\pi_{jt+1}(\ell^*, B, B) > \pi_{jt+1}(\ell', B, A)$ for every $j \in \mathcal{F}$, $j \neq n'$ if $\sigma_{n't} = (\ell', B, A)$;

Moreover, if $\sigma_{n't} = (\ell', B, A)$ or $\sigma_{n't} = (\ell', B, B)$, then $\pi_{nt+1}(\ell^*, B, B) = \pi_{nt+1}(\ell', B, B)$ for every $n \in \mathcal{N}$ and, so, every agent $n \neq n'$ is indifferent between keeping strategy (ℓ^*, B, B) and adopting strategy (ℓ', B, B). But then, with positive probability at then end of the next period of time every agent $n \in \mathcal{N}$, $n \neq n'$ will be given a revision opportunity and will adopt strategy (ℓ', B, B). Consequently, the system will reach a different absorbing state of the type B-BB which does not belong to the basin of attraction of S'.

Hence, if $S' = \{$A-AA, A-AB, A-BA, B-AB$\}$ and the population is large enough, then $M(\text{B-BB}, S') > 1$. □

Lemma B.4. *Let $S' = \{$A-AA, A-AB, A-BA, B-AB$\}$. If the population is large enough, then one mistake can lead with positive probability the system outside the basin of attraction of S' starting from any absorbing state belonging to S', i.e., $M(S', S'^{-1}) = 1$.*

Proof. Let $S' = \{$A-AA, A-AB, A-BA, B-AB$\}$ and assume that the population is large enough.

We need to show that a single mistake can lead with positive probability the system outside the basin of attraction of S' starting from any absorbing state belonging to S' and, so, starting from absorbing states of the type A-AA, A-AB, A-BA, and B-AB. In the following, we prove this separately for the four types of absorbing states.

Case: From an absorbing state of the type A-AA
Assume that the system is in an absorbing state of the type A-AA. Then, every $n \in \mathcal{N}$ adopts strategy (ℓ^*, A, A). Moreover, by Theorem 1 states of the type A-AA are absorbing if $p, q \in (0, 1)$.

Assume that at the end of time t a fine reasoner $j' \in \mathcal{F}$ adopts by mistake strategy $\sigma_{j't} = (\ell', B, A)$ with $\ell' \neq \ell^*$. Then, at then end of time $t + 1$ for every $j \in \mathcal{F}$, $j \neq j'$

$$\pi_{jt+1}(\ell', B, A) = pb + (1 - p)a > a = \pi_{jt+1}(\ell^*, A, A)$$

and, so, every fine reasoner $j \neq j'$ strictly prefers strategy (ℓ', B, A) to strategy (ℓ^*, A, A). But then with positive probability at then end of the next period of time every fine reasoner $j \in \mathcal{F}$, $j \neq j'$ will be given a revision opportunity and will adopt strategy (ℓ', B, A). Consequently, if the population is large enough the system will reach a state of the type A/BA which does not belong to the basin of attraction of S' (see Lemma B.1).

Therefore, if the system is in an absorbing state of the type A-AA and the population is large enough one mistake can lead with positive probability the system outside the basin of attraction of S'.

Case: From an absorbing state of the type A-AB
Assume that the system is in an absorbing state of the type A-AB. Then, every $i \in \mathcal{C}$ adopts strategy (ℓ^*, A, A) while every $j \in \mathcal{F}$ adopts strategy (ℓ^*, A, B).

According to Theorem 1 states of type A-AB are absorbing states if condition $q \in \left(\alpha, \min\left\{\frac{\alpha}{1-p}, 1\right\}\right)$ holds. Moreover, by Theorem 1 if states of type A-AB are not absorbing the system will reach another absorbing state and, consequently, either the analysis of a different absorbing state belonging to S' is the relevant one or the system will exit without any mistake the basin of attraction of S'.

Consider the case $q \in \left(\alpha, \min\left\{\frac{\alpha}{1-p}, 1\right\}\right)$. Assume that at the end of time t a fine reasoner $j' \in \mathcal{F}$ adopts by mistake strategy $\sigma_{j't} = (\ell', B, B)$ with $\ell' \neq \ell^*$. Then, at then end of time $t + 1$ for every fine reasoner $j \neq j'$

$$\begin{aligned}\pi_{jt+1}(\ell', B, B) = pb + (1 - p)[qb + (1 - q)d] > \dots\\ \dots > pa + (1 - p)[qb + (1 - q)d] = \pi_{jt+1}(\ell^*, A, B)\end{aligned}$$

and, so, every fine reasoner $j \neq j'$ strictly prefers strategy (ℓ', B, B) to strategy (ℓ^*, A, B). But then with positive probability at then end of the next period of time every fine reasoner $j \in \mathcal{F}$, $j \neq j'$ will be given a revision opportunity and will adopt strategy (ℓ', B, B). Consequently, the system will reach a state of type A/BB which by Theorem 1 is never an absorbing state. More precisely, given $q \in \left(\alpha, \min\left\{\frac{\alpha}{1-p}, 1\right\}\right)$, in such state for every coarse reasoner $i \in \mathcal{C}$

$$\begin{aligned}\pi_{it+2}(\ell', B, B) = pb + (1-p)[qb + (1-q)d] > \dots \\ \dots > pa + (1-p)[qc + (1-q)a] = \pi_{it+2}(\ell^*, A, A)\end{aligned}$$

and, so, every coarse reasoner strictly prefers strategy (ℓ', B, B) to strategy (ℓ^*, A, A). But then with positive probability at then end of the next period of time every coarse reasoner $i \in \mathcal{C}$ will be given a revision opportunity and will adopt strategy (ℓ', B, B). But then the system will reach an absorbing state of the type B-BB which does not belong to the basin of attraction of S'.

Therefore, if the system is in an absorbing state of the type A-AB, then one mistake may lead with positive probability the system outside the basin of attraction of S'.

Case: From an absorbing state of the type A-BA

Assume that the system is in an absorbing state of the type A-BA. Then, every $i \in \mathcal{C}$ adopts strategy (ℓ^*, A, A) while every $j \in \mathcal{F}$ adopts strategy (ℓ^*, B, A).

By Theorem 1 states of the type A-BA are absorbing states if $q \in \left(\alpha, \min\left\{\frac{\alpha}{p}, 1\right\}\right)$. Moreover, by Theorem 1 if states of the type A-BA are not absorbing the system will reach another absorbing state and, consequently, either the analysis of a different absorbing state belonging to S' is the relevant one or the system will exit without any mistake the basin of attraction of S'.

Consider the case $q \in \left(\alpha, \min\left\{\frac{\alpha}{p}, 1\right\}\right)$. Assume that at the end of time t a fine reasoner $j' \in \mathcal{F}$ adopts by mistake strategy $\sigma_{j't} = (\ell', B, A)$ with $\ell' \neq \ell^*$. Then, at then end of time $t+1$ for every other fine reasoner $j \in \mathcal{F}$, $j \neq j'$

$$\pi_{jt+1}(\ell', B, A) = pb + (1-p)a > p[qb + (1-q)a] + (1-p)a \approx \pi_{jt+1}(\ell^*, B, A)$$

and, so, every fine reasoner $j \neq j'$ strictly prefers strategy (ℓ', B, A) to strategy (ℓ^*, B, A). But then with positive probability at then end of the next period of time every agent $j \in \mathcal{F}$, $j \neq j'$ will be given a revision opportunity and will adopt strategy (ℓ', B, A). Consequently, the system will reach a state of the type A/BA which does not belong to the basin of attraction of S' (see Lemma B.1).

Therefore, if the system is in an absorbing state of the type A-BA and the population is large enough one mistake may lead with positive probability the system outside the basin of attraction of S'.

Case: From an absorbing state of the type B-AB
Assume that the system is in an absorbing state of the type B-AB. Then, every $i \in \mathcal{C}$ adopts strategy (ℓ^*, B, B) while every $j \in \mathcal{F}$ adopts strategy (ℓ^*, A, B).

By Theorem 1 states of the type B-AB are absorbing states if $q \in \left(1 - \alpha, \min\left\{\frac{1-\alpha}{p}, 1\right\}\right)$. Moreover, by Theorem 1 if states of the type B-AB are not absorbing the system will reach another absorbing state and, consequently, either the analysis of a different absorbing state belonging to S' is the relevant one or the system will exit without any mistake the basin of attraction of S'.

Consider the case $q \in \left(1-\alpha, \min\left\{\frac{1-\alpha}{p}, 1\right\}\right)$. Assume that at the end of time t a fine reasoner $j' \in \mathcal{F}$ adopts by mistake strategy $\sigma_{j't} = (\ell', B, B)$ with $\ell' \neq \ell^*$. Then, at then end of time $t + 1$ for every other fine reasoner $j \in \mathcal{F}, j \neq j'$

$$\pi_{jt+1}(\ell', B, B) = b > p[qa + (1 - q)d] + (1 - p)b \approx \pi_{jt+1}(\ell^*, A, B)$$

and, so, every fine reasoner $j \neq j'$ strictly prefers strategy (ℓ', B, B) to strategy (ℓ^*, A, B). Moreover, for every coarse reasoner $i \in \mathcal{C}$

$$\pi_{it+1}(\ell', B, B) = b > p[qd + (1 - q)b] + (1 - p)b \approx \pi_{it+1}(\ell^*, B, B)$$

and, so, every coarse reasoner strictly prefers strategy (ℓ', B, B) to strategy (ℓ^*, B, B). But then with positive probability at the end of the next period of time every agent $n \in \mathcal{N}, n \neq n'$ will be given a revision opportunity and will adopt strategy (ℓ', B, B). Consequently, the system will reach an absorbing state of the type B-BB which does not belong to the basin of attraction of S'.

Therefore, if the system is in an absorbing state of the type B-AB one mistake can lead with positive probability the system outside the basin of attraction of S'.

But then, we can conclude that if $S' = \{$A-AA, A-AB, A-BA, B-AB$\}$ and the population is large enough, then $M(S', S'^{-1}) = 1$. □

Theorem 2. *If the population is large enough, then all the absorbing states of the types A-AA, A-AB, A-BA, and B-AB are never stochastically stable.*

Proof. We consider separately the cases $q \in \left(0, \min\left\{\frac{1-\alpha}{1-p}, 1\right\}\right)$ and $q \in \left(\min\left\{\frac{1-\alpha}{1-p}, 1\right\}, 1\right)$ and for each case we show that the statement holds.

Consider the case $q \in \left(0, \min\left\{\frac{1-\alpha}{1-p}, 1\right\}\right)$. In this region of the parameter space all the absorbing states in Theorem 1 may be absorbing states. Therefore, consider the following partition of the absorbing states of the system:

- $\Omega = \{$A/BA, B-BA, B-BB$\}$;
- $\Omega^{-1} = \{$A-AA, A-AB, A-BA, B-AB$\}$

Given $q \in \left(0, \min\left\{\frac{1-\alpha}{1-p}, 1\right\}\right)$, then either states of the type B-BA are absorbing or they belong to the basin of attraction of the set of states of the type B-BB. In addition, either states of the type A/BA are absorbing or they belong to the basin of attraction of the set of states of the types B-BA and B-BB. Therefore, by combining Lemma B.1-Lemma B.3 we can conclude that one mistake is never enough to lead the system with positive probability outside the basin of attraction of Ω, i.e., $M(\Omega, \Omega^{-1}) > 1$ and, consequently, the radius of the basin of attraction of Ω is strictly larger than one, i.e., $R(\Omega) > 1$.
Moreover, by Lemma B.4 one mistake is enough to lead the system outside the basin of attraction of Ω^{-1} starting from any absorbing state belonging to Ω^{-1}. But given that Ω and Ω^{-1} form a partition of the set of absorbing states of the system this means that one mistake is enough to make the system enter the basin of attraction of Ω starting from any absorbing state belonging to Ω^{-1} and, so, the coradius of Ω is equal to one, i.e., $CR(\Omega) = 1$.
By combining these results we can conclude that if $q \in \left(0, \min\left\{\frac{1-\alpha}{1-p}, 1\right\}\right)$ and the population is large enough then

$$R(\Omega) > 1 = CR(\Omega)$$

and, so, all the stochastically stable states of the system are contained in Ω and, consequently, all the absorbing states of the types A-AA, A-AB, A-BA, and B-AB are never stochastically stable.

Consider now the case $q \in \left(\min\left\{\frac{1-\alpha}{1-p}, 1\right\}, 1\right)$. In this region of the parameter space all the absorbing states in Theorem 1 except states of the type B-BA may be absorbing states. Therefore, consider the following partition of the absorbing states of the system:

- $\Omega = \{$A/BA, B-BB$\}$;
- $\Omega^{-1} = \{$A-AA, A-AB, A-BA, B-AB$\}$

Given $q \in \left(\min\left\{ \frac{1-\alpha}{1-p}, 1 \right\}, 1 \right)$, states of the type A/BA are always absorbing. Therefore, by combining Lemma B.1 and Lemma B.3 we can conclude that one mistake is never enough to lead the system with positive probability outside the basin of attraction of Ω, i.e., $M(\Omega, \Omega^{-1}) > 1$ and, consequently, the radius of the basin of attraction of Ω is strictly larger than one, i.e., $R(\Omega) > 1$.
Moreover, by Lemma B.4 one mistake is enough to lead the system outside the basin of attraction of Ω^{-1} starting from any absorbing state belonging to Ω^{-1}. But given that Ω and Ω^{-1} form a partition of the set of absorbing states of the system – in the (p, q)-region considered – this means that one mistake is enough to make the system enter the basin of attraction of Ω starting from any absorbing state belonging to Ω^{-1} and, so, the coradius of Ω is equal to one, i.e., $CR(\Omega) = 1$.
By combining these results we can conclude that if $q \in \left(\min\left\{ \frac{1-\alpha}{1-p}, 1 \right\}, 1 \right)$ and the population is large enough then

$$R(\Omega) > 1 = CR(\Omega)$$

and, so, all the stochastically stable states of the system are contained in Ω and, consequently, all the absorbing states of the types A-AA, A-AB, A-BA, and B-AB are never stochastically stable.

Given that if the population is large enough all the absorbing states of the types A-AA, A-AB, A-BA, and B-AB are neither stochastically stable if $q \in \left(0, \min\left\{ \frac{1-\alpha}{1-p} \right\}, 1 \right)$ nor if $q \in \left(\min\left\{ \frac{1-\alpha}{1-p}, 1 \right\}, 1 \right)$, then they are never stochastically stable. $\square$

Appendix C

Chapter 2: Proof of Theorem 3

As in Appendix B, we will denote with $M(S, S')$ the minimum number of mistakes required to lead with positive probability the system into the basin of attraction of S' starting from an absorbing state belonging to S – where both S and S' are sets of absorbing states. Moreover, we will refer to the set of states of a given type, say X-YZ, by simply writing X-YZ.

We begin by stating some preliminary results that are needed to prove Theorem 3.

Lemma C.1. *Let* $S = \{A\text{-}AB, A\text{-}BA, B\text{-}AB, B\text{-}BA, B\text{-}BB\}$ *and let* $S' \subseteq S$, *then* $(1-\alpha)N$ *mistakes can lead with positive probability the system into an absorbing state of the type A-AA starting from any absorbing state in* S', *i.e.,* $M(S', A\text{-}AA) \leq (1-\alpha)N$.

Proof. Let $S = \{\text{A-AB, A-BA, B-AB, B-BA, B-BB}\}$ and let $S' \subseteq S$. By assumption in each absorbing state belonging to S' all agents stay in the same location.

If $(1-\alpha)N$ agents adopt by mistake strategy (ℓ^*, A, A), then both Equation (A.1) and Equation (A.2) will not hold for every fine reasoner $j \in \mathcal{F}$; in addition, Equation (A.3) will not hold for every coarse reasoner $i \in \mathcal{C}$. In other words, all agents will be willing to adopt strategy (ℓ^*, A, A). But then, with positive probability at the end of the next period of time all agents except the ones who made a mistake will be given a revision opportunity and will adopt strategy (ℓ^*, A, A). Consequently, the system will reach an absorbing state of the type A-AA.

Therefore, if $S = \{\text{A-AB}, \text{A-BA}, \text{B-AB}, \text{B-BA}, \text{B-BB}\}$ and $S' \subseteq S$, then $M(S', \text{A-AA}) \leq (1 - \alpha)N$. □

Lemma C.2. *Let $S = \{A\text{-}AB, A\text{-}BA, B\text{-}AB, B\text{-}BA, B\text{-}BB\}$ and let $S' \subseteq S$, then the minimum number of mistakes required to lead with positive probability the system into the basin of attraction of S' starting from an absorbing state of the type A-AA is*

$$M(A\text{-}AA, S') = \min\left\{\alpha N, M(A/BA, S') + 1\right\}$$

Proof. Let $S = \{\text{A-AB}, \text{A-BA}, \text{B-AB}, \text{B-BA}, \text{B-BB}\}$ and let $S' \subseteq S$. Moreover, assume that the system is in an absorbing state of the type A-AA. Then, every agent $n \in \mathcal{N}$ adopts strategy (ℓ^*, A, A).

The minimum number of mistakes required to make the system enter directly – without mistakes involving agents moving into empty locations – the basin of attraction of S' is αN. Moreover, by Lemma B.4 if the system is in an absorbing state of the type A-AA a single mistake can lead with positive probability the system into a state of the type A/BA. From there $M(\text{A/BA}, S')$ mistakes can lead with positive probability the system into the basin of attraction of S'. Consequently, this path requires $M(\text{A/BA}, S') + 1$ mistakes.

Hence, if $S = \{\text{A-AB}, \text{A-BA}, \text{B-AB}, \text{B-BA}, \text{B-BB}\}$ and $S' \subseteq S$, then

$$M(\text{A-AA}, S') = \min\left\{\alpha N, M(\text{A/BA}, S') + 1\right\}$$

□

Note that if states of the type A/BA are not absorbing states then they belong to the basin of attraction of the set of states of the type B-BA if $p \in \left(\frac{a-d}{b-d}, 1\right)$ and $q \in \left(1 - \alpha, 1\right)$. But then, if $p \in \left(\frac{a-d}{b-d}, 1\right)$ and $q \in \left(1 - \alpha, 1\right)$ we can restate Lemma C.2 as

$$M(\text{A-AA}, S') = \min\left\{\alpha N, M(\text{B-BA}, S') + 1\right\}$$

Lemma C.3. *Let $S = \{A\text{-}AB, A\text{-}BA, B\text{-}AB, B\text{-}BA, B\text{-}BB\}$ and let $S' = S \setminus A\text{-}BA$. If the population is large enough, then the minimum number of mistakes required to lead with positive probability the system into the basin of attraction of S starting from an absorbing state of the type A/BA is either larger than*

$(1-\alpha)N$ *or such that:*

$$M(A/BA, S) = \begin{cases} (1-q)N - 1 & \textit{if } q \in \left(\min\left\{\frac{1-\alpha}{1-p}, 1\right\}, 1\right) \\ \left(\frac{\alpha - p\frac{b-d}{a-d+b-c}}{(1-p)+\frac{p}{1-q}\frac{a-c}{a-d+b-c}}\right)N & \textit{otherwise} \end{cases}$$

Moreover, the minimum number of mistakes required to lead with positive probability the system into the basin of attraction of S' starting from an absorbing state of the type A/BA is either larger than $(1-\alpha)N$ or such that:

$$M(A/BA, S') = \left(\frac{\alpha - p\frac{b-d}{a-d+b-c}}{(1-p)+\frac{p}{1-q}\frac{a-c}{a-d+b-c}}\right)N$$

Proof. Let $S = \{\text{A-AB}, \text{A-BA}, \text{B-AB}, \text{B-BA}, \text{B-BB}\}$ and let $S' = S \setminus \text{A-BA}$. Assume that the system is in an absorbing state of the type A/BA in which every coarse reasoner $i \in \mathcal{C}$ adopts strategy (ℓ^*, A, A) while every fine reasoner $j \in \mathcal{F}$ adopts strategy (ℓ^{**}, B, A). By Theorem 1 it must be $p \in \left(0, \frac{a-d}{b-d}\right)$.

We begin by computing separately for each absorbing state belonging to the set S the minimum number of mistakes required to lead the system into it starting from an absorbing state of the type A/BA.

Case: To reach an absorbing state of the type A-AB

Assume that also states of the type A-AB are absorbing. Then, by Theorem 1 it must be $q \in \left(\alpha, \min\left\{\frac{\alpha}{1-p}, 1\right\}\right)$.

Given that in an absorbing state of the type A/BA all agents play action A in global interactions, $\gamma_j = 0$ for every fine reasoner $j \in \mathcal{F}$. Moreover, given that in the arrival state coarse reasoners still play action A in both local and global interactions, at least αN mistakes will be required for Equation (A.1) to be satisfied for every fine reasoner and, so, to reach an absorbing state of the type A-AB.

Therefore, if both states of the type A/BA and states of the type A-AB are absorbing, then

$$M(\text{A/BA}, \text{A-AB}) \geq \alpha N$$

Case: To reach an absorbing state of the type A-BA

Assume that also states of the type A-BA are absorbing. Then, by Theorem 1 it must be $q \in \left(\alpha, \min\left\{\frac{\alpha}{p}, 1\right\}\right)$.

Given that states of the type A/BA differ from states of the type A-BA only in having agents segregated according to type, the system can reach a state of the type A-BA in two possible ways: (i) by mistake $(1-q)N-1$ coarse reasoners adopt strategy (ℓ^{**}, A, A) or (ii) by mistake $qN-1$ fine reasoners adopt strategy (ℓ^{*}, B, A). However, given that states of type A-BA are absorbing if $q \in \big(\alpha, \min\big\{\frac{\alpha}{p}, 1\big\}\big)$, then the first path is the one requiring the minimum amount of mistakes.

Hence, if both states of type A/BA and states of type A-BA are absorbing states, then

$$M(\text{A/BA}, \text{A-BA}) = (1-q)N - 1$$

Case: To reach an absorbing state of the type B-AB

Assume that also states of the type B-AB are absorbing. Then, by Theorem 1 it must be $q \in \big(1-\alpha, \min\big\{\frac{1-\alpha}{p}, 1\big\}\big)$.

Assume that by mistake all coarse reasoners but one adopt strategy (ℓ^{**}, B, B) and, so, $(1-q)N-1$ mistakes are made. Then, the system will reach with positive probability a state of the type B-BA. In a state of this type all agents stay in the same location and play B in local interactions; therefore, at least additional $(1-\alpha)N$ mistakes will be needed to make it convenient for fine reasoners to adopt strategy (ℓ^{**}, A, B). But then, this path requires more than $(1-\alpha)N$ mistakes.

Assume now that by mistake all fine reasoners but one adopt strategy (ℓ^{*}, A, B) and, so, $qN-1$ mistakes are made. But then, given that $q > 1-\alpha$, this path requires more than $(1-\alpha)N$ mistakes.

Assume that by mistake fN fine reasoners adopt strategy (ℓ^{**}, A, B). Then given that in a state of the type A/BA all agents play A in global interactions at least $\alpha N > (1-\alpha)N$ mistakes will be required to make it convenient for fine reasoners to play B globally.

Assume that by mistake kN coarse reasoners adopt strategy (ℓ^{*}, B, B) and that such mistakes are enough to make it convenient for coarse reasoners to adopt strategy (ℓ^{**}, B, B). Then the system will reach with positive probability a state of the type B-BA. At this point at least $(1-\alpha)N$ additional mistakes will be needed to make it convenient for fine reasoners to adopt strategy (ℓ^{**}, A, B) as in a state of the type B-BA all agents play action B locally. But then this path requires at least $(1-\alpha)N$ mistakes.

Hence, if both states of the type A/BA and states of the type B-AB are absorbing states, then

$$M(\text{A/BA}, \text{B-AB}) > (1-\alpha)N$$

Case: To reach an absorbing state of the type B-BA

Assume that also states of the type B-BA are absorbing. Then, by Theorem 1 it must be $q \in \left(1-\alpha, \min\left\{\frac{1-\alpha}{1-p}, 1\right\}\right)$.

If by mistake all coarse reasoners but one adopt strategy (ℓ^{**}, B, B) and, so, $(1-q)N-1$ mistakes are made, then the system will reach with positive probability a state of type B-BA.

If by mistake all fine reasoners but one adopt strategy (ℓ^{*}, B, A), then $qN-1$ mistakes are made. But given that $q > 1-\alpha$ this process will require at least $(1-\alpha)N$ mistakes.

If by mistake kN coarse reasoners adopt strategy (ℓ^{*}, B, B), then it must be $\lambda^B_{i\ell^*} = \frac{k}{1-q}$, $\lambda_{i\ell^{**}} = 1$, $\gamma_i = k$, and Equation (A.3) will hold for every coarse reasoner $i \in \mathcal{C}$ if

$$p\left(\frac{b-d}{a-d+b-c} + \frac{k}{1-q}\frac{a-c}{a-d+b-c}\right) + (1-p)k > \alpha \Leftrightarrow$$
$$\Leftrightarrow p\frac{k}{1-q}\frac{a-c}{a-d+b-c} + (1-p)k > \alpha - p\frac{b-d}{a-d+b-c} \Leftrightarrow$$
$$\Leftrightarrow k > \frac{\alpha - p\frac{b-d}{a-d+b-c}}{(1-p) + \frac{p}{1-q}\frac{a-c}{a-d+b-c}}$$

If this holds, the system will reach with positive probability a state of the type B-BA. Note that given $q < \frac{1-\alpha}{1-p}$ this alternative always requires less mistakes than $(1-q)N$ (under the large population assumption the -1 can be disregarded). In fact,

$$1-q < \frac{\alpha - p\frac{b-d}{a-d+b-c}}{(1-p) + \frac{p}{1-q}\frac{a-c}{a-d+b-c}} \Leftrightarrow$$
$$\Leftrightarrow (1-p)(1-q) + p\frac{a-c}{a-d+b-c} < \alpha - p\frac{b-d}{a-d+b-c} \Leftrightarrow$$
$$\Leftrightarrow (1-p) - q(1-p) < \alpha - p \Leftrightarrow$$
$$\Leftrightarrow q > \frac{1-\alpha}{1-p}$$

If kN coarse reasoners adopt by mistake strategy (ℓ^{**}, B, B), more mistakes will be required to make the system reach with positive probability a state of the type B-BA with respect to the previous case as now $\lambda^B_{i\ell^*} = 0$, $\lambda_{i\ell^{**}} = 1$, $\gamma_i = k$. A similar reasoning applies in case fN fine reasoners adopt by mistake strategy (ℓ^{*}, B, A) as this would imply

$\lambda^B_{i\ell^*} = \frac{f}{1-q+f}$, $\lambda_{i\ell^{**}} = 1$, $\gamma_i = 0$.

Hence, if both states of type A/BA and states of type B-BA are absorbing states, then

$$M(\text{A/BA}, \text{B-BA}) = \left(\frac{\alpha - p\frac{b-d}{a-d+b-c}}{(1-p) + \frac{p}{1-q}\frac{a-c}{a-d+b-c}} \right) N$$

Case: To reach an absorbing state of the type B-BB
By Theorem 1 states of the type B-BB are absorbing in the entire parameter space and, so, no additional assumption are needed.

The minimum number of mistakes required to lead the system into an absorbing state of the type B-BB starting from an absorbing state of the type A/BA is attained different ways depending on the area of the parameter space considered.

If $q \in \left(0, \min\left\{\frac{1-\alpha}{1-p}, 1\right\}\right)$, the easiest way for the system to reach a state of the type B-BB starting from an absorbing state of the type A/BA requires passing through a state of the type B-BA. $M(\text{A/BA}, \text{B-BA})$ mistakes are enough to make the system reach with positive probability a state of the type B-BA starting from an absorbing state of the type A/BA. At this point

1. If $q \in (0, 1-\alpha)$, then states of the type B-BA are not absorbing states and fine reasoners prefer strategy (ℓ^{**}, B, B) to strategy (ℓ^{**}, B, A). Consequently, without additional mistakes the system will reach an absorbing state of the type B-BB;

2. If $q \in \left(1-\alpha, \min\left\{\frac{1-\alpha}{1-p}, 1\right\}\right)$, states of the type B-BA are absorbing states and additional $[\alpha - (1-q)]N$ mistakes – by fine reasoners adopting strategy (ℓ^{**}, B, B) – will be needed to make the system reach with positive probability an absorbing state of the type B-BB.

If, instead, $q \in \left(\min\left\{\frac{1-\alpha}{1-p}, 1\right\}, 1\right)$ passing through a state of the type B-BA is no longer the best option. In fact, in this case states of the type B-BA are not absorbing and in such states coarse reasoners prefer strategy (ℓ^*, A, A) to strategy (ℓ^*, B, B). However, if αN agents adopt by mistake strategy (ℓ^*, B, B) – this is the minimum number of mistakes required to make it convenient for fine reasoners to play B globally – then the system will reach with positive probability an absorbing state of the type B-BB.

Hence, if states of the type A/BA are absorbing, then $M(\text{A/BA}, \text{B-BB})$

is equal to

$$\begin{cases} M(\text{A/BA}, \text{B-BA}) & \text{if } q \in \big(0, 1-\alpha\big) \\ M(\text{A/BA}, \text{B-BA}) + \big[\alpha - (1-q)\big]N & \text{if } q \in \big(0, \min\big\{\frac{1-\alpha}{1-p}, 1\big\}\big) \\ \alpha N & \text{if } q \in \big(\min\big\{\frac{1-\alpha}{1-p}, 1\big\}, 1\big) \end{cases}$$

But then we can conclude that

$$M(\text{A/BA}, S) = \begin{cases} (1-q)N - 1 & \text{if } q \in \big(\min\big\{\frac{1-\alpha}{1-p}, 1\big\}, 1\big) \\ \left(\frac{\alpha - p\frac{b-d}{a-d+b-c}}{(1-p) + \frac{p}{1-q}\frac{a-c}{a-d+b-c}}\right)N & \text{otherwise} \end{cases}$$

Moreover, it must be

$$M(\text{A/BA}, S') = \left(\frac{\alpha - p\frac{b-d}{a-d+b-c}}{(1-p) + \frac{p}{1-q}\frac{a-c}{a-d+b-c}}\right)N$$

□

Lemma C.4. *Let $S = \{A\text{-}AB, B\text{-}AB, B\text{-}BB\}$. If the population is large enough, then the minimum number of mistakes required to lead with positive probability the system into the basin of attraction of S starting from an absorbing state of the type A-BA is:*

$$M(A\text{-}BA, S) = [\alpha - p(1-\alpha)]N$$

Proof. Let $S = \{\text{A-AB}, \text{B-AB}, \text{B-BB}\}$. Assume that the system is in an absorbing state of the type A-BA in which every coarse reasoner $i \in \mathcal{C}$ adopts strategy (ℓ^*, A, A), while every fine reasoner $j \in \mathcal{F}$ adopts strategy (ℓ^*, B, A). By Theorem 1 it must be $p \in (0, 1)$ and $q \in \big(\alpha, \min\big\{\frac{\alpha}{p}, 1\big\}\big)$.

We begin by computing separately for each absorbing state belonging to S the minimum number of mistakes required to lead the system into it starting from an absorbing state of the type A-BA.

Case: To reach an absorbing state of the type A-AB
Assume that also states of the type A-AB are absorbing. Then, by Theorem 1 it must be $q \in \big(\alpha, \min\big\{\frac{\alpha}{1-p}, 1\big\}\big)$.

Given that in an absorbing state of the type A-BA all agents play action A in global interactions, $\gamma_j = 0$ for every fine reasoner $j \in \mathcal{F}$. Moreover, given that in the arrival state coarse reasoners still play action A in both local and global interactions, at least αN mistakes will be required for Equation (A.1) to be satisfied for every fine reasoner and, so, to reach an absorbing state of the type A-AB.

Therefore, if both states of the type A-BA and states of the type A-AB are absorbing, then

$$M(\text{A-BA}, \text{A-AB}) \geq \alpha N$$

Case: To reach an absorbing state of the type B-AB

Assume that also states of the type B-AB are absorbing. Then, by Theorem 1 it must be $q \in \left(1-\alpha, \min\left\{\frac{1-\alpha}{p}, 1\right\}\right)$.

If kN coarse reasoners adopt by mistake strategy (ℓ^*, B, B) and such amount of mistakes is enough for Equation (A.3) to hold, then with positive probability the system will reach an absorbing state of the type B-BA. But given that in a state of the type B-BA all agents play strategy B locally, at least additional $(1-\alpha)N$ mistakes will be required to reach an absorbing state of the type B-AB. Therefore, this path requires at least $(1-\alpha)N$ mistakes.

If fN fine reasoners adopt by mistake strategy (ℓ^*, A, B), then at least αN mistakes will be required for Equation (A.1) to hold.

If fN fine reasoners adopt by mistake strategy (ℓ^*, A, B) and kN coarse reasoners adopt strategy (ℓ^*, B, B), then:

$$\begin{gathered}
\mathbb{E}[\pi_i((\ell^*, A, B), \sigma_{-i})] \geq \mathbb{E}[\pi_i((\ell^*, A, A), \sigma_{-i})] \text{ for every } i \in \mathcal{C} \Leftrightarrow \\
\Leftrightarrow (f+k)b + (1-f-k)d \geq (f+k)c + (1-f-k)a \Leftrightarrow \\
\Leftrightarrow (f+k)(a-d+b-c) \geq (a-d) \Leftrightarrow \\
\Leftrightarrow f+k \geq \alpha > 1-\alpha
\end{gathered}$$

Therefore, if both states of the type A-BA and states of the type A-AB are absorbing, then

$$M(\text{A-BA}, \text{B-AB}) > (1-\alpha)N$$

Case: To reach an absorbing state of the type B-BB

By Theorem 1 states of the type B-BB are absorbing in the entire parameter space and, so, no additional assumption are needed.

Assume that by mistake kN coarse reasoners adopt strategy (ℓ^*, B, B)

and fN fine reasoners adopt strategy (ℓ^*, B, B). Then,

$$\begin{gathered}
\mathbb{E}[\pi_i((\ell^*, B, B), \sigma_{-i})] \geq \mathbb{E}[\pi_i((\ell^*, A, A), \sigma_{-i})] \text{ for every } i \in \mathcal{C} \Leftrightarrow \\
\Leftrightarrow p[(q+k)b + (1-q-k)d] + (1-p)[(k+f)b + (1-k-f)d] \geq \\
\geq p[(q+k)c + (1-q-k)a] + (1-p)[(k+f)c + (1-k-f)a] \Leftrightarrow \\
\Leftrightarrow pq(a-d+b-c) + k(a-d+b-c) + (1-p)f(a-d+b-c) \geq a-d \Leftrightarrow \\
\Leftrightarrow k + (1-p)f \geq \alpha - pq
\end{gathered}$$

This amount of mistakes is enough to reach a state of type B-BA. However, additional mistakes are needed to reach a state of type B-BB and, more precisely, $M(\text{B-BA}, \text{B-BB}) = [q-(1-\alpha)]N$ additional mistakes are needed.

But then, the minimum number of mistakes required to lead with positive probability the system into an absorbing state of the type B-BB starting from an absorbing state of the type A-BA is such that:

$$\begin{gathered}
\begin{cases} k = \alpha - pq - (1-p)f \\ f = q - (1-\alpha) \end{cases} \Rightarrow \\
\Rightarrow k + f = \alpha - pq - (1-p)[q-(1-\alpha)] + q - (1-\alpha) \Leftrightarrow \\
\Leftrightarrow k + f = \alpha - pq + p[q-(1-\alpha)] \Leftrightarrow \\
\Leftrightarrow k + f = \alpha - p(1-\alpha)
\end{gathered}$$

Hence, if states of the type A-BA are absorbing, then $M(\text{A-BA}, \text{B-BB}) = [\alpha - p(1-\alpha)]N$.

But then, we can conclude that

$$M(\text{A-BA}, S) = [\alpha - p(1-\alpha)]N$$

□

Lemma C.5. *Let $S = \{A\text{-}AB, B\text{-}AB, B\text{-}BB\}$. If the population is large enough, then the minimum number of mistakes required to lead with positive probability the system into the basin of attraction of S starting from an absorbing state of the type B-BA is:*

$$M(B\text{-}BA, S) = [q-(1-\alpha)]N$$

Proof. Let $S = \{\text{A-AB}, \text{B-AB}, \text{B-BB}\}$. Assume that the system is in an absorbing state of the type B-BA in which every coarse reasoner $i \in$

$\mathcal{C}$ adopts strategy (ℓ^*, B, B) while every fine reasoner $j \in \mathcal{F}$ adopts strategy (ℓ^*, B, A). By Theorem 1 it must be $p \in (0, 1)$ and $q \in \big(1 - \alpha, \min\big\{\frac{1-\alpha}{1-p}, 1\big\}\big)$.

We begin by computing separately for each absorbing state belonging to S the minimum number of mistakes required to lead the system into it starting from an absorbing state of the type B-BA.

Case: To reach an absorbing state of the type A-AB
Assume that also states of the type A-AB are absorbing. Then, by Theorem 1 it must be $q \in \big(\alpha, \min\big\{\frac{\alpha}{1-p}, 1\big\}\big)$.

If the population is large enough, in an absorbing state of the type B-BA $\gamma_n = 1 - q$ and $\lambda_{n\ell^*} = 1$ for every agent $n \in \mathcal{N}$. But then, at least $[\alpha - (1 - q)]N = [q - (1 - \alpha)]N$ fine reasoners must adopt by mistake strategy (ℓ^*, A, B) in order for Equation (A.1) to hold. Note that eventually additional mistakes may be needed in order to reach an absorbing state of the type A-AB.

Therefore, if both states of the type B-BA and states of the type A-AB are absorbing, then $M(\text{B-BA}, \text{A-AB}) \geq [q - (1 - \alpha)]N$.

Case: To reach an absorbing state of the type B-AB
Assume that also states of the type B-AB are absorbing. Then, by Theorem 1 it must be $q \in \big(1 - \alpha, \min\big\{\frac{1-\alpha}{p}, 1\big\}\big)$.

If the population is large enough, in an absorbing state of the type B-BA $\gamma_n = 1 - q$ and $\lambda_{n\ell^*} = 1$ for every agent $n \in \mathcal{N}$. But then, at least $[\alpha - (1 - q)]N = [q - (1 - \alpha)]N$ fine reasoners must adopt by mistake strategy (ℓ^*, A, B) in order for Equation (A.1) to hold. Note that additional mistakes may be required in order for Equation (A.2) not to hold and, so, to lead with positive probability the system into an absorbing state of the type B-AB.

Therefore, if both states of the type B-BA and states of the type B-AB are absorbing, then $M(\text{B-BA}, \text{B-AB}) \geq [q - (1 - \alpha)]N$.

Case: To reach an absorbing state of the type B-BB
By Theorem 1 states of the type B-BB are absorbing in the entire parameter space and, so, no additional assumption are needed.

If the population is large enough, in an absorbing state of the type B-BA $\gamma_n = 1 - q$ and $\lambda_{n\ell^*} = 1$ for every agent $n \in \mathcal{N}$. But then, at least $[q - (1 - \alpha)]N$ fine reasoners must adopt by mistake strategy (ℓ^*, A, B) in order for Equation (A.1) to hold and, consequently, to lead with positive probability the system into an absorbing state of the type B-BB.

Therefore, if states of the type B-BA are absorbing, then it must be $M(\text{B-BA}, \text{B-BB}) = [q - (1 - \alpha)]N$.

But then, we can conclude that

$$M(\text{B-BA}, S) = [q - (1 - \alpha)]N$$

$\square$

Given these results, we can now prove Theorem 3

Theorem 3. *If the population is large enough, then all and only absorbing states of the following types are stochastically stable:*

(3.1) A/BA, if $p \in \big(0, \frac{2\alpha-1}{\alpha}\big)$ *and* $q \in (0, 1)$*;*

(3.2) B-BA, if $p \in \big(\frac{a-d}{b-d}, 1\big)$ *and* $q \in \big(2(1-\alpha), 1\big)$*;*

(3.3) B-BB, if $p \in \big(\frac{a-d}{b-d}, 1\big)$ *and* $q \in \big(0, 1-\alpha\big)$*.*

Proof. We prove the three statements separately.

We begin with (3.1). The proof of this part of the theorem is divided in two main parts that consider different regions of the parameter space: in the first we consider $p \in \big(0, \frac{a-d}{b-d}\big)$ and $q \in \big(0, \min\big\{\frac{1-\alpha}{1-p}, 1\big\}\big)$ while in the second we consider $p \in \big(0, \frac{a-d}{b-d}\big)$ and $q \in \big(\min\big\{\frac{1-\alpha}{1-p}, 1\big\}, 1\big)$. This distinction is required in order to obtain valid radius-coradius arguments.

Consider first the case $p \in \big(0, \frac{a-d}{b-d}\big)$ and $q \in \big(0, \min\big\{\frac{1-\alpha}{1-p}, 1\big\}\big)$. Within this region all the absorbing states in Theorem 1 may be absorbing states. Let $\Omega = \{\text{A/BA, A-AA}\}$ and $\Omega^{-1} = \{\text{B-BA, B-BB, A-AB, A-BA, B-AB}\}$.

Given this partition of the set of absorbing states, by Lemma C.1 $(1 - \alpha)N$ are enough to lead with positive probability the system into the basin of attraction of Ω. But then,

$$CR(\Omega) \leq (1 - \alpha)N$$

Moreover, by combining Lemma C.2 and Lemma C.3 we have

$$R(\Omega) = \frac{(1-p)(a-d) - p(b-a)}{(1-p)(a-d+b-c) + p\frac{1}{1-q}(a-c)}N$$

Therefore, if $p \in \big(0, \frac{a-d}{b-d}\big)$ and $q \in \big(0, \min\big\{\frac{1-\alpha}{1-p}, 1\big\}\big)$, then all the stochastically stable states of the system are contained in Ω if:

$$\frac{(1-p)(a-d)-p(b-a)}{(1-p)(a-d+b-c)+p\frac{1}{1-q}(a-c)}N > (1-\alpha)N \Leftrightarrow$$

$$\Leftrightarrow (a-d)-(b-c) > p\bigg[(a-d)+(b-a)-(b-c)+\frac{1-\alpha}{1-q}(a-c)\bigg] \Leftrightarrow$$

$$\Leftrightarrow \alpha-(1-\alpha) > p\bigg[\alpha-\frac{a-c}{a-d+b-c}+\frac{1-\alpha}{1-q}\frac{a-c}{a-d+b-c}\bigg] \Leftrightarrow$$

$$\Leftrightarrow 2\alpha-1 > p\bigg[\alpha-\frac{a-c}{a-d+b-c}\bigg(1-\frac{1-\alpha}{1-q}\bigg)\bigg] \Leftrightarrow$$

$$\Leftrightarrow p < \frac{2\alpha-1}{\alpha-\frac{a-c}{a-d+b-c}\Big(1-\frac{1-\alpha}{1-q}\Big)}$$

which for $q < \alpha$ is always larger than $\frac{2\alpha-1}{\alpha}$. This together with the fact that states of the type A-AA are never stochastically stable – by Theorem 2 – allows us to conclude the following: if the population is large enough, then all and only absorbing states of the type A/BA are stochastically stable if $p \in \big(0, \frac{2\alpha-1}{\alpha}\big)$ and $q \in \big(0, \min\big\{\frac{1-\alpha}{1-p}, 1\big\}\big)$.

Consider now the case $p \in \big(0, \frac{a-d}{b-d}\big)$ and $q \in \big(\min\big\{\frac{1-\alpha}{1-p}, 1\big\}, 1\big)$. Within this area of the parameter space states of the type B-BA are never absorbing states. Consider, then, the following partition of absorbing states: $\Omega' = \{$A-AA, A-BA, A/BA$\}$ and $\Omega'^{-1} = \{$A-AB, B-AB, B-BB$\}$.

Given this partition of the set of absorbing states, by Lemma C.1 $(1-\alpha)N$ are enough to lead with positive probability the system into the basin of attraction of Ω'. But then,

$$CR(\Omega') \leq (1-\alpha)N$$

By combining Lemma C.2-Lemma C.4, if $q \in \big(\min\big\{\frac{1-\alpha}{1-p}, 1\big\}, 1\big)$ and $p \in \big(0, \frac{a-d}{b-d}\big)$, then the radius of the basin of attraction of Ω'^{-1} is

$$R(\Omega') = [\alpha - p(1-\alpha)]N$$

Therefore, if $p \in \big(0, \frac{a-d}{b-d}\big)$ and $q \in \big(\min\big\{\frac{1-\alpha}{1-p}, 1\big\}, 1\big)$, then all the stochastically stable states of the system are contained in Ω if:

$$\begin{gathered}R(\Omega') > CR(\Omega') \Leftrightarrow \\ \Leftrightarrow [\alpha - p(1-\alpha)]N > (1-\alpha)N \Leftrightarrow \\ \Leftrightarrow 2\alpha - 1 > p(1-\alpha) \\ \Leftrightarrow p < \frac{2\alpha-1}{1-\alpha}\end{gathered}$$

which is always larger than $\frac{2\alpha-1}{\alpha}$. This together with the fact that states of the types A-AA and A-BA are never stochastically stable – by Theorem 2 – allows us to conclude the following: if the population is large enough, then all and only absorbing states of the type A/BA are stochastically stable if $p \in \big(0, \frac{2\alpha-1}{\alpha}\big)$ and $q \in \big(\min\big\{\frac{1-\alpha}{1-p}, 1\big\}, 1\big)$.

Given that all and only absorbing states of the type A/BA are stochastically stable both if $p \in \big(0, \frac{2\alpha-1}{\alpha}\big)$ and $q \in \big(0, \min\big\{\frac{1-\alpha}{1-p}, 1\big\}\big)$ and if $p \in \big(0, \frac{2\alpha-1}{\alpha}\big)$ and $q \in \big(\min\big\{\frac{1-\alpha}{1-p}, 1\big\}, 1\big)$, the we can conclude that they are stochastically stable if $p \in \big(0, \frac{2\alpha-1}{\alpha}\big)$ and $q \in \big(0, 1\big)$.

We now prove (3.2). Consider the case $p \in \big(\frac{a-d}{b-d}, 1\big)$ and $q \in (2(1-\alpha), 1)$; moreover, consider the following partition of absorbing states: $\Omega'' = \{$A-AA, A-BA, B-BA$\}$ and $\Omega''^{-1} = \{$B-BB, A-AB, B-AB$\}$.

Given this partition of the set of absorbing states, by Lemma C.1 $(1-\alpha)N$ are enough to lead with positive probability the system into the basin of attraction of Ω''. But then,

$$CR(\Omega'') \leq (1-\alpha)N$$

Within the area of the parameter space considered states of the type A/BA are never absorbing states and, in particular, if the system is in a state of the type A/BA then without any mistake it reaches with positive probability an absorbing state of the type B-BA. But then, we can restate Lemma C.2 as follows:

$$M(\text{A-AA}, \Omega''^{-1}) = \min\big\{(\alpha N, M(\text{B-BA}, \Omega''^{-1}))\big\}$$

This together with $CR(\Omega'') \leq (1-\alpha)N$ implies that the area such that $R(\Omega'') < CR(\Omega'')$ is not tightened by the the presence of states of the type A-AA into Ω''.

Moreover, by combining Lemma C.4 and Lemma C.5 we can conclude that the minimum number of mistakes required to lead with positive probability the system outside the basin of attraction of Ω'' is

$$R(\Omega'') = [q - (1-\alpha)]N$$

Therefore, if $p \in \left(\frac{a-d}{b-d}, 1\right)$ and $q \in (2(1-\alpha), 1)$, then the stochastically stable states of the system are contained in Ω'' if

$$R(\Omega'') > CR(\Omega'') \Leftrightarrow [q-(1-\alpha)]N > (1-\alpha)N \Leftrightarrow q > 2(1-\alpha)$$

But given that by Theorem 2 states of the types A-AA and A-BA are never stochastically stable we can conclude that if the population is large enough, then all and only absorbing states of the type B-BA are stochastically stable if $p \in \left(\frac{a-d}{b-d}, 1\right)$ and $q \in \left(2(1-\alpha), 1\right)$.

We, finally, prove (3.3). Consider the case $p \in \left(\frac{a-d}{b-d}, 1\right)$ and $q \in \left(0, 1-\alpha\right)$. By Theorem 1 if $p \in \left(\frac{a-d}{b-d}, 1\right)$ and $q \in (0, (1-\alpha))$, only states of the types A-AA and B-BB are absorbing states and, consequently, candidate stochastically stable states. However, by Theorem 2 states of the type A-AA are never stochastically stable.

Therefore, if $p \in \left(\frac{a-d}{b-d}, 1\right)$ and $q \in (0, (1-\alpha))$, all and only states of the type B-BB are stochastically stable.

Hence, if the population is large enough, then all and only absorbing states of the following types are stochastically stable:

(3.1) A/BA, if $p \in \left(0, \frac{2\alpha-1}{\alpha}\right)$ and $q \in (0, 1)$;

(3.2) B-BA, if $p \in \left(\frac{a-d}{b-d}, 1\right)$ and $q \in \left(2(1-\alpha), 1\right)$;

(3.3) B-BB, if $p \in \left(\frac{a-d}{b-d}, 1\right)$ and $q \in \left(0, 1-\alpha\right)$.

□

Appendix D

Chapter 2: More Simulation Results

We report the simulation results obtained by considering variations of the payoffs characterizing the Stag Hunt game with respect to the one reported in the main text. In particular, we consider two extreme and opposite scenarios. In the first scenario the payoffs characterizing the Stag Hunt game imply an extremely high value of α which in turn implies that the risk dominant convention tends to provide a higher expected payoff even if only few agents adopt it. Instead, in the second scenario the payoffs are such that α is close to 0.50 and, so, even though A is the risk dominant convention it provides a higher expected payoff only if almost 50% of agents adopt it.

D.1 Simulation results with high α

First, we consider a case in which the basin of attraction of the risk dominant action is large by assuming $a = 4.5$, $b = 5$, $c = 4.3$, $d = 1$. Given these payoffs, we have $\alpha = 0.83$ and $\frac{a-d}{b-d} = 0.87$. In Table D.1 for each type of absorbing states in Theorem 1 we report the area in the (p, q) parameter space in which each state is absorbing and the sufficient conditions for it being stochastically stable.

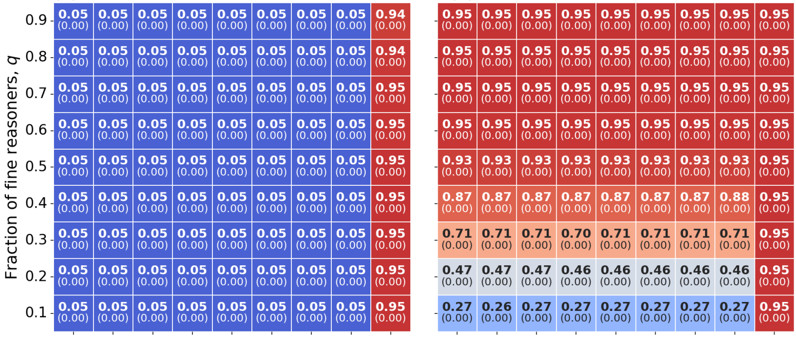
Type	Absorbing		Stochastically Stable	
	p	q	p	q
A-AA	$(0.00, 1.00)$	$(0.00, 1.00)$	$\emptyset$	$\emptyset$
A-AB	$(0.00, 1.00)$	$(0.83, \frac{0.83}{1-p})$	$\emptyset$	$\emptyset$
A-BA	$(0.00, 1.00)$	$(0.83, \frac{0.83}{p})$	$\emptyset$	$\emptyset$
A/BA	$(0.00, 0.87)$	$(0.00, 1.00)$	$(0.00, 0.79)$	$(0.00, 1.00)$
B-AB	$(0.00, 1.00)$	$(0.17, \frac{0.17}{p})$	$\emptyset$	$\emptyset$
B-BA	$(0.00, 1.00)$	$(0.17, \frac{0.17}{1-p})$	$(0.87, 1.00)$	$(0.34, 1.00)$
B-BB	$(0.00, 1.00)$	$(0.00, 1.00)$	$(0.87, 1.00)$	$(0.00, 0.17)$

Table D.1: Theoretical predictions for the simulations setup, high α. Values of p and q such that states of a given type are absorbing states and values of p and q such that they are stochastically stable. Case $a = 4.5$, $b = 5$, $c = 4.3$, $d = 1$.

Given our theoretical results, we expect the system to spend most of the time in the set of states of the type:

- A/BA if $p \in (0.00, 0.79)$ irrespective of the fraction of fine reasoners, q;
- B-BA if $p \in (0.87, 1.00)$ and $q \in (0.34, 1.00)$;
- B-BB if $p \in (0.87, 1.00)$ and $q \in (0.00, 0.17)$.

In Figure D.1 for each combination of p and q considered we report by type of agent and type of interaction the average diffusion of the the payoff dominant action. As the figure shows, the system spends most of the time in the set of states of the type A/BA for values of p lower than or equal to 0.80. Instead, for higher values of p the system spends most of the time in the set of states of the type B-BB if the fraction of fine reasoners in the population, q, is lower than or equal to 0.20; while if fine reasoning is more widespread the system spends most of the time in the set of states of the type B-BA.

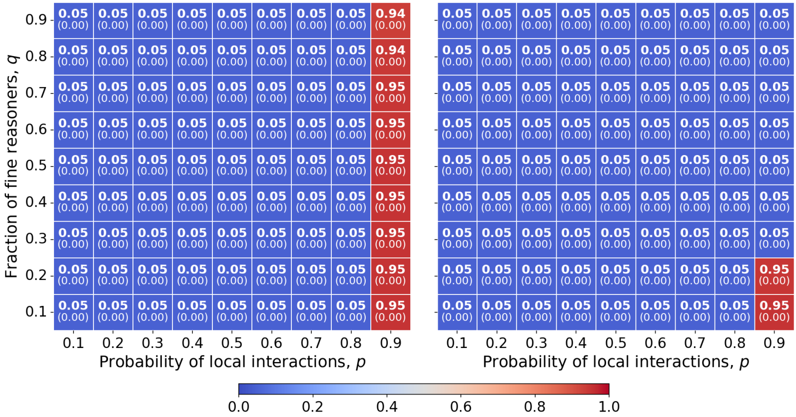


(a) Coarse Reasoners, Local Interactions

0.9	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.94 (0.00)
0.8	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.94 (0.00)
0.7	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.95 (0.00)
0.6	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.95 (0.00)
0.5	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.95 (0.00)
0.4	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.95 (0.00)
0.3	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.95 (0.00)
0.2	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.95 (0.00)
0.1	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.95 (0.00)

(b) Fine Reasoners, Local Interactions

0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)
0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)
0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)
0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)
0.93 (0.00)	0.93 (0.00)	0.93 (0.00)	0.93 (0.00)	0.93 (0.00)	0.93 (0.00)	0.93 (0.00)	0.93 (0.00)	0.95 (0.00)
0.87 (0.00)	0.87 (0.00)	0.87 (0.00)	0.87 (0.00)	0.87 (0.00)	0.87 (0.00)	0.87 (0.00)	0.88 (0.00)	0.95 (0.00)
0.71 (0.00)	0.71 (0.00)	0.71 (0.00)	0.70 (0.00)	0.71 (0.00)	0.71 (0.00)	0.71 (0.00)	0.71 (0.00)	0.95 (0.00)
0.47 (0.00)	0.47 (0.00)	0.47 (0.00)	0.46 (0.00)	0.46 (0.00)	0.46 (0.00)	0.46 (0.00)	0.46 (0.00)	0.95 (0.00)
0.27 (0.00)	0.26 (0.00)	0.27 (0.00)	0.27 (0.00)	0.27 (0.00)	0.27 (0.00)	0.27 (0.00)	0.27 (0.00)	0.95 (0.00)



(c) Coarse Reasoners, Global Interactions

0.9	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.94 (0.00)
0.8	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.94 (0.00)
0.7	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.95 (0.00)
0.6	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.95 (0.00)
0.5	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.95 (0.00)
0.4	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.95 (0.00)
0.3	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.95 (0.00)
0.2	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.95 (0.00)
0.1	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.95 (0.00)
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9

(d) Fine Reasoners, Global Interactions

0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)
0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)
0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)
0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)
0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)
0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)
0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)
0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.95 (0.00)
0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.05 (0.00)	0.95 (0.00)
0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9

Figure D.1: Simulation results, high α. Average proportion of agents playing the payoff dominant action B by type of agent (coarse vs fine reasoner) and type of interaction (local vs global interaction); standard errors in parentheses. Case $a = 4.5$, $b = 5$, $c = 4.3$, $d = 1$.

D.2 Simulation results with low α

We now consider a case in which the basin of attraction of the risk dominant action is relatively small by setting $a = 4$, $b = 5.5$, $c = 3$, $d = 1.3$ which in turn implies $\alpha = 0.52$ and $\frac{a-d}{b-d} = 0.64$. In Table D.2 for each type of absorbing state in Theorem 1 we report the area in the (p, q) parameter space in which states of such type are absorbing and the sufficient conditions for it being stochastically stable.

Type	Absorbing		Stochastically Stable	
	p	q	p	q
A-AA	$(0.00, 1.00)$	$(0.00, 1.00)$	$\emptyset$	$\emptyset$
A-AB	$(0.00, 1.00)$	$(0.52, \frac{0.52}{1-p})$	$\emptyset$	$\emptyset$
A-BA	$(0.00, 1.00)$	$(0.52, \frac{0.52}{p})$	$\emptyset$	$\emptyset$
A/BA	$(0.00, 0.64)$	$(0.00, 1.00)$	$(0.00, 0.08)$	$(0.00, 1.00)$
B-AB	$(0.00, 1.00)$	$(0.48, \frac{0.48}{p})$	$\emptyset$	$\emptyset$
B-BA	$(0.00, 1.00)$	$(0.48, \frac{0.48}{1-p})$	$(0.64, 1.00)$	$(0.96, 1.00)$
B-BB	$(0.00, 1.00)$	$(0.00, 1.00)$	$(0.64, 1.00)$	$(0.00, 0.48)$

Table D.2: Theoretical predictions for the simulations setup, low α. Values of p and q such that states of a given type are absorbing states and values of p and q such that they are stochastically stable. Case $a = 4$, $b = 5.5$, $c = 3$, $d = 1.3$.

Given our theoretical results, we expect the system to spend most of the time in the set of states of the type:

- A/BA if $p \in (0.00, 0.08)$ irrespective of the fraction of fine reasoners, q;
- B-BA if $p \in (0.64, 1.00)$ and $q \in (0.96, 1.00)$;
- B-BB if $p \in (0.64, 1.00)$ and $q \in (0.00, 0.48)$.

In Figure D.2 for each combination of p and q considered we report by type of agent and type of interaction the average diffusion of the payoff dominant action. As the figure shows, the system spends most of the time in the set of states of the type A/BA for values of p lower than or equal to 0.50. Instead, for higher values of p the system spends most of the time in the set of states of the type B-BB if the fraction of fine reasoners in the population, q, is lower than or equal to 0.50; otherwise, the system spends most of the time in the set of states of the type B-BA. We also notice that in this case the simulations provide less clear-cut predictions. We believe that this may be caused by the relatively low value of α which implies that, even though action A is risk dominant, almost half of the population must adopt action A for it to be the best reply and we believe that this weakens the diffusion of the risk dominant action.

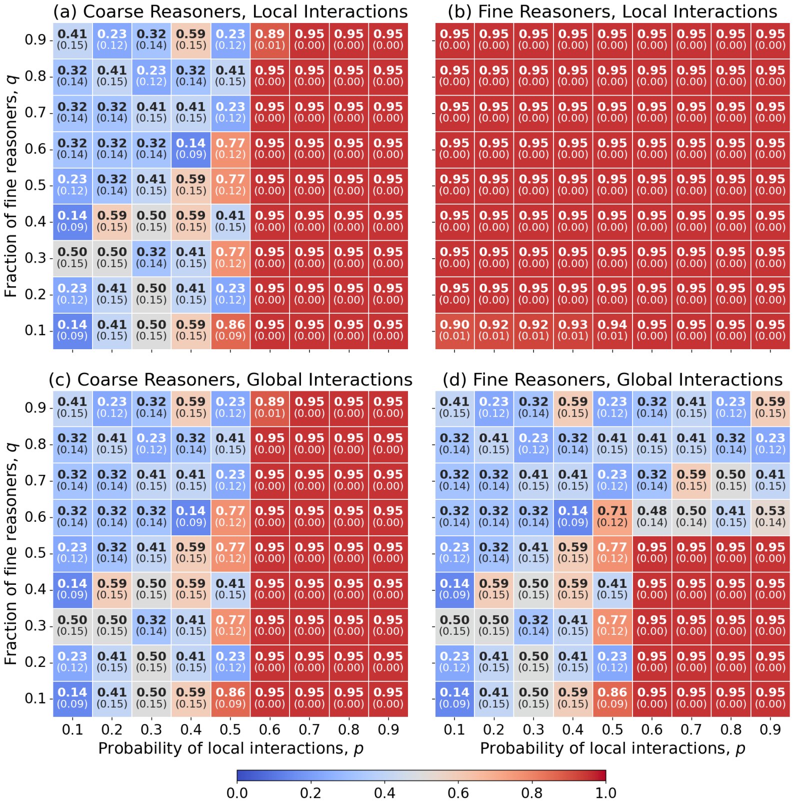


Figure D.2: Simulation results, low α. Average proportion of agents playing the payoff dominant action B by type of agent (coarse vs fine reasoner) and type of interaction (local vs global interaction); standard errors in parentheses. Case $a = 4$, $b = 5.5$, $c = 3$, $d = 1.3$.

Appendix E

Chapter 3: Proof of Theorem 4

Consider the system under unperturbed dynamics ($\varepsilon = 0$). Assume that at time t the system is in state S_t. Let $\tau^r_{a,\ell}(S_t)$ [$\tau^o_{a,\ell}(S_t)$] be agent a's probability to receive the cooperation benefit, b, from a random opponent conditional on the game played being an infinitely repeated [one-shot and anonymous] prisoners' dilemma under the assumption that agent a stays in location ℓ if this is not the case in state S_t.

Let $A_{\ell,t}$ denote the number of agents in location ℓ at time t, index a generic agent in location ℓ at time t by $\alpha = 1, \ldots, A_{\ell,t}$, and let $I_{a,\ell,t}$ be an indicator function taking value 1 if player a stays in location ℓ at time t and, so, if $\ell_{a,t} = \ell$. Then $\tau^r_{a,\ell}(S_t)$ can be expressed as: $\tau^r_{a,\ell}(S_t) =$

$$= \begin{cases} \dfrac{1}{A_{\ell,t} - I_{a,\ell,t}} \displaystyle\sum_{\substack{\alpha=1 \\ \alpha \neq a}}^{A_{\ell,t}} \left[G(k_{\alpha,t}) r_{\alpha,t} + (1 - G(k_{\alpha,t})) i_{\alpha,t} \right] & \text{if } A_{\ell,t} - I_{a,\ell,t} > 0 \\ 0 & \text{if } A_{\ell,t} - I_{a,\ell,t} = 0 \end{cases}$$

while $\tau^o_{a,\ell}(S_t) =$

$$= \begin{cases} \dfrac{1}{A_{\ell,t} - I_{a,\ell,t}} \displaystyle\sum_{\substack{\alpha=1 \\ \alpha\neq a}}^{A_{\ell,t}} \left[G(k_{\alpha,t})o_{\alpha,t} + (1-G(k_{\alpha,t}))i_{\alpha,t}\right] & \text{if } A_{\ell,t} - I_{a,\ell,t} > 0 \\ 0 & \text{if } A_{\ell,t} - I_{a,\ell,t} = 0 \end{cases}$$

Given that agents know the current state of the system, each agent a has all the information required to compute $\tau^r_{a,\ell}(S_t)$ and $\tau^o_{a,\ell}(S_t)$ for each location ℓ.

In the following we will denote with $BR_a(S_t)$ the set of best replies of agent a given the current state of the system.

Proposition E.1. *Consider the system under unperturbed dynamics ($\varepsilon = 0$). Assume that at the end of time t agent a is given a game-play revision opportunity and let $(\ell, i_{a,t}, o_{a,t}, r_{a,t}, k_{a,t})$ denote its current strategy. Then,*

$$BR_a(S_t) \subseteq \{(\ell, 0, 0, 1, k^d), (\ell, 1, 0, 1, k^c)\} \tag{E.1}$$

with $k^d = p(b-c)\tau^r_{a,\ell}(S_t)$ and $k^c = (1-p)c$.

Proof. Let $(\ell, i_{a,t}, o_{a,t}, r_{a,t}, k_{a,t})$ denote the strategy adopted by agent a at the beginning of time t. Assume that at the end of time t agent a is given a game-play revision opportunity.

Given that agent a has been given a game-play revision opportunity, it cannot change its location choice and, consequently, it will remain in location ℓ. Formally, $\ell_{a,t+1} = \ell$.

In addition, agent a will set $o_{a,t+1} = 0$ and $r_{a,t+1} = 1$ as playing AllD is a strictly dominant strategy in the one-shot and anonymous prisoners' dilemma while playing TFT is a weakly dominant strategy in the infinitely repeated prisoners' dilemma.

Given that our interest lies in finding the stochastically stable states of the system and given that equilibria in mixed strategies are never stochastically stable, we focus on the case in which agent a chooses either $i_{a,t+1} = 0$ or $i_{a,t+1} = 1$.

If agent a sets $i_{a,t+1} = 0$ (i.e., it decides to play AllD under intuition), then its expected payoff will be

$$\pi^D_{a,\ell}(S_t) = G(k^d)p(b-c)\tau^r_{a,\ell}(S_t) + (1-p)b\tau^o_{a,\ell}(S_t) - \int_0^{k^d} tg(t)\,\mathrm{d}t$$

where k^d is the optimal threshold cost of deliberation that satisfies the first order condition

$$\frac{\partial \pi^D_{a,\ell}(S_t)}{\partial k^d} = g(k^d)p(b-c)\tau^r_{a,\ell}(S_t) - g(k^d)k^d = 0 \Leftrightarrow k^d = p(b-c)\tau^r_{a,\ell}(S_t)$$

Therefore, if agent a sets $i_{a,t+1} = 0$, then it finds it optimal to choose $k_{a,t+1} = k^d$.

If, instead, agent a chooses $i_{a,t+1} = 1$ (i.e., it decides to play TFT under intuition), then its expected payoff will be

$$\pi^C_{a,\ell}(S_t) = p(b-c)\tau^r_{a,\ell}(S_t) + (1-p)b\tau^o_{a,\ell}(S_t) - (1-G(k^c))(1-p)c - \int_0^{k^c} tg(t)\,\mathrm{d}t$$

where k^c is the optimal threshold cost of deliberation that satisfies the first order condition

$$\frac{\partial \pi^C_{a,\ell}(S_t)}{\partial k^c} = g(k^c)(1-p)c - g(k^c)k^c = 0 \Leftrightarrow k^c = (1-p)c$$

Therefore, if agent a chooses $i_{a,t+1} = 1$, then it finds it optimal to set $k_{a,t+1} = k^c$.

But then, given that the location choice is fixed to the current location ℓ, agent a will myopically best reply to the current state of the system by choosing either strategy $(\ell, 0, 0, 1, k^d)$ or strategy $(\ell, 1, 0, 1, k^c)$ with $k^d = p(b-c)\tau^r_{a,\ell}(S_t)$ and $k^c = (1-p)c$.

Hence, in the system under unperturbed dynamics if at the end of time t agent a is given a game-play revision opportunity and the current location choice of agent a is $\ell_{a,t} = \ell$, then

$$BR_a(S_t) \subseteq \{(\ell, 0, 0, 1, k^d), (\ell, 1, 0, 1, k^c)\}$$

□

Corollary E.1. *Consider the system under unperturbed dynamics ($\varepsilon = 0$). Assume that at the end of time t agent a is given a full-strategy revision opportunity. Then,*

$$BR_a(S_t) \subseteq \left\{ \left((\ell, 0, 0, 1, k^d), (\ell, 1, 0, 1, k^c) \right)_{\ell=1}^{L} \right\} \tag{E.2}$$

with $k^d = p(b-c)\tau^r_{a,\ell}(S_t)$ and $k^c = (1-p)c$.

The result follows immediately from Proposition E.1. More precisely, it can be obtained by simply iterating the arguments used in Proposition E.1 for each possible location $\ell \in \mathcal{L}$.

We now introduce some definitions that are useful to prove Theorem 4.

We say that location ℓ' is at least as good as location ℓ for agent a given the state of the system at time t, formally $\ell' \succsim_{a,S_t} \ell$, if $\pi^D_{a,\ell'} \geq \pi^D_{a,\ell}$ and $\pi^C_{a,\ell'} \geq \pi^C_{a,\ell}$.

We say that location ℓ' is indifferent to location ℓ for agent a, $\ell' \sim_{a,S_t} \ell$, if $\ell' \succsim_{a,S_t} \ell$ and $\ell \succsim_{a,S_t} \ell'$. Finally, we say that location ℓ' is preferred to location ℓ for agent a, formally $\ell' \succ_{a,S_t} \ell$, if $\ell' \succsim_{a,S_t} \ell$ but not $\ell \succsim_{a,S_t} \ell'$.

Such location-ordering is informative about the desirability for agent a to move into one or another location and this turns out to be useful to analyse an agent's best reply if given a full-strategy revision opportunity. For example, if $\ell' \succsim_{a,S_t} \ell$ and $\ell' \notin BR_a(S_t)$ then it must also be that $\ell \notin BR_a(S_t)$. If, instead, $\ell' \succ_{a,S_t} \ell$, then $\ell \notin BR_a(S_t)$. More interestingly, if $\ell' \succ_{a,S_t} \ell$ for all $\ell \in \mathcal{L}$, $\ell \neq \ell'$, then we can conclude that agent a will move (or remain) into location ℓ' if given a full-strategy revision opportunity.

Proposition E.2. *The following statements hold:*

P.E.2.1 *- If* $\left(\tau^o_{a,\ell'}(S_t), \tau^r_{a,\ell'}(S_t)\right) \geq \left(\tau^o_{a,\ell}(S_t), \tau^r_{a,\ell}(S_t)\right)$*, then* $\ell' \succsim_{a,S_t} \ell$*;*

P.E.2.2 *- If* $\left(\tau^o_{a,\ell'}(S_t), \tau^r_{a,\ell'}(S_t)\right) > \left(\tau^o_{a,\ell}(S_t), \tau^r_{a,\ell}(S_t)\right)$*, then* $\ell' \succ_{a,S_t} \ell$*;*

P.E.2.3 *- If* $\left(\tau^o_{a,\ell'}(S_t), \tau^r_{a,\ell'}(S_t)\right) = \left(\tau^o_{a,\ell}(S_t), \tau^r_{a,\ell}(S_t)\right)$*, then* $\ell' \sim_{a,S_t} \ell$*;*

Proof. Consider the expected payoff obtained by agent a in case it adopts the optimal dual process defection strategy in a generic location ℓ

$$\pi^D_{a,\ell}(S_t) = G\big(p(b-c)\tau^r_{a,\ell}(S_t)\big)p(b-c)\tau^r_{a,\ell}(S_t) + \dots$$

$$\dots + (1-p)b\tau^o_{a,\ell}(S_t) - \int_0^{p(b-c)\tau^r_{a,\ell}(S_t)} tg(t)\,\mathrm{d}t$$

Such payoff depends on the location choice via $\tau^o_{a,\ell}(S_t)$ and $\tau^r_{a,\ell}(S_t)$. In particular,

$$\frac{\partial \pi^D_{a,\ell}(S_t)}{\partial \tau^r_{a,\ell}(S_t)} = g(k^d)k^d p(b-c) + G(k^d)p(b-c) - g(k^d)k^d p(b-c) > 0$$

$$\frac{\partial \pi^D_{a,\ell}(S_t)}{\partial \tau^o_{a,\ell}(S_t)} = (1-p)b > 0$$

and, so, the following implications must hold:

DD1 - If $\left(\tau^o_{a,\ell'}(S_t), \tau^r_{a,\ell'}(S_t)\right) \geq \left(\tau^o_{a,\ell}(S_t), \tau^r_{a,\ell}(S_t)\right)$, then $\pi^D_{a,\ell'}(S_t) \geq \pi^D_{a,\ell}(S_t)$;

DD2 - If $\left(\tau^o_{a,\ell'}(S_t), \tau^r_{a,\ell'}(S_t)\right) > \left(\tau^o_{a,\ell}(S_t), \tau^r_{a,\ell}(S_t)\right)$, then $\pi^D_{a,\ell'}(S_t) > \pi^D_{a,\ell}(S_t)$;

DD3 - If $\left(\tau^o_{a,\ell'}(S_t), \tau^r_{a,\ell'}(S_t)\right) = \left(\tau^o_{a,\ell}(S_t), \tau^r_{a,\ell}(S_t)\right)$, then $\pi^D_{a,\ell'}(S_t) = \pi^D_{a,\ell}(S_t)$.

Consider now the expected payoff obtained by agent a in case it adopts the optimal dual process cooperation strategy in a generic location ℓ

$$\pi^C_{a,\ell}(S_t) = p(b-c)\tau^r_{a,\ell}(S_t) + (1-p)b\tau^o_{a,\ell}(S_t) - (1-G(k^c))(1-p)c - \int_0^{k^c} tg(t)\, \mathrm{d}t$$

Such payoff depends on the location choice via $\tau^o_{a,\ell}(S_t)$ and $\tau^r_{a,\ell}(S_t)$. In particular

$$\frac{\partial \pi^C_{a,\ell}(S_t)}{\partial \tau^r_{a,\ell}(S_t)} = p(b-c) > 0,$$
$$\frac{\partial \pi^C_{a,\ell}(S_t)}{\partial \tau^o_{a,\ell}(S_t)} = (1-p)b > 0,$$

and, so, the following implications must hold:

DC1 - If $\left(\tau^o_{a,\ell'}(S_t), \tau^r_{a,\ell'}(S_t)\right) \geq \left(\tau^o_{a,\ell}(S_t), \tau^r_{a,\ell}(S_t)\right)$, then $\pi^C_{a,\ell'}(S_t) \geq \pi^C_{a,\ell}(S_t)$;

DC2 - If $\left(\tau^o_{a,\ell'}(S_t), \tau^r_{a,\ell'}(S_t)\right) > \left(\tau^o_{a,\ell}(S_t), \tau^r_{a,\ell}(S_t)\right)$, then $\pi^C_{a,\ell'}(S_t) > \pi^C_{a,\ell}(S_t)$;

DC3 - If $\left(\tau^o_{a,\ell'}(S_t), \tau^r_{a,\ell'}(S_t)\right) = \left(\tau^o_{a,\ell}(S_t), \tau^r_{a,\ell}(S_t)\right)$, then $\pi^C_{a,\ell'}(S_t) = \pi^C_{a,\ell}(S_t)$.

DD1 together with **DC1**, **DD2** together with **DC2**, and **DD3** together with **DC3** imply, respectively, **P.E.2.1**, **P.E.2.2**, and **P.E.2.3**. □

Proposition E.2 tells us that $\left(\tau^o_{a,\ell}(S_t), \tau^r_{a,\ell}(S_t)\right)$ may provide sufficient information to compare different locations in terms of agents' location preferences.

The previous results are silent on whether agent a will choose a strategy prescribing to play AllD or TFT under intuition. If given a game-play revision opportunity agent a will decide to play TFT under intuition if $\pi^C_{a,\ell}(S_t) > \pi^D_{a,\ell}(S_t)$ and, so, if

$$\begin{aligned}(1-G(k^d))p(b-c)\tau^r_{a,\ell}(S_t) + \int_0^{k^d} tg(t)\ \mathrm{d}t > \dots \\ \dots > (1-G(k^c))(1-p)c + \int_0^{k^c} tg(t)\ \mathrm{d}t\end{aligned} \tag{E.3}$$

where the LHS contains all the elements that depend on $\tau^r_{a,\ell}(S_t)$ as $k^d = p(b-c)\tau^r_{a,\ell}(S_t)$. The derivative of the LHS with respect to $\tau^r_{a,\ell}(S_t)$ is

$$\begin{aligned}\frac{\partial LHS}{\partial \tau^r_{a,\ell}(S_t)} &= (1-G(k^d))p(b-c) - g(k^d)k^d p(b-c) + g(k^d)k^d p(b-c) \\ &= (1-G(k^d))p(b-c) \geq 0\end{aligned}$$

Therefore, the higher $\tau^r_{a,\ell}(S_t)$ the higher the likelihood that Equation (E.3) holds. To make this more salient we can also express Equation (E.3) as

$$\tau^r_{a,\ell}(S_t) > \frac{(1-G(k^c))(1-p)c + \int_{k^d}^{k^c} tg(t)\ \mathrm{d}t}{(1-G(k^d))p(b-c)} \tag{E.4}$$

We now introduce an ordering of agents according to their level of cooperation and, in particular, by their probability to provide the benefit of cooperation.

We say that agent a' is at least as cooperative as agent a if the following inequalities hold:

$$\begin{aligned}G(k_{a',t})r_{a',t} + (1-G(k_{a',t}))i_{a',t} \geq G(k_{a,t})r_{a,t} + (1-G(k_{a,t}))i_{a,t}) \\ G(k_{a',t})o_{a',t} + (1-G(k_{a',t}))i_{a',t} \geq G(k_{a,t})o_{a,t} + (1-G(k_{a,t}))i_{a,t})\end{aligned}$$

Moreover, we say that agent a' is more cooperative than agent a if it is at least as cooperative as agent a but agent a is not at least as cooperative as agent a'. Finally, we say that agent a' is maximally cooperative if there does not exist any agent a that is more cooperative than a'. Note that there may be more than one maximally cooperative agent.

In the following we will often use the concept of maximally cooperative agent by considering only the subset of agents staying in a given location; in that case, we will say that agent a' is maximally cooperative in location ℓ.

Proposition E.3. *Assume that agent a' and agent a stay in the same location ℓ'. If agent a' is at least as cooperative as agent a and if Equation* (E.3) *holds in location ℓ for agent a', then Equation* (E.3) *must necessarily hold in location ℓ for agent a too.*

Proof. Assume that agent a' and agent a stay in the same location ℓ', that agent a' is at least as cooperative as agent a, and that Equation (E.3) holds in location ℓ for agent a'.

For any location $\ell \in \mathcal{L}$ we can express $\tau^r_{a',\ell}(S_t)$ as

$$\tau^r_{a',\ell}(S_t) = \frac{1}{A_{\ell,t} - I_{a',\ell,t}} \Big[Const + I_{a,\ell,t} \big[G(k_{a,t}) r_{a,t} + (1 - G(k_{a,t})) i_{a,t} \big] \Big]$$

and $\tau^r_{a,\ell}(S_t)$ as

$$\tau^r_{a,\ell}(S_t) = \frac{1}{A_{\ell,t} - I_{a,\ell,t}} \Big[Const + I_{a',\ell,t} \big[G(k_{a',t}) r_{a',t} + (1 - G(k_{a',t})) i_{a',t} \big] \Big]$$

where

$$Const = \sum_{\substack{\alpha=1 \\ \alpha \neq a,a'}}^{A_{\ell,t}} \big[G(k_{\alpha,t}) r_{\alpha,t} + (1 - G(k_{\alpha,t})) i_{\alpha,t} \big]$$

But given that agent a' is at least as cooperative as agent a, it must be that

$$G(k_{a',t}) r_{a',t} + (1 - G(k_{a',t})) i_{a',t} \geq G(k_{a,t}) r_{a,t} + (1 - G(k_{a,t})) i_{a,t})$$

which together with $I_{a',\ell,t} = I_{a,\ell,t}$ for every $\ell \in \mathcal{L}$ – because agent a' and agent a stay in the same location – implies that $\tau^r_{a,\ell}(S_t) \geq \tau^r_{a',\ell}(S_t)$ for every $\ell \in \mathcal{L}$. But then, for every location $\ell \in \mathcal{L}$ it must be

$$(1 - G(k^d_{a,t})) p(b-c) \tau^r_{a,\ell}(S_t) + \int_0^{k^d_{a,t}} t g(t) \, \mathrm{d}t \geq$$

$$\ldots \geq (1 - G(k^d_{a',t})) p(b-c) \tau^r_{a',\ell}(S_t) + \int_0^{k^d_{a',t}} t g(t) \, \mathrm{d}t$$

as this expression is increasing in $\tau^r_{a,\ell}(S_t)$.

Given that by assumption Equation (E.3) holds in location ℓ for agent a', it must be

$$(1 - G(k^d_{a',t}))p(b-c)\tau^r_{a',\ell}(S_t) + \int_0^{k^d_{a',t}} tg(t)\,\mathrm{d}t > \ldots$$

$$\ldots > (1 - G(k^c))(1-p)c + \int_0^{k^c} tg(t)\,\mathrm{d}t$$

which together with the previous result implies

$$(1 - G(k^d_{a,t}))p(b-c)\tau^r_{a,\ell}(S_t) + \int_0^{k^d_{a,t}} tg(t)\,\mathrm{d}t > \ldots$$

$$\ldots > (1 - G(k^c))(1-p)c + \int_0^{k^c} tg(t)\,\mathrm{d}t$$

and, so, Equation (E.3) must hold in location ℓ for agent a too.

Hence, if agent a' is at least as cooperative as agent a and if Equation (E.3) holds in location ℓ for agent a', then Equation (E.3) must necessarily hold in location ℓ for agent a too. ☐

Corollary E.2. *Assume that agent a' is maximally cooperative in its current location ℓ'. If Equation* (E.3) *holds in location ℓ for agent a', then Equation* (E.3) *must necessarily hold in location ℓ for every other agent staying in location ℓ'.*

The result follows immediately from Proposition E.3 and the definition of maximally cooperative agent in a given location. Corollary E.2 tells us that if an agent that is maximally cooperative in a given location $\ell' \in \mathcal{L}$ is willing to adopt a dual process cooperation strategy if moving (or staying) in location $\ell'' \in \mathcal{L}$ then any other agent staying in location ℓ' must be willing to adopt a dual process cooperation strategy if moving or staying in location $\ell'' \in \mathcal{L}$.

At this point we can provide the prove of our first main result.

Theorem 4. *Consider the system under unperturbed dynamics ($\varepsilon = 0$). The absorbing sets of the system are:*

- *($\mathcal{L}$, $\mathbf{0}$, $\mathbf{0}$, $\mathbf{1}$, $\mathbf{0}$) if $p \in (0,1)$;*
- *($\boldsymbol{\ell}^*$, $\mathbf{0}$, $\mathbf{0}$, $\mathbf{1}$, $\mathbf{k^d}$) if $p \in \left(0, \frac{c}{c+G(k^d)(b-c)}\right)$ with every k^d such that $k^d = p(b-c)G(k^d)$ if any;*

- *($\boldsymbol{\ell}^*$, $\mathbf{1}$, $\mathbf{0}$, $\mathbf{1}$, $\mathbf{k^c}$) if $p \in \left(\frac{c}{b}, 1\right)$ with the unique $k^c = (1-p)c$.*

There are no other absorbing sets.

Proof.

Lemma E.1. *The intuitive defection set ($\mathcal{L}$, $\mathbf{0}$, $\mathbf{0}$, $\mathbf{1}$, $\mathbf{0}$) is absorbing if $p \in (0, 1)$.*

Proof. Assume that at time t the system is in a state belonging to the set ($\mathcal{L}$, $\mathbf{0}$, $\mathbf{0}$, $\mathbf{1}$, $\mathbf{0}$). In such state each player adopts an intuitive defection strategy and stays in a randomly chosen location. Therefore, it must be $\tau^r_{a,\ell}(S_t) = 0$ and $\tau^o_{a,\ell}(S_t) = 0$ for each agent a and for each location ℓ. But then, by Proposition E.2 it must be $\ell' \sim_{a,S_t} \ell$ for each $\ell \in \mathcal{L}$ and for every $a \in \mathcal{A}$. Consequently, if at the end of time t an agent $a \in \mathcal{A}$ is given a full-strategy revision opportunity then such agent is indifferent between staying in its current location and moving into any other location. Moreover, by Corollary E.1

$$BR_a(S_t) \subseteq \left\{\{(\ell, 0, 0, 1, k^d), (\ell, 1, 0, 1, k^c)\}_{\ell=1}^L\right\}$$

with $k^d = p(b-c)\tau^r_{a,\ell}(S_t) = 0$ and $k^c = (1-p)c$ for each location $\ell \in \mathcal{L}$.

Adopting an intuitive defection strategy (as $k^d = 0$) in any location ℓ provides a null expected payoff, while adopting a dual process cooperation strategy in any location ℓ provides an expected payoff of

$$\pi^C_{a,\ell}(S_t) = -(1 - G(k^c))(1-p)c - \int_0^{k^c} tg(t)\, \mathrm{d}t < 0$$

But then, no agent has an incentive to adopt a dual process cooperation strategy and

$$BR_a(S_t) \subseteq \left\{\{(\ell, 0, 0, 1, k^d)\}_{\ell=1}^L\right\}$$

with $k^d = p(b-c)\tau^r_{a,\ell}(S_t) = 0$ for every location $\ell \in \mathcal{L}$. Therefore, if the system is in a state belonging to the intuitive defection set then every player $a \in \mathcal{A}$ best replies to the current state of the system by keeping an intuitive defection strategy randomizing the location-choice. But then, the system will remain within the intuitive defection set.

Hence, the intuitive defection set ($\mathcal{L}$, $\mathbf{0}$, $\mathbf{0}$, $\mathbf{1}$, $\mathbf{0}$) is absorbing if $p \in (0, 1)$. □

Lemma E.2. *Dual process defection states ($\boldsymbol{\ell}^*$, $\mathbf{0}$, $\mathbf{0}$, $\mathbf{1}$, $\mathbf{k^d}$) are absorbing if $p \in \left(0, \frac{c}{c+G(k^d)(b-c)}\right)$ with every $k^d : k^d = p(b-c)G(k^d)$ if any.*

Proof. Assume that at time t the system is in a state of type ($\boldsymbol{\ell}^*$, **0**, **0**, **1**, $\mathbf{k^d}$). In such state every agent plays AllD under intuition, stays in location ℓ^*, and sets its threshold cost of deliberation to its optimal value $k^d = p(b-c)G(k^d) > 0$[1]. Note that depending on $G(\cdot)$ there may be none, one or multiple values of $k^d > 0$ such that $k^d = p(b-c)G(k^d)$. In the following we consider the case in which $G(\cdot)$ is such that there exists at least one $k^d > 0$ satisfying the previous condition. If this is the case, in a dual process defection state ($\boldsymbol{\ell}^*$, **0**, **0**, **1**, $\mathbf{k^d}$) for each agent $a \in \mathcal{A}$ it must be

- $\tau^r_{a,\ell}(S_t) = 0$ and $\tau^o_{a,\ell}(S_t) = 0$ for each location $\ell \neq \ell^*$;
- $\tau^r_{a,\ell^*}(S_t) = G(k^d)$ and $\tau^o_{a,\ell^*}(S_t) = 0$

But then, by Proposition E.2 it must be $\ell^* \succ_{a,S_t} \ell$ for each agent $a \in \mathcal{A}$ and each location $\ell \in \mathcal{L}, \ell \neq \ell^*$. Consequently, every agent a strictly prefers staying in its current location ℓ^* (i.e., no agent has an incentive to change its location-choice). This together with Corollary E.1 implies that for each agent $a \in \mathcal{A}$

$$BR_a(S_t) \subseteq \{(\ell^*, 0, 0, 1, k^d), (\ell^*, 1, 0, 1, k^c)\}$$

with $k^d = p(b-c)G(k^d) > 0$ and $k^c = (1-p)c$.

Adopting a dual process defection strategy in location ℓ^* with the optimal threshold cost of deliberation $k^d = p(b-c)G(k^d) > 0$ provides an expected payoff of

$$\pi^D_{a,\ell^*}(S_t) = G(k^d)p(b-c)G(k^d) - \int_0^{k^d} tg(t)\,\mathrm{d}t$$

while adopting a dual process cooperation strategy in location ℓ^* with the optimal threshold cost of deliberation $k^c = (1-p)c$ provides an expected payoff of

$$\pi^C_{a,\ell^*}(S_t) = p(b-c)G(k^d) - (1-G(k^c))(1-p)c - \int_0^{k^c} tg(t)\,\mathrm{d}t$$

Agents prefer keeping a dual process defection strategy if

$$(1-G(k^d))p(b-c)G(k^d) - (1-G(k^c))(1-p)c < \int_{k^d}^{k^c} tg(t)\,\mathrm{d}t$$

---

[1] For the case $k^d = p(b-c)G(k^d) = 0$ see Lemma E.1.

Integration by parts yields

$$\int_{k^d}^{k^c} tg(t)\ \mathrm{d}t = G(k^c)(1-p)c - G(k^d)p(b-c)G(k^d) - \int_{k^d}^{k^c} G(z)\ \mathrm{d}z$$

But then, the previous inequality is satisfied if and only if

$$\begin{aligned}(1-G(k^d))p(b-c)G(k^d) - (1-G(k^c))(1-p)c < \dots\\ \dots < G(k^c)(1-p)c - G(k^d)p(b-c)G(k^d) - \int_{k^d}^{k^c} G(z)\ \mathrm{d}z \Leftrightarrow\\ \Leftrightarrow p(b-c)G(k^d) - (1-p)c < -\int_{k^d}^{k^c} G(z)\ \mathrm{d}z \Leftrightarrow\\ \Leftrightarrow -\int_{k^d}^{k^c} 1\ \mathrm{d}z < -\int_{k^d}^{k^c} G(z)\ \mathrm{d}z \Leftrightarrow\\ \Leftrightarrow \int_{k^d}^{k^c} \big(1-G(z)\big)\ \mathrm{d}z > 0\end{aligned}$$

which holds if and only if $k^d < k^c$ and, so,

$$\Leftrightarrow p(b-c)G(k^d) < (1-p)c \Leftrightarrow p < \frac{c}{c+G(k^d)(b-c)}$$

Therefore, if $p \in \big(0, \frac{c}{c+G(k^d)(b-c)}\big)$ and the system is in a dual process defection state it must be $BR_a(S_t) = (\ell^*, 0, 0, 1, k^d)$ with $k^d = p(b-c)G(k^d)$ for every $a \in \mathcal{A}$. But then, the system will remain in the dual process defection state $(\boldsymbol{\ell}^*, \mathbf{0}, \mathbf{0}, \mathbf{1}, \mathbf{k^d})$.

Hence, dual process defection states $(\boldsymbol{\ell}^*, \mathbf{0}, \mathbf{0}, \mathbf{1}, \mathbf{k^d})$ are absorbing if $p \in \big(0, \frac{c}{c+G(k^d)(b-c)}\big)$ with every $k^d : k^d = p(b-c)G(k^d)$ if any. □

Lemma E.3. *Dual process cooperation states* $(\boldsymbol{\ell}^*, \mathbf{1}, \mathbf{0}, \mathbf{1}, \mathbf{k^c})$ *are absorbing if* $p \in (\frac{c}{b}, 1)$ *with the unique* $k^c = (1-p)c$.

Proof. Assume that at time t the system is in a dual process cooperation state $(\boldsymbol{\ell}^*, \mathbf{1}, \mathbf{0}, \mathbf{1}, \mathbf{k^c})$. In such state every agent plays TFT under intuition, stays in location ℓ^*, and sets its threshold cost of deliberation to the optimal value $k^c = (1-p)c$. But then, for every agent $a \in \mathcal{A}$ it must be

- $\tau^r_{a,\ell}(S_t) = 0$ and $\tau^o_{a,\ell}(S_t) = 0$ for each location $\ell \neq \ell^*$;
- $\tau^r_{a,\ell^*}(S_t) = 1$ and $\tau^o_{a,\ell^*}(S_t) = 1 - G(k^c)$.

But then, by Proposition E.2 it must be $\ell^* \succ_{a,S_t} \ell$ for each agent $a \in \mathcal{A}$ and each location $\ell \in \mathcal{L}, \ell \neq \ell^*$. Consequently, every agent a strictly prefers staying in its current location (i.e., no agent has an incentive to change its location-choice). This together with Corollary E.1 implies that for each agent $a \in \mathcal{A}$

$$BR_a(S_t) \subseteq \{(\ell^*, 0, 0, 1, k^d), (\ell^*, 1, 0, 1, k^c)\}$$

with $k^d = p(b-c)$ and $k^c = (1-p)c$.

Adopting a dual process defection strategy in location ℓ^* with the optimal threshold cost of deliberation $k^d = p(b-c)$ provides an expected payoff of

$$\pi^D_{a,\ell^*}(S_t) = G(k^d)p(b-c) + (1-p)b(1-G(k^c)) - \int_0^{k^d} tg(t)\ \mathrm{d}t$$

while adopting a dual process cooperation strategy in location ℓ^* with the optimal threshold cost of deliberation $k^c = (1-p)c$ provides an expected payoff of

$$\pi^C_{a,\ell^*}(S_t) = p(b-c) + (1-p)b(1-G(k^c)) - (1-G(k^c))(1-p)c - \int_0^{k^c} tg(t)\ \mathrm{d}t$$

Agents prefer keeping a dual process cooperation strategy if

$$p(b-c) - (1-G(k^c))(1-p)c - \int_0^{k^c} tg(t) > G(k^d)p(b-c) - \int_0^{k^d} tg(t)\ \mathrm{d}t \Leftrightarrow$$

$$\Leftrightarrow (1-G(k^d))p(b-c) - (1-G(k^c))(1-p)c > \int_{k^d}^{k^c} tg(t)\ \mathrm{d}t$$

Integration by parts yields

$$\int_{k^d}^{k^c} tg(t)\ \mathrm{d}t = G(k^c)(1-p)c - G(k^d)p(b-c) - \int_{k^d}^{k^c} G(t)\ \mathrm{d}t$$

But then, the previous inequality is satisfied if and only if

$$\Leftrightarrow p(b-c) - (1-p)c > -\int_{k^d}^{k^c} G(z)\,\mathrm{d}z \Leftrightarrow$$
$$\Leftrightarrow -\int_{k^d}^{k^c} 1\,\mathrm{d}z > -\int_{k^d}^{k^c} G(z)\,\mathrm{d}z \Leftrightarrow$$
$$\Leftrightarrow \int_{k^d}^{k^c} \big(1 - G(z)\big)\,\mathrm{d}z < 0$$

which holds if and only if $k^d > k^c$ and, so, if

$$\Leftrightarrow p(b-c) > (1-p)c \Leftrightarrow p > \frac{c}{b}$$

Therefore, if $p \in (\frac{c}{b}, 0)$ it must be $BR_a(S_t) = (\ell^*, 1, 0, 1, k^c)$ with $k^c = (1-p)c$ for each agent $a \in \mathcal{A}$. Consequently, the system will remain in the dual process cooperation state ($\boldsymbol{\ell}^*$, $\mathbf{1}$, $\mathbf{0}$, $\mathbf{1}$, $\mathbf{k^c}$).

Hence, dual process cooperation states ($\boldsymbol{\ell}^*$, $\mathbf{1}$, $\mathbf{0}$, $\mathbf{1}$, $\mathbf{k^c}$) are absorbing if $p \in (\frac{c}{b}, 1)$ with the unique $k^c = (1-p)c$. □

Lemma E.4. *There are no other absorbing sets than the ones in Lemma E.1, Lemma E.3, and Lemma E.2.*

Proof. Assume that the system is in a generic state of the system S_0 that does not belong to the absorbing sets ($\mathcal{L}$, $\mathbf{0}$, $\mathbf{0}$, $\mathbf{1}$, $\mathbf{0}$), ($\boldsymbol{\ell}^*$, $\mathbf{0}$, $\mathbf{0}$, $\mathbf{1}$, $\mathbf{k^d}$) with every $k^d : k^d = p(b-c)G(k^d)$, and ($\boldsymbol{\ell}^*$, $\mathbf{1}$, $\mathbf{0}$, $\mathbf{1}$, $\mathbf{k^c}$) with the unique $k^c = (1-p)c$.

With positive probability at the end of time 0 all the players will be given a game-play revision opportunity. In either case, by Corollary E.1 at the beginning of time 1 all the players will adopt either a strategy of type $(\ell_{a,1}, 0, 0, 1, k^d)$ with $k^d = p(b-c)\tau^r_{a,\ell_{a,1}}(S_0) \geq 0$ or of type $(\ell_{a,1}, 1, 0, 1, k^c)$ with $k^c = (1-p)c$.

We now show that with positive probability the system can reach a state in which each non-empty location ℓ is characterized by having all agents adopting the same strategy among $(\ell, 1, 0, 1, k^c)$ with $k^c = (1-p)c$, $(\ell, 0, 0, 1, k^d)$ with $k^d > 0 : k^d = p(b-c)G(k^d)$, and $(\ell, 0, 0, 1, 0)$.

For each non-empty location ℓ, let $a', a'', \ldots$ denote the agents currently in location ℓ ordered from the maximally cooperative to the least cooperative (eventually breaking ties randomly). Then, fixed a location

ℓ, with positive probability at the end of time 1 only agent a', a maximally cooperative agent in location ℓ, will be given a game-play revision opportunity. Then, the following scenarios are possible:

1 Agent a' adopts or keeps strategy $(\ell, 1, 0, 1, k^c)$. Given that a' is a maximally cooperative agent in location ℓ, then it is at least as cooperative as any other agent in location ℓ; moreover, given that a' has adopted or kept a dual process cooperation strategy Equation (E.3) must hold for agent a'. Consequently, by Corollary E.2 Equation (E.3) must hold for any other agent in location ℓ. But then, with positive probability at the end of time 2 all the agents staying in location ℓ (eventually except agent a') will be given a game-play revision opportunity and each of them will adopt strategy $(\ell, 1, 0, 1, k^c)$. But then, with positive probability the system will reach a state in which each agent in location ℓ adopts strategy $(\ell, 1, 0, 1, k^c)$ with $k^c = (1-p)c$.

2 Agent a' adopts an intuitive defection strategy. But then, it must be that $k^d = p(b-c)\tau_{a',\ell}(S_1) = 0$ and, so, all the other agents in location ℓ must be adopting an intuitive defection strategy. Moreover, given that agent a' has adopted an intuitive defection strategy too, now all the agents in location ℓ adopt strategy $(\ell, 0, 0, 1, 0)$.

3 Consider the case in which agent a' adopts a dual process defection strategy $(\ell, 0, 0, 1, k^d)$ with $k^d > 0$. Then, with positive probability at the end of period 2 only agent a'', who at the end of time 1 was the second maximally cooperative agent, will be given a game-play revision opportunity. Two scenarios are possible[2]:

   3.1 Agent a'' adopts a dual process cooperation strategy. But then, at the end of time 3 agent a'' will be at least as cooperative as any the other agent in location ℓ; moreover, Equation (E.3) must hold for agent a''. But then, by Corollary E.2 Equation (E.3) must hold for any other agent staying in location ℓ. Then, with positive probability at the end of time 3 every agent in location ℓ (eventually except a'') will be given a game-play revision opportunity and each of them will adopt a dual

---

[2] Note that agent a'' cannot adopt an intuitive defection strategy as agent a' has adopted a dual process defection strategy at the end of the previous period and, consequently, $k^d = p(b-c)\tau_{a'',\ell}(S_1) > 0$.

process cooperation strategy. But then, with positive probability the system will reach a state in which each agent in location ℓ adopts strategy $(\ell, 1, 0, 1, k^c)$ with $k^c = (1-p)c$.

3.2 Agent a'' adopts a dual process defection strategy. In this case we repeat all the steps from 3 by considering the next player in the original ranking. If this process is repeated until all the players in location ℓ have been given a second revision opportunity, then the system will have reached a situation in which all the agents in location ℓ adopt a dual process defection strategy (eventually with agents choosing different threshold costs of deliberation). Now,

  3.2.1 If at least one agent has the incentive to adopt a dual process cooperation strategy, then by the arguments in point 1 with positive probability all the agents in location ℓ will eventually adopt a dual process cooperation strategy;

  3.2.2 If no agent has the incentive to adopt a dual process cooperation strategy, then let $\mathcal{D}_{\ell,t}$ denote the set of agents in location ℓ that are currently willing to decrease their threshold cost of deliberation if given a game-play revision opportunity. Then, with positive probability all the agents in $\mathcal{D}_{\ell,t}$ will be given a game-play revision opportunity at the end of the current period of time and by construction each of them will decrease its threshold cost of deliberation. With positive probability this process of providing a game-play revision opportunity only to players willing to decrease their threshold cost of deliberation will continue until the set $\mathcal{D}_{\ell,t}$ will become empty. At this point three scenarios are possible: (a) all agents in location ℓ have adopted an intuitive defection strategy $(\ell, 0, 0, 1, 0)$ (ii), (b) all agents in location ℓ have adopted a dual process defection strategy $(\ell, 0, 0, 1, k^d)$ with $k^d > 0 : k^d = p(b-c)G(k^d)$, or (c) all the agents in location ℓ are now willing to increase (or keep unchanged) their threshold cost of deliberation. In this last scenario, with positive probability at the end of the next period of time all the agents in location ℓ will be given a game-play revision opportunity and each of them will increase (or keep unchanged) its threshold cost of deliberation. This process will continue with positive probability until either (a′) all agents in location ℓ adopt a dual process defection strategy $(\ell, 0, 0, 1, k^d)$ with $k^d > 0 : k^d = p(b - c)G(k^d)$ or (b′) all agents in location ℓ adopt a dual process cooperation strategy $(\ell, 1, 0, 1, k^c)$.

By repeating these steps separately for each non-empty location, with positive probability the system will reach a state in which in each non-empty location all the players adopt the same strategy among $(\ell, 1, 0, 1, k^c)$ with $k^c = (1 - p)c$, $(\ell, 0, 0, 1, k^d)$ with $k^d > 0 : k^d = p(b - c)G(k^d)$ and $(\ell, 0, 0, 1, 0)$. However, in different locations different strategies may have emerged. Then,

- If in all the non-empty locations the intuitive defection strategy has emerged, then the system has reached the intuitive defection set (Lemma E.1);

- If in at least one location, say ℓ', the dual process defection strategy has emerged[3], but in no location the dual process cooperation strategy has emerged, then the system will reach a dual process defection absorbing state. More precisely, such absorbing state will be characterized by the highest $k^d > 0 : k^d = p(b - c)G(k^d)$ that has emerged. To see this denote with $k^{D'}$ the maximum threshold cost of deliberation that has emerged and denote with ℓ' the location in which it has emerged (randomly selecting one if $k^{D'}$ has emerged in more than one location). Then, for every agent $a \in \mathcal{A}$ it must be $\tau^r_{a,\ell'}(S_t) = G(k^{D'}) \geq \tau^r_{a,\ell}(S_t)$ and $\tau^o_{a,\ell'}(S_t) = 0 \geq 0 = \tau^r_{a,\ell}(S_t)$ for all $\ell \neq \ell'$ as by construction $k^{D'}$ is the highest threshold. Consequently, if at the end of the current period of time all the agents not staying in location ℓ' will be given a full-strategy revision opportunity, then by Proposition E.2 and Lemma E.2 they will all adopt strategy $(\ell', 0, 0, 1, k^{D'})$ and, so, the system will have reached a dual process defection absorbing state.

- If in at least one location, say ℓ', the dual process cooperation strategy has emerged[4], then the system will reach a dual process cooperation absorbing state. In fact, for every agent it must be $\tau^r_{a,\ell'}(S_t) =$

---

[3] Note that in this scenario it must be $p \in \left(0, \frac{c}{c+G(k^d)(b-c)}\right)$; otherwise in no location the dual process defection strategy would have emerged.

[4] Note that in this scenario it must be that $p \in (\frac{c}{b}, 1)$; otherwise, in no location the dual process cooperation strategy would have emerged.

$1 \geq \tau^r_{a,\ell}(S_t)$ and $\tau^o_{a,\ell'}(S_t) = (1 - G(k^c)) \geq \tau^r_{a,\ell}(S_t)$ for all $\ell \neq \ell'$. But then, with positive probability at the end of the current period of time all the agents staying in location $\ell \neq \ell'$ will be given a full-strategy revision opportunity and then by Proposition E.2 and by Lemma E.3 they will all adopt strategy $(\ell', 1, 0, 1, k^c)$ and, so, the system will have reached a dual process cooperation absorbing state.

We have shown that starting from a generic state of the system that does not belong to the absorbing sets ($\mathcal{L}$, **0**, **0**, **1**, **0**), ($\boldsymbol{\ell}^*$, **0**, **0**, **1**, $\mathbf{k^d}$) with every $k^d : k^d = p(b - c)G(k^d)$, and ($\boldsymbol{\ell}^*$, **1**, **0**, **1**, $\mathbf{k^c}$) with the unique $k^c = (1 - p)c$ the system will enter with positive probability either a dual process cooperation absorbing state, or a dual process defection absorbing state, or the intuitive defection state.

Hence, there are no other absorbing sets other than the ones in Lemma E.1, Lemma E.3, and Lemma E.2.

□

The results in Lemma E.1, Lemma E.3, Lemma E.2, and Lemma E.4 together imply Theorem 4. □

Appendix F

Chapter 3: Proof of Theorem 5

Theorem 5. *Consider the system under perturbed dynamics* ($\varepsilon > 0$)*. If* $p \in \left(\frac{c}{b}, 1\right)$ *then all the stochastically stable states of the system are contained in the set of dual process cooperation states* $\{(\boldsymbol{\ell}^*, \mathbf{1}, \mathbf{0}, \mathbf{1}, \mathbf{k^c})\}_{\boldsymbol{\ell}^*=\mathbf{1}}^{\mathbf{L}}$ *with* $\mathbf{k^c} = (k^c, \ldots, k^c)$ *and* $k^c = (1-p)c$*.*

Proof. Consider the system under perturbed dynamics ($\varepsilon > 0$) and assume $p \in (\frac{c}{b}, 1)$. Let Ω and Ω^{-1} be a partition of the set of absorbing sets of the system under unperturbed dynamics (see Theorem 4) such that Ω is the set of all the dual process cooperation states of type ($\boldsymbol{\ell}$, $\mathbf{1}$, $\mathbf{0}$, $\mathbf{1}$, $\mathbf{k^c}$) with $\mathbf{k^c} = (k^c, \ldots, k^c)$ and $k^c = (1-p)c$ while Ω^{-1} contains the intuitive defection set ($\mathcal{L}$, $\mathbf{0}$, $\mathbf{0}$, $\mathbf{1}$, $\mathbf{0}$) and all the dual process defection states of type ($\boldsymbol{\ell}$, $\mathbf{0}$, $\mathbf{0}$, $\mathbf{1}$, $\mathbf{k^d}$) with $\mathbf{k^d} = (k^d, \ldots, k^d)$ and $k^d = p(b-c)G(k^d)$ for any $k^d : k^d = p(b-c)G(k^d)$ if any.

Given this partition of the set of absorbing sets of the system under unperturbed dynamics, we show that under perturbed dynamics one mistake can lead with positive probability the system into the basin of attraction of Ω starting from any state in Ω^{-1} and, in particular, from any state belonging to the intuitive defection set ($\mathcal{L}$, $\mathbf{0}$, $\mathbf{0}$, $\mathbf{1}$, $\mathbf{0}$) [1] and from any dual process defection state of type ($\boldsymbol{\ell}$, $\mathbf{0}$, $\mathbf{0}$, $\mathbf{1}$, $\mathbf{k^d}$) with $k^d = p(b-c)G(k^d)$ [2]. This in turn implies that the coradius of Ω is equal to one. We then show that one mistake is never enough to lead with positive probability the system outside the basin of attraction of Ω starting from any dual process cooperation state of type ($\boldsymbol{\ell}$, $\mathbf{1}$, $\mathbf{0}$, $\mathbf{1}$, $\mathbf{k^c}$) with

$k^c = (1-p)c$ [3]. But then the radius of Ω must be strictly larger than one.

[1] ***Minimum number of mistakes from the intuitive defection set to*** Ω***:*** Assume that the system is in a state belonging to the intuitive defection set ($\mathcal{L}$, **0**, **0**, **1**, **0**). In such a state, if given a full-strategy revision opportunity every agent best replies to the current state of the system by keeping an intuitive defection strategy and randomizing the location choice. But then, with positive probability (without any mistake) the system will reach a state – still belonging to the intuitive defection set – in which at least one location, say ℓ', is empty. Once the system has reached a state of this kind assume that at the end of the next period of time t an agent, say a', is given a full-strategy revision opportunity and by mistake it adopts a dual process cooperation strategy in the empty location ℓ', i.e., it adopts strategy (ℓ', $1, 0, 1, k^c$) with $k^c = (1-p)c$.

In period $t+1$ the system will be in a state in which agent a' is the only agent in location ℓ' and it adopts a dual process cooperation strategy, while all other agents are distributed across the remaining locations (or a subset of them) and each of such agents adopts an intuitive defection strategy. But then, for every agent $a \neq a'$ it must be

$$\big(\tau^o_{a,\ell'}(S_{t+1}), \tau^r_{a,\ell'}(S_{t+1})\big) = (1-G(k^c), 1) > (0,0) = \big(\tau^o_{a,\ell}(S_{t+1}), \tau^r_{a,\ell}(S_{t+1})\big)$$

for each location $\ell \neq \ell'$ and, so, by Proposition E.2 for each agent $a \neq a'$ it must be $\ell' \succ_{a,t+1} \ell$ for each $\ell \neq \ell'$. Therefore, if an agent $a \neq a'$ is given a full-strategy revision opportunity, then it will best reply to the current state of the system by moving into location ℓ'. This together with Corollary E.1 implies that, if given a full-strategy revision opportunity, each agent $a \neq a'$ will either adopt strategy (ℓ', 1, 0, 1, k^c) with $k^c = (1-p)c$ or strategy (ℓ', 0, 0, 1, k^d) with $k^d = p(b-c)$. However, given $p \in (\frac{c}{b}, 1)$ they will strictly prefer adopting a dual process cooperation strategy.

But then, with positive probability at the end of time $t+1$ every agent $a \neq a'$ will be given a full-strategy revision opportunity and each of them will best reply to the current state of the system by adopting strategy (ℓ', $1, 0, 1, k^c$) with $k^c = (1-p)c$. At this point the system will have reached a dual process cooperation state of type ($\boldsymbol{\ell'}$, $\mathbf{1}$, $\mathbf{0}$, $\mathbf{1}$, $\mathbf{k^c}$) with $k^c = (1-p)c$.

Therefore, if $p \in (\frac{c}{b}, 1)$ one mistake can lead with positive probability the system into the basin of attraction of Ω starting from a generic state belonging to the intuitive defection set.

[2] ***Minimum number of mistakes from a dual process defection state to*** Ω***:*** Assume that at the end of time t the system is in a dual process defection state of type ($\boldsymbol{\ell}^*$, **0**, **0**, **1**, $\mathbf{k^d}$) with $k^d : k^d = p(b-c)G(k^d)$. In such a state all agents stay in the same location ℓ^* and adopt a dual process defection strategy.

Assume that at the end of time t an agent, say a', is given a full-strategy revision opportunity and by mistake it adopts a dual process cooperation strategy in an empty location, i.e., it adopts strategy (ℓ', 1, 0, 1, k^c) with $k^c = (1-p)c$.

In period $t+1$ the system will be in a state in which all agents except agent a' stay in location ℓ^* and adopt a dual process defection strategy, while agent a' stays in location ℓ' and adopts a dual process cooperation strategy. But then, for every agent $a \neq a'$ it must be

$$\begin{gathered}\left(\tau^o_{a,\ell'}(S_{t+1}), \tau^r_{a,\ell'}(S_{t+1})\right) = (1-G(k^c), 1)\\ \left(\tau^o_{a,\ell^*}(S_{t+1}), \tau^r_{a,\ell^*}(S_{t+1})\right) = (0, G(k^d))\\ \left(\tau^o_{a,\ell}(S_{t+1}), \tau^r_{a,\ell}(S_{t+1})\right) = (0,0) \ \forall \ell \neq \ell^*, \ell'\end{gathered}$$

But then $\left(\tau^o_{a,\ell'}(S_{t+1}), \tau^r_{a,\ell'}(S_{t+1})\right) > \left(\tau^o_{a,\ell}(S_{t+1}), \tau^r_{a,\ell}(S_{t+1})\right)$ for each $\ell \neq \ell'$ and, so, by Proposition E.2 it must be $\ell' \succ_{a,t+1} \ell$ for each $\ell \neq \ell'$ and for each agent $a \neq a'$. Therefore, if at the end of time $t+1$ an agent $a \neq a'$ is given a full-strategy revision opportunity, then it will best reply to the current state of the system by moving into location ℓ'. This together with Corollary E.1 implies that, if given a full-strategy revision opportunity at the end of period $t+1$, each agent $a \neq a'$ will either adopt strategy (ℓ', 1, 0, 1, k^c) with $k^c = (1-p)c$ or strategy (ℓ', 0, 0, 1, k^d) with $k^d = p(b-c)$. However, given $p \in (\frac{c}{b}, 1)$ each agent $a \neq a'$ will strictly prefer adopting a dual process cooperation strategy.

But then, with positive probability at the end of time $t+1$ every agent $a \neq a'$ will be given a full-strategy revision opportunity and each of them will best reply to the current state of the system by adopting strategy (ℓ', 1, 0, 1, k^c) with $k^c = (1-p)c$. At this point the system will have reached a dual process cooperation state of type ($\boldsymbol{\ell}'$, **1**, **0**, **1**, $\mathbf{k^c}$) with $k^c = (1-p)c$.

Therefore, if $p \in (\frac{c}{b}, 1)$ one mistake can lead with positive probability the system into the basin of attraction of Ω starting from a dual process defection state of type ($\boldsymbol{\ell}$, **0**, **0**, **1**, $\mathbf{k^d}$) with $k^d = p(b-c)G(k^d)$.

Hence, if $p \in (\frac{c}{b}, 1)$ one mistake can lead with positive probability the system into the basin of attraction of Ω starting from any state belonging to Ω^{-1}. Consequently, $CR(\Omega) = 1$.

[3] ***Lower bound for the minimum number of mistakes from Ω to Ω^{-1}:*** Assume that at the end of time t the system is in a dual process cooperation state of type ($\boldsymbol{\ell}^*$, $\mathbf{1}$, $\mathbf{0}$, $\mathbf{1}$, $\mathbf{k^c}$) with $k^c = (1-p)c$. In such state all agents stay in the same location ℓ^* and adopt the optimal dual process cooperation strategy.

Assume that at the end of the current period of time t an agent, say a', is given a revision opportunity and it makes a mistake by deciding to: (i) stay in its current location ℓ^*[1], (ii) play AllD under intuition, and (iii) deliberate if the cost of deliberation is lower than or equal to $k' \geq 0$. In other words, assume that by mistake agent a' adopts strategy (ℓ^*, 0, 0, 1, k').

Then, at time $t+1$ for every agent $a \neq a'$ it must be

$$\big(\tau^o_{a,\ell^*}(S_{t+1}), \tau^r_{a,\ell^*}(S_{t+1})\big) = \left(\frac{(A-2)[1-G(k^c)]}{A-1}, \frac{(A-2)+G(k')}{A-1}\right)$$
$$\big(\tau^o_{a,\ell}(S_{t+1}), \tau^r_{a,\ell}(S_{t+1})\big) = (0,0)\ \forall \ell \neq \ell^*$$

and, so, by Proposition E.2 for every agent $a \neq a'$ and each location $\ell \neq \ell^*$ it will be $\ell^* \succ_{a,t+1} \ell$ and, consequently, if given a full-strategy revision opportunity each agent $a \neq a'$ will decide to stay in location ℓ^*.

Consider the case in which at the end of time $t+1$ an agent $a \neq a'$ is given a revision opportunity. Then, agent a prefers keeping its current dual process cooperation strategy rather than adopting a dual process defection strategy with the optimal threshold cost of deliberation $k'' = p(b-c)\tau^r_{a,\ell^*}(S_{t+1})$ if

$$\begin{aligned}
&\tau^r_{a,\ell^*}(S_{t+1}) > \frac{(1-G(k^c))(1-p)c + \int_{k''}^{k^c} tg(t)\,\mathrm{d}t}{(1-G(k''))p(b-c)} \Leftrightarrow \\
&\Leftrightarrow \frac{(A-2)+G(k')}{A-1} > \dots \\
&\dots > \frac{(1-G(k^c))(1-p)c + \int_{p(b-c)\frac{(A-2)+G(k')}{A-1}}^{(1-p)c} tg(t)\,\mathrm{d}t}{(1-G(k''))p(b-c)}
\end{aligned} \tag{F.1}$$

---

[1] Note that the scenario in which agent a' decides to stay in its current location rather than moving into an empty location is the one in which agents $a \neq a'$ who will be given a revision opportunity in the next period of time have a higher incentive to switch to AllD under intuition. In fact, if a' moved into an empty location then no agent $a \neq a'$ would have an incentive to change location choice unless agent a' adopted a dual process cooperation strategy (or an even more cooperative strategy). But in this case the system would still remain in the basin of attraction of Ω.

Such condition is most demanding for $G(k') = 0$, i.e., in case agent a' has adopted an intuitive defection strategy. Moreover, such condition is satisfied independently of $G(\cdot)$ if

$$\frac{A-2}{A-1} > \frac{(1-G(k^c))(1-p)c + G(k^c)(1-p)c - p(b-c)\frac{A-2}{A-1}G(k'')}{(1-G(k''))p(b-c)} \Leftrightarrow$$

$$\Leftrightarrow \frac{A-2}{A-1}p(b-c) > (1-p)c$$

As the number of agents A grows large this condition converges to $p \in (\frac{c}{b}, 1)$.

But then, if $p \in (\frac{c}{b}, 1)$ and the population is large enough one mistake is never enough to make the system exit with positive probability the basin of attraction of Ω. Hence, the radius of the basin of attraction of Ω is at least equal to two.

We have shown that if the population is large enough and $p \in (\frac{c}{b}, 1)$ then:

1. One mistake can lead with positive probability the system into the basin of attraction of Ω starting from any state belonging to Ω^{-1} and, so, $CR(\Omega) = 1$;
2. One mistake is never enough to lead with positive probability the system outside the basin of attraction of Ω starting from a state belonging to Ω and, so, $R(\Omega) \geq 2$.

But then, we have $R(\Omega) \geq 2 > 1 = CR(\Omega)$. Therefore, if $p \in (\frac{c}{b}, 1)$ and the number of agents A is large enough then all the stochastically stable states of the system are contained in the set of dual process cooperation states $\Omega = \left\{\{(\boldsymbol{\ell}, \mathbf{1}, \mathbf{0}, \mathbf{1}, \mathbf{k^c})\}_{\boldsymbol{\ell}=1}^{L}\right\}$.

Note that all the states belonging to $\Omega = \left\{\{(\boldsymbol{\ell}, \mathbf{1}, \mathbf{0}, \mathbf{1}, \mathbf{k^c})\}_{\boldsymbol{\ell}=1}^{L}\right\}$ are stochastically stable. In fact, we can show that one mistake is enough to lead with positive probability the system into any dual process cooperation state of type $(\boldsymbol{\ell}^*, \mathbf{1}, \mathbf{0}, \mathbf{1}, \mathbf{k^c})$ with $k^c = (1-p)c$ to any other dual process cooperation state, say $(\boldsymbol{\ell}', \mathbf{1}, \mathbf{0}, \mathbf{1}, \mathbf{k^c})$ with $\ell' \neq \ell^*$.

Consider the case in which at time t the system is in a dual process cooperation state of type $(\boldsymbol{\ell}^*, \mathbf{1}, \mathbf{0}, \mathbf{1}, \mathbf{k^c})$ with $k^c = (1-p)c$. Moreover, assume that at the end of time t an agent, say a', is given a full-strategy revision opportunity and by mistake it adopts a dual process cooperation strategy in an empty location, i.e., it adopts strategy $(\ell', 1, 0, 1, k^c)$[2].

---

[2] Any other strategy that is more cooperative than this strategy – as the intuitive cooper-

Then, at time $t+1$ for every agent $a \neq a'$ it will be

$$\begin{aligned}\left(\tau^o_{a,\ell'}(S_{t+1}), \tau^r_{a,\ell'}(S_{t+1})\right) &= (1-G(k^c), 1)\\ \left(\tau^o_{a,\ell^*}(S_{t+1}), \tau^r_{a,\ell^*}(S_{t+1})\right) &= (1-G(k^c), 1)\\ \left(\tau^o_{a,\ell}(S_{t+1}), \tau^r_{a,\ell}(S_{t+1})\right) &= (0,0)\ \forall \ell \neq \ell^*, \ell'\end{aligned}$$

and, so, by Proposition E.2 every agent $a \neq a'$ will be $\ell' \succ_{a,t+1} \ell^*$ and, consequently, if given a full-strategy revision opportunity each agent $a \neq a'$ will be indifferent between staying in its current location or moving into location ℓ' keeping a dual process cooperation strategy[3]. But then, with positive probability at the end of time $t+1$ every agent $a \neq a'$ will be given a full-strategy revision opportunity and each of them will adopt strategy $(\ell', 1, 0, 1, k^c)$ with $k^c = (1-p)c$. Consequently, with one mistake the system will have reached a dual process cooperation state of type $(\boldsymbol{\ell}', \mathbf{1}, \mathbf{0}, \mathbf{1}, \mathbf{k^c})$ with $\ell' \neq \ell^*$.

Given that this reasoning can be applied to any couple of locations (ℓ, ℓ') one mistake can lead with positive probability the system from any dual process cooperation state in Ω to any other dual process cooperation state state in Ω.

Hence, if the population is large enough and $p \in (\frac{c}{b}, 1)$ then all the states belonging to the set $\Omega = \left\{\{(\boldsymbol{\ell}^*, \mathbf{1}, \mathbf{0}, \mathbf{1}, \mathbf{k^c})\}_{\boldsymbol{\ell}^*=1}^{L}\right\}$ are stochastically stable. □

---

ation strategy – would be fine too

[3] This must be the case because $\tau^r_{a,\ell^*}(S_t) = 1$ was such that Equation (E.3) was satisfied in location ℓ^* for each agent. But, given $\tau^r_{a,\ell'}(S_{t+1}) = 1 \geq \tau^r_{a,\ell^*}(S_t)$, Equation (E.3) must hold in location ℓ' at time $t+1$ for every agent $a \neq a'$.

Appendix G

Chapter 4: Different Cost Distributions

I repeat the analyses reported in Chapter 4 by considering different cost distributions. More precisely, in the main text deliberation costs were drawn from a uniform distribution and different widths of the uniform distribution have been considered. Instead, here I fix the width of the distribution to $C = 1$ and consider cost distributions other than the uniform. In Figure G.1 I report the probability density functions (PDF) of the distributions analyzed: a truncated normal distribution, a cost distribution characterized by a continuously increasing PDF, and one with a continuously decreasing PDF. Despite being defined in the same interval (for comparability purposes) these distributions present quite different properties. In particular, if the uniform and the normal distribution present the same expected value, the increasing [decreasing] distribution has a relatively high [low] expected value.

The findings reported in the next figures suggest that the shape of the cost distribution does not play a crucial role. More generally, I expect that the main results hold for any sufficiently smooth deliberation cost distribution – i.e. if the cost distribution is such that small changes in the threshold cost do not lead to huge differences in the probability to incur deliberation.

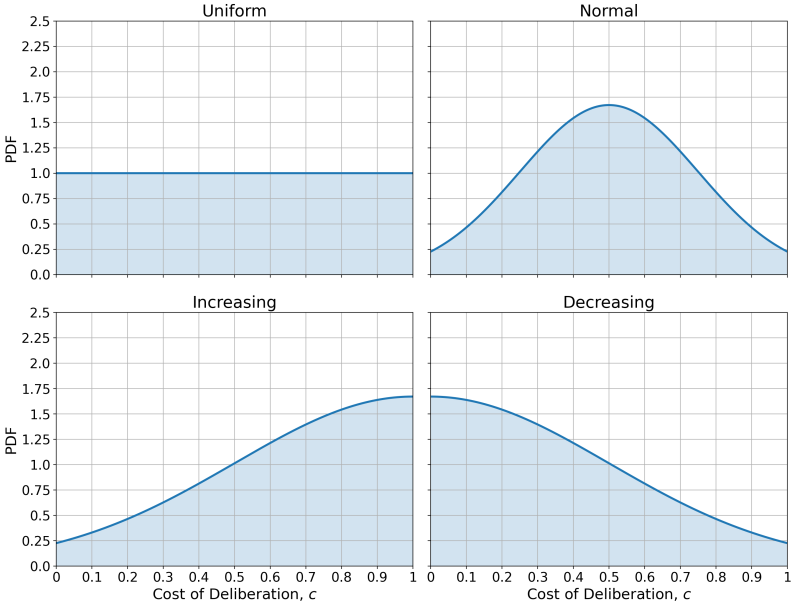


Figure G.1: Distributions analyzed. Probability density functions of the distributions analyzed.

The most significant differences are observed in the case of a decreasing PDF which has a relatively low expected cost of deliberation. The results associated with the decreasing PDF are quite comparable with the ones found in the case of a uniform distribution in the case $C = 0.5$ (see Chapter 4). This suggests that regarding the cost distribution what really matters is the expected cost associated with that distribution rather than the actual shape or the width of the distribution.

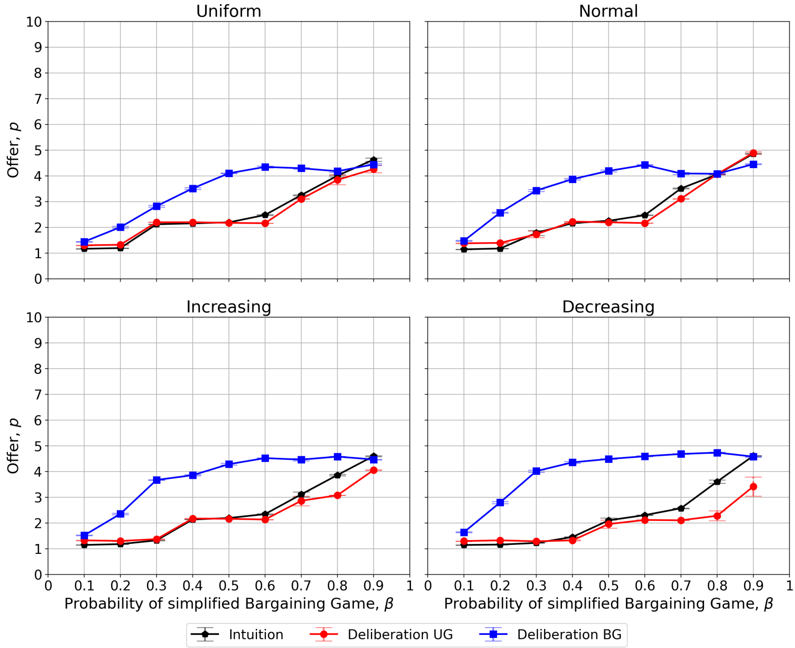


Figure G.2: Proposers' behavior. Proposers' average offer under intuition, under deliberation if the game is an UG, and under deliberation if the game is a simplified BG and their standard errors as a function of β by cost distribution.

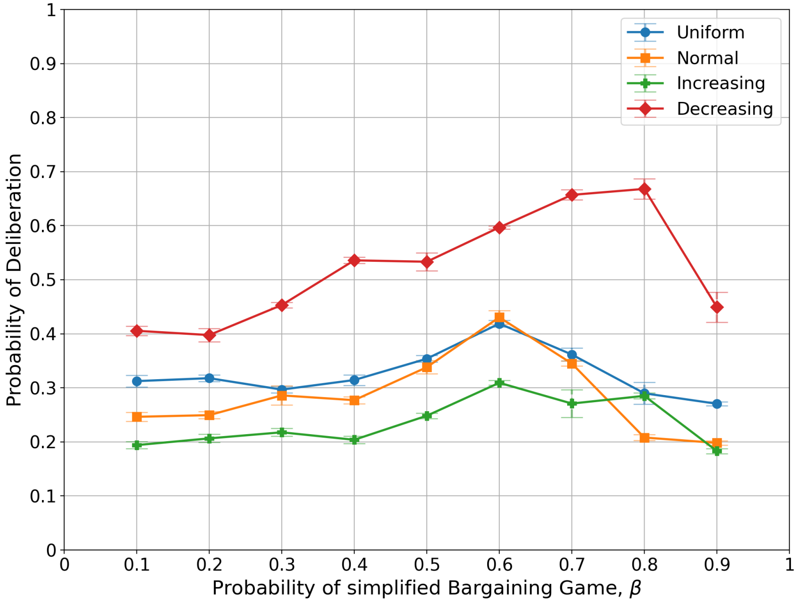


Figure G.3: Proposers' deliberation patterns. Average proposers' probability to deliberate and their standard errors as a function of β by cost distribution.

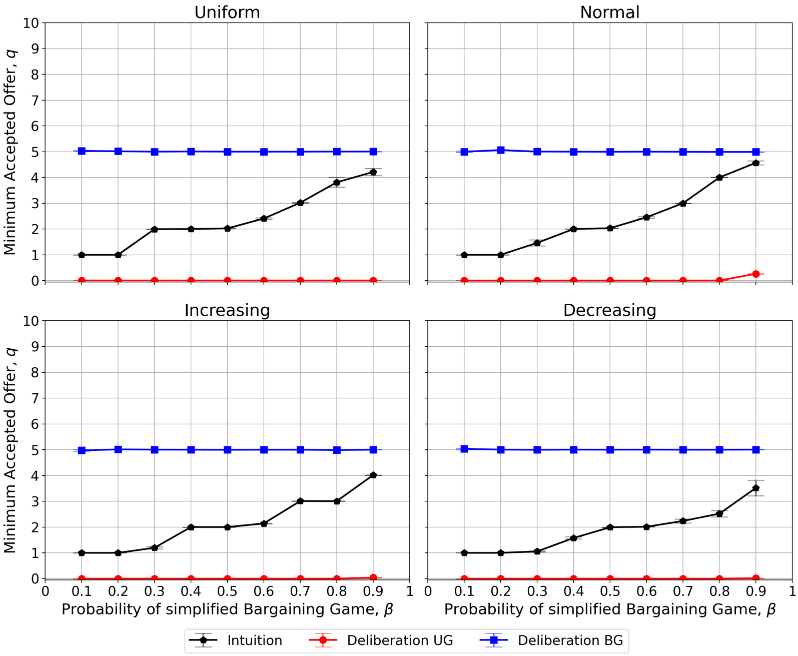


Figure G.4: Receivers' behavior. Receivers' average minimum accepted offer under intuition, under deliberation if the game is an UG, and under deliberation if the game is a simplified BG and their standard errors as a function of β by assumed cost distribution.

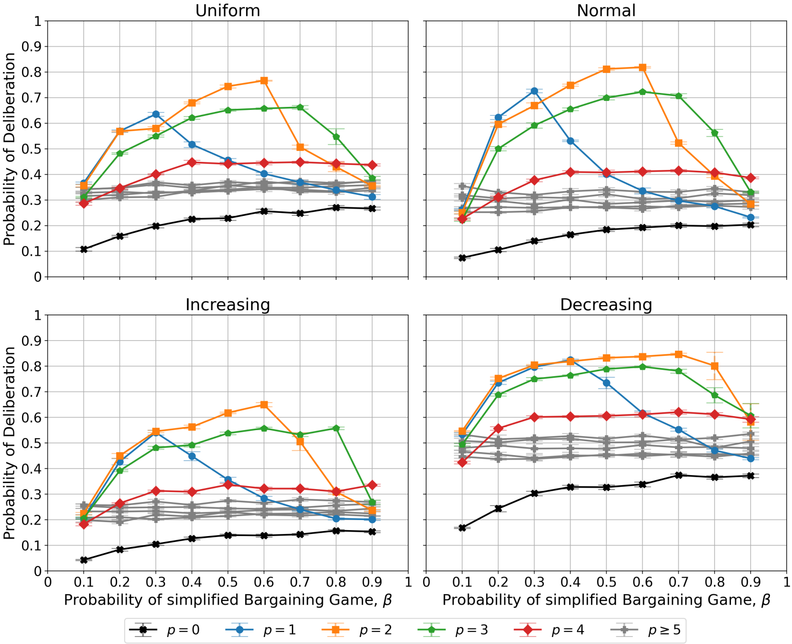


Figure G.5: Receivers' deliberation patterns. Average receivers' probability to deliberate given that the proposer has offered an amount p and their standard errors as a function of β by cost distribution.

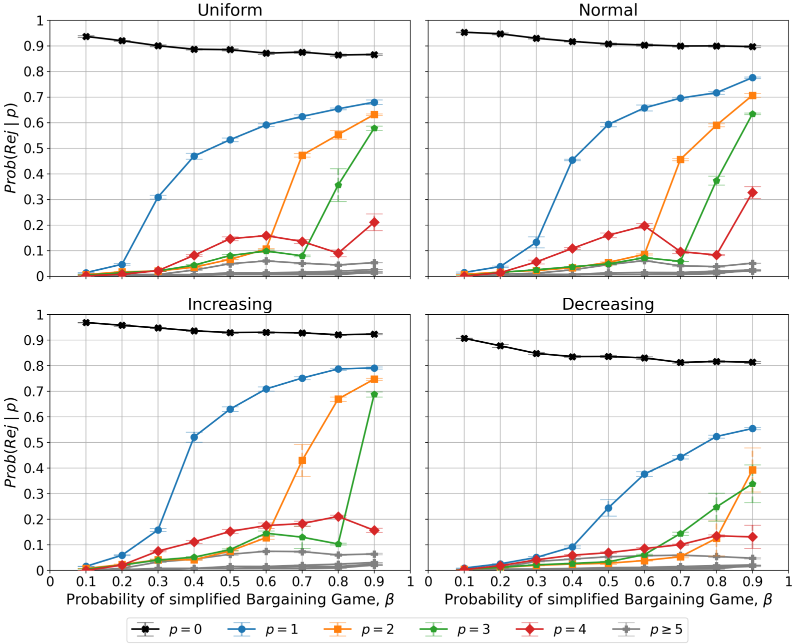


Figure G.6: Conditional rejection rates. Average receivers' conditional rejection rates in the Ultimatum Game and their standard errors as a function of β by cost distribution.

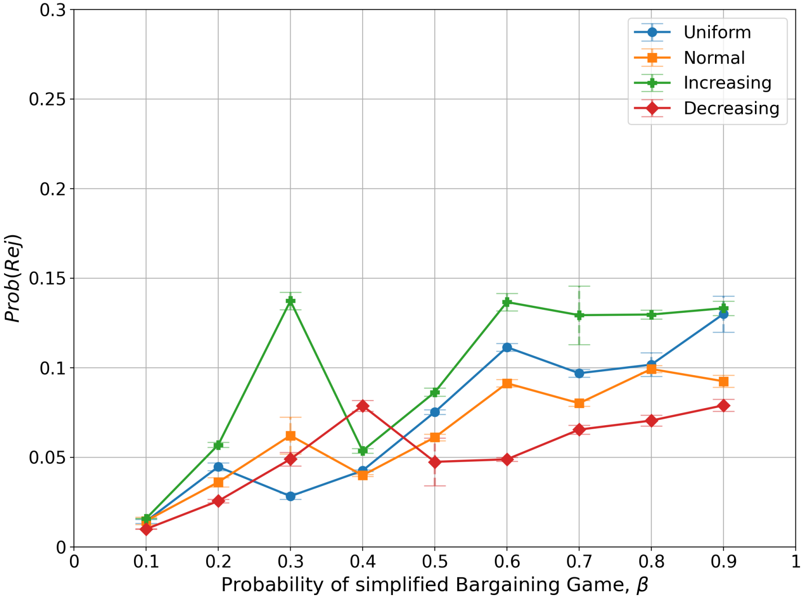


Figure G.7: Expected rejection rates. Overall expected rejection rates in the Ultimatum Game and their standard errors as a function of β by cost distribution.

Appendix H

Chapter 4: Different Patience Factors

I repeat the analyses reported in Chapter 4 by considering different patience factors. More precisely, in the main text I considered the case in which both proposers and receivers had a patience factor of $\delta = 0.8$.

In the model, the patience factor determines the agents' payoff in the simplified Bargaining Game in case the receiver rejects the offer made by the proposer. In other words, it affects the receivers' outside option in the simplified BG and it also indirectly affects the receivers' outside option under intuition. Consequently, from a theoretical point of view, I expect that if agents are less patient then receivers should be less demanding and should tend to accept more often unfair offers.

The next figures report the simulation results obtained by considering patience factors $\delta \in \{0.2, 0.5, 0.8\}$. These findings substantially confirm the theoretical predictions just made.

One thing that is worth mentioning is the potential role played by cognitive manipulations on agents' level of patience. In particular, cognitive manipulations may affect agents' patience in such a way that the change in patience counteracts the expected effects of increased/decreased reliance on intuition. Consider, for example, the case of a time pressure treatment that increases subjects' reliance on intuition. According to the results obtained in Chapter 4 this should increase receivers' MAOs and, consequently, make them more demanding. However, if time pressure also makes receivers more impatient and, consequently, decreases their MAO, then – depending on which effect dominates – receivers might both become more demanding, less demanding, or not change their behavior (if the two effects compensate each other). Further research on this topic might be of interest.

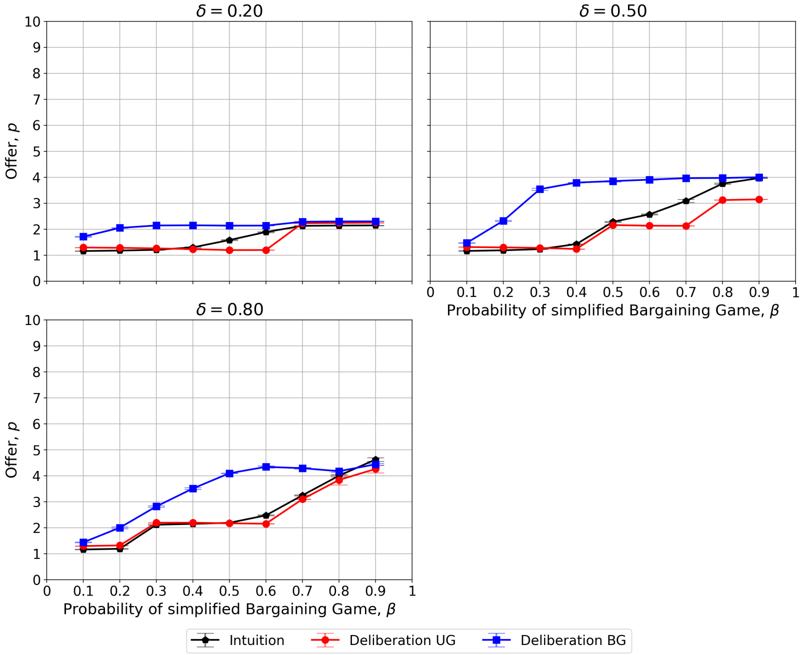


Figure H.1: Proposers' behavior. Proposers' average offer under intuition, under deliberation if the game is an UG, and under deliberation if the game is a simplified BG and their standard errors as a function of β by cost distribution.

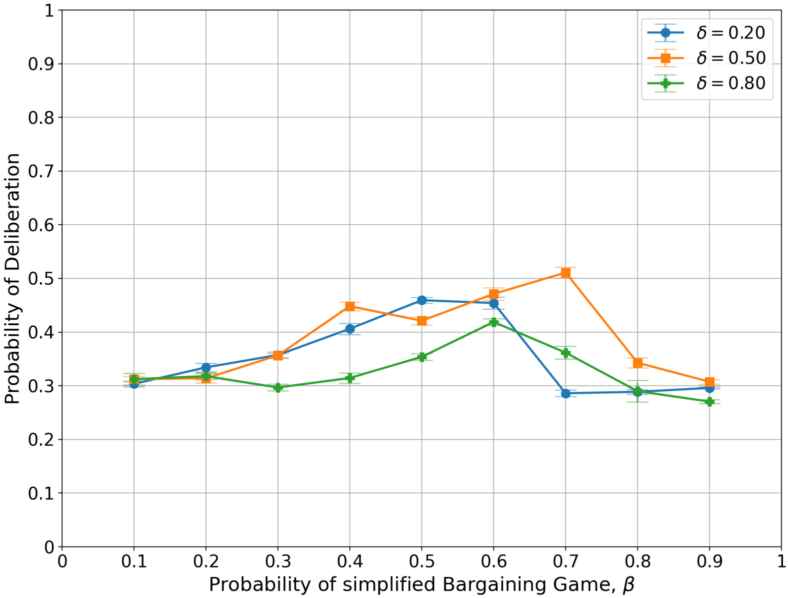


Figure H.2: Proposers' deliberation patterns. Average proposers' probability to deliberate and their standard errors as a function of β by cost distribution.

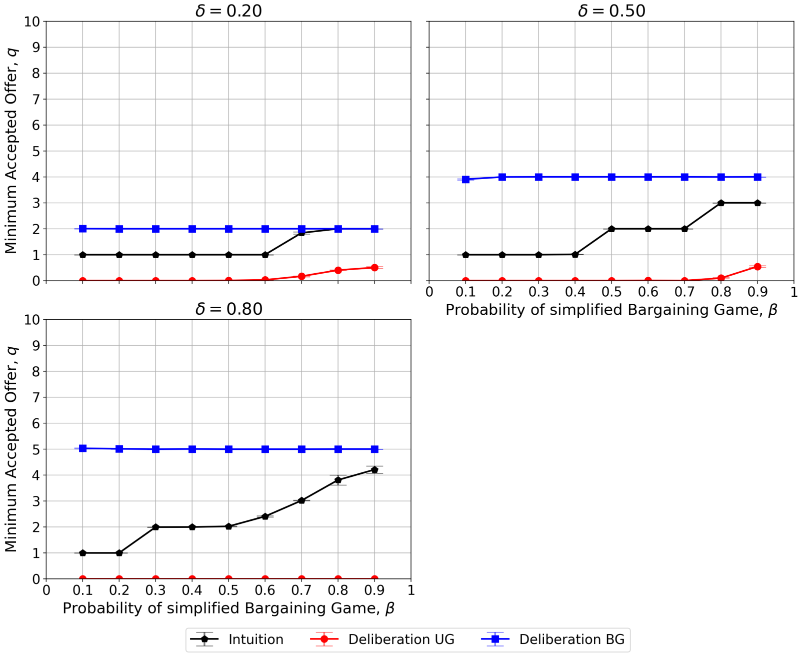


Figure H.3: Receivers' behavior. Receivers' average minimum accepted offer under intuition, under deliberation if the game is an UG, and under deliberation if the game is a simplified BG and their standard errors as a function of β by assumed cost distribution.

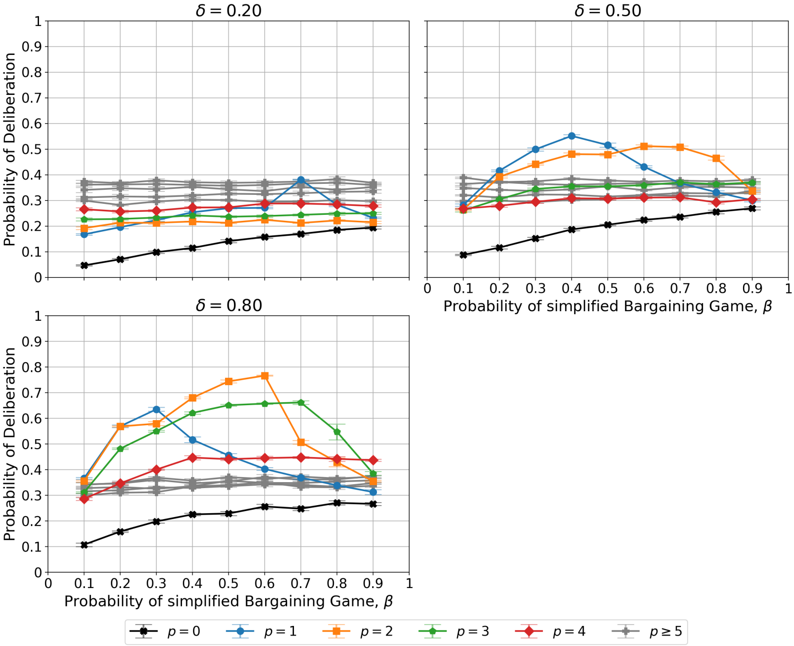


Figure H.4: Receivers' deliberation patterns. Average receivers' probability to deliberate given that the proposer has offered an amount p and their standard errors as a function of β by cost distribution.

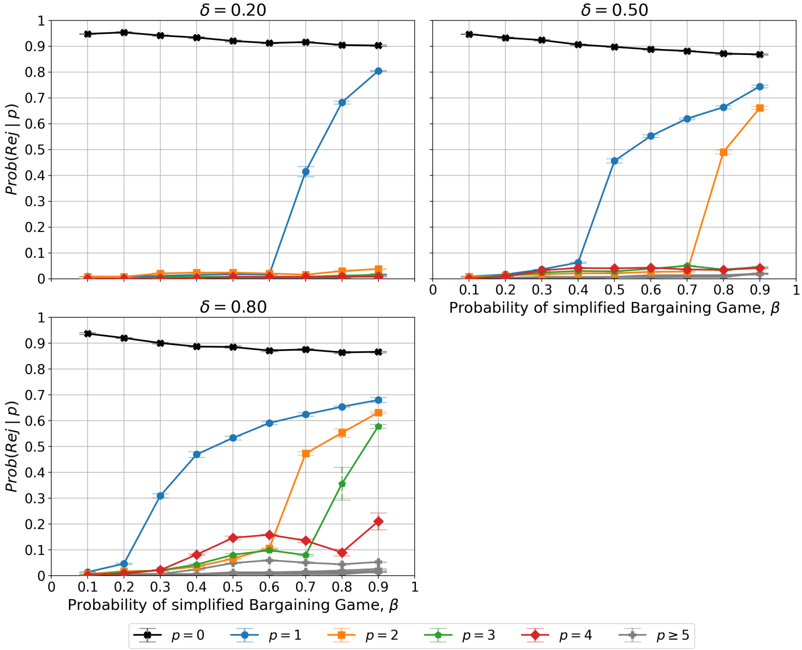


Figure H.5: Conditional rejection rates. Average receivers' conditional rejection rates in the Ultimatum Game and their standard errors as a function of β by cost distribution.

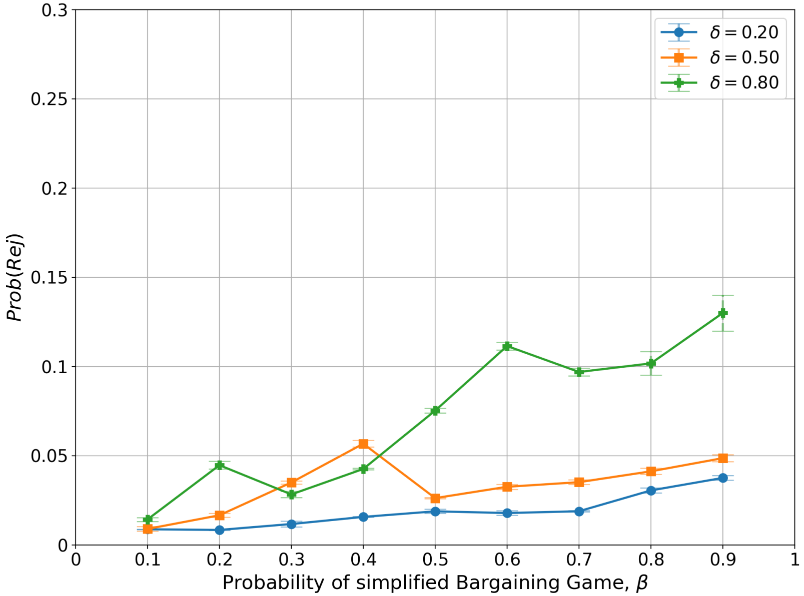


Figure H.6: Expected rejection rates. Overall expected rejection rates in the Ultimatum Game and their standard errors as a function of β by cost distribution. These rejection rates are computed by taking into account both proposers' behavior and receivers' conditional rejection rates.

Bibliography

Achtziger, Anja, Carlos Alós-Ferrer, and Alexander K Wagner (2016). "The impact of self-control depletion on social preferences in the ultimatum game". In: *Journal of Economic Psychology* 53, pp. 1–16.

Alexander, J McKenzie (2007). "The structural evolution of morality". In.

Alós-Ferrer, Carlos and Simon Weidenholzer (2006). "Imitation, local interactions, and efficiency". In: *Economics Letters* 93.2, pp. 163–168.

— (2008). "Contagion and efficiency". In: *Journal of Economic Theory* 143.1, pp. 251–274.

André, Jean-Baptiste and Nicolas Baumard (2011). "Social opportunities and the evolution of fairness". In: *Journal of theoretical biology* 289, pp. 128–135.

Anwar, Ahmed W (2002). "On the Co-existence of Conventions". In: *Journal of Economic Theory* 107.1, pp. 145–155.

Axelrod, Robert and William D Hamilton (1981). "The evolution of cooperation". In: *Science* 211.4489, pp. 1390–1396.

Bear, Adam, Ari Kagan, and David G Rand (2017). "Co-evolution of cooperation and cognition: the impact of imperfect deliberation and context-sensitive intuition". In: *Proceedings of the Royal Society B* 284.1851, p. 20162326.

Bear, Adam and David G Rand (2016). "Intuition, deliberation, and the evolution of cooperation". In: *Proceedings of the National Academy of Sciences* 113.4, pp. 936–941.

Belloc, Marianna et al. (2019). "Intuition and Deliberation in the Stag Hunt Game". In: *Scientific Reports* 9.1, pp. 1–7.

Bhaskar, Venkataraman and Fernando Vega-Redondo (2004). "Migration and the evolution of conventions". In: *Journal of Economic Behavior & Organization* 55.3, pp. 397–418.

Bilancini, Ennio and Leonardo Boncinelli (2018). "Social coordination with locally observable types". In: *Economic Theory* 65.4, pp. 975–1009.

Bilancini, Ennio, Leonardo Boncinelli, and Roberto Di Paolo (2021). "Manipulating cognition in the one-shot Stag-Hunt game played online". In: *Mimeo*.

Bilancini, Ennio, Leonardo Boncinelli, and Eugenio Vicario (2022). "Assortativity in cognition". In: *arXiv preprint arXiv:2205.15114*.

Cappelletti, Dominique, Werner Güth, and Matteo Ploner (2011). "Being of two minds: Ultimatum offers under cognitive constraints". In: *Journal of Economic Psychology* 32.6, pp. 940–950.

Capraro, Valerio (2019). "The dual-process approach to human sociality: A review". In: *Available at SSRN 3409146*.

Chen, Hsiao-Chi, Yunshyong Chow, and Li-Chau Wu (2013). "Imitation, local interaction, and coordination". In: *International Journal of Game Theory* 42.4, pp. 1041–1057.

Clare Anderson, David L Dickinson (2010). "Bargaining and trust: the effects of 36-h total sleep deprivation on socially interactive decisions". In: *Journal of Sleep Research* 19, pp. 54–63. DOI: 10.1111/j.1365-2869.2009.00767.x.

Cui, Zhiwei (2014). "More neighbors, more efficiency". In: *Journal of Economic Dynamics and Control* 40, pp. 103–115.

Cui, Zhiwei and Simon Weidenholzer (2021). "Lock-in through passive connections". In: *Journal of Economic Theory* 192, p. 105187.

Debove, Stéphane, Nicolas Baumard, and Jean-Baptiste André (2016). "Models of the evolution of fairness in the ultimatum game: a review and classification". In: *Evolution and Human Behavior* 37.3, pp. 245–254. ISSN: 1090-5138. DOI: https://doi.org/10.1016/j.evolhumbehav.2016.01.001. URL: https://www.sciencedirect.com/science/article/pii/S1090513816000039.

Dieckmann, Tone (1999). "The evolution of conventions with mobile players". In: *Journal of Economic Behavior & Organization* 38.1, pp. 93–111.

Ellison, Glenn (1993). "Learning, local interaction, and coordination". In: *Econometrica*, pp. 1047–1071.

— (2000). "Basins of attraction, long-run stochastic stability, and the speed of step-by-step evolution". In: *Review of Economic Studies* 67.1, pp. 17–45.

Ely, Jeffrey C (2002). "Local conventions". In: *Advances in Theoretical Economics* 2.1.

Erev, Ido and Alvin E Roth (1998). "Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria". In: *American economic review*, pp. 848–881.

Evans, Jonathan St BT and Keith E Stanovich (2013). "Dual-process theories of higher cognition: Advancing the debate". In: *Perspectives on Psychological Science* 8.3, pp. 223–241.

Foster, Dean and Peyton Young (1990). "Stochastic evolutionary game dynamics". In: *Theoretical Population Biology* 38.2, pp. 219–232.

Freidlin, Mark Iosifovich and Alexander D Wentzell (1984). *Random Perturbations of Dynamical Systems*. Springer.

Gigerenzer, Gerd and Wolfgang Gaissmaier (2011). "Heuristic decision making". In: *Annual review of psychology* 62.1, pp. 451–482.

Goyal, Sanjeev and Fernando Vega-Redondo (2005). "Network formation and social coordination". In: *Games and Economic Behavior* 50.2, pp. 178–207.

Grimm, Veronika and Friederike Mengel (2011). "Let me sleep on it: Delay reduces rejection rates in ultimatum games". In: *Economics Letters* 111.2, pp. 113–115. URL: https://EconPapers.repec.org/RePEc:eee:ecolet:v:111:y:2011:i:2:p:113-115.

Güth, Werner and Martin G Kocher (2014). "More than thirty years of ultimatum bargaining experiments: Motives, variations, and a survey of the recent literature". In: *Journal of Economic Behavior & Organization* 108, pp. 396–409.

Güth, Werner, Rolf Schmittberger, and Bernd Schwarze (1982). "An experimental analysis of ultimatum bargaining". In: *Journal of Economic Behavior Organization* 3.4, pp. 367–388. ISSN: 0167-2681. DOI: https://doi.org/10.1016/0167-2681(82)90011-7. URL: https://www.sciencedirect.com/science/article/pii/0167268182900117.

H. P. Young (1993). "The Evolution of Conventions". In: *Econometrica* 61, pp. 57–84.

Halali, Eliran, Yoella Bereby-Meyer, and Axel Ockenfels (2013). "Is it all about the self? The effect of self-control depletion on ultimatum game proposers". In: *Frontiers in Human Neuroscience* 7.240.

Jackson, Matthew O and Alison Watts (2002). "On the formation of interaction networks in social coordination games". In: *Games and Economic Behavior* 41.2, pp. 265–291.

Jagau, Stephan and Matthijs van Veelen (2017). "A general evolutionary framework for the role of intuition and deliberation in cooperation". In: *Nature Human Behaviour* 1.8, pp. 1–6.

Kandori, Michihiro, George J Mailath, and Rafael Rob (1993). "Learning, mutation, and long run equilibria in games". In: *Econometrica*, pp. 29–56.

Kandori, Michihiro and Rafael Rob (1995). "Evolution of equilibria in the long run: A general theory and applications". In: *Journal of Economic Theory* 65.2, pp. 383–414.

Knoch, Daria et al. (2006). "Disruption of Right Prefrontal Cortex by Low-Frequency Repetitive Transcranial Magnetic Stimulation Induces Risk-Taking Behavior". In: *Journal of Neuroscience* 26.24, pp. 6469–6472. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.0804-06.2006. eprint: https://www.jneurosci.org/content/26/24/6469.full.pdf. URL: https://www.jneurosci.org/content/26/24/6469.

Mosleh, Mohsen and David G Rand (2018). "Population structure promotes the evolution of intuitive cooperation and inhibits deliberation". In: *Scientific reports* 8.1, pp. 1–8.

Neo, Wei Siong et al. (2013). "The effects of time delay in reciprocity games". In: *Journal of Economic Psychology* 34, pp. 20–35. ISSN: 0167-4870. DOI: https://doi.org/10.1016/j.joep.2012.11.001. URL: https://www.sciencedirect.com/science/article/pii/S0167487012001389.

Newton, Jonathan (2018). "Evolutionary game theory: A renaissance". In: *Games* 9.2, p. 31.

Nowak, Martin A, Karen M Page, and Karl Sigmund (2000). "Fairness versus reason in the ultimatum game". In: *Science* 289.5485, pp. 1773–1775.

Oechssler, Jörg (1997). "Decentralization and the coordination problem". In: *Journal of Economic Behavior & Organization* 32.1, pp. 119–135.

Page, Karen M, Martin A Nowak, and Karl Sigmund (2000). "The spatial ultimatum game". In: *Proceedings of the Royal Society of London. Series B: Biological Sciences* 267.1458, pp. 2177–2182.

Peski, Marcin (2010). "Generalized risk-dominance and asymmetric dynamics". In: *Journal of Economic Theory* 145.1, pp. 216–248.

Pin, Paolo, Elke Weidenholzer, and Simon Weidenholzer (2017). "Constrained mobility and the evolution of efficient outcomes". In: *Journal of Economic Dynamics and Control* 82, pp. 165–175.

Rand, David G. and Martin A. Nowak (2013). "Human Cooperation". In: *Trends in Cognitive Sciences* 17.8, pp. 413–425.

Rand, David G., Corina E. Tarnita, et al. (2013). "Evolution of fairness in the one-shot anonymous Ultimatum Game". In: *Proceedings of the National Academy of Sciences* 110.7, pp. 2581–2586. DOI: 10.1073/pnas.1214167110. eprint: https://www.pnas.org/doi/pdf/10.1073/pnas.

1214167110. URL: https://www.pnas.org/doi/abs/10.1073/pnas.1214167110.

Robson, Arthur J and Fernando Vega-Redondo (1996). "Efficient equilibrium selection in evolutionary games with random matching". In: *Journal of Economic Theory* 70.1, pp. 65–92.

Rubinstein, Ariel (1982). "Perfect Equilibrium in a Bargaining Model". In: *Econometrica* 50.1, pp. 97–109. ISSN: 00129682, 14680262. URL: http://www.jstor.org/stable/1912531 (visited on 10/20/2022).

Santos, Francisco C., Jorge M. Pacheco, and Tom Lenaerts (2006). "Cooperation prevails when individuals adjust their social ties". In: *PLoS Computational Biology* 2.10.

Shi, Fei (2013). "Comment on "On the co-existence of conventions"[J. Econ. Theory 107 (2002) 145–155]". In: *Journal of Economic Theory* 1.148, pp. 418–421.

Skyrms, Brian (2004). *The stag hunt and the evolution of social structure*. Cambridge University Press.

Staudigl, Mathias and Simon Weidenholzer (2014). "Constrained interactions and social coordination". In: *Journal of Economic Theory* 152, pp. 41–63.

Sutter, Matthias, Martin Kocher, and Sabine Strauß (2003). "Bargaining under time pressure in an experimental ultimatum game". In: URL: https://EconPapers.repec.org/RePEc:lmu:muenar:18220.

Weidenholzer, Simon (2010). "Coordination games and local interactions: a survey of the game theoretic literature". In: *Games* 1.4, pp. 551–585.

Zaki, Jamil and Jason P Mitchell (2013). "Intuitive prosociality". In: *Current Directions in Psychological Science* 22.6, pp. 466–470.