Inauguraldissertation zur Erlangung des akademischen Grades eines Doktors der Wirtschaftswissenschaften der Universität Mannheim

Titel:

Essays on Learning in Games and Social Contexts

Autor:

Alexander Gaigl

Abteilungssprecher:

Prof. Tom Krebs, Ph.D.

Prüfungsausschuss:

Prof. Dr. Paul Gans *(Vorsitzender)*

Prof. Jörg Oechssler, PhD (*Erstreferent)*

Prof. Dr. Ernst-Ludwig von Thadden (*Zweitreferent)*

Prof. Dr. Erik Theissen

Datum der mündlichen Prüfung: 12.05.2010

Kurzlebenslauf des Autors

2006-20010	Universität Mannheim, VWL, Promotion
2001-2006	Universität Karlsruhe, Wirtschaftsingenieurwesen, Diplom
2000-2001	Zivildienst, Leonberg
1991-2000	Albert-Schweitzer Gymnasium, Leonberg
1987-1991	Mörike Schule, Leonberg

Erklärung

Hiermit bestätige ich, dass ich die vorliegende Dissertation mit dem Titel:

Essays on Learning in Games and Social Contexts

unter Verwendung der unten genannten Hilfsmittel selbstständig erstellt habe. Ich versichere, dass ich ausschließlich die angegebenen Quellen und Hilfen in Anspruch genommen habe. Der erste Teil der Dissertation ist das Resultat gemeinschaftlicher Arbeit, und er ist deutlich als solcher gekennzeichnet.

Verwendete Hilfsmittel:

- MATLAB, LaTex, Excel
- Hochleistungscomputer des Rechenzentrums der Universität Mannheim
- referenzierte Literautur

Alexander Gaig

Introduction

Models of learning and adaptive behavior have received much attention in the Economic literature in recent years. Different classes of learning approaches have been proposed, which vary in a variety of dimensions, such as the requirements on individuals' rationality, the degree of interaction, the number of interacting individuals as well as the time-horizon. In this dissertation, I focus on two classes of learning models, which are particularly important.

The first two papers study the long-run outcome of learning in situations in which a large number of players interact repeatedly. A central assumption here is that players do not carry out complex strategic considerations but rather follow relatively simple decision rules which are based on past experience. It is often argued that such approaches do a better job in capturing the behavior of real-world decision makers than models which assume ideally rational behavior. In the paper *On the Dynamics in a Market for long-term Relationships*, I study a model in which players are faced with a modified version of the Prisoner's Dilemma which gives them the choice between either maintaining or quitting the relationship with their current partner. I show that limited information about the environment combined with simple learning behavior may induce players to adopt cooperation in the long run. Moreover, in the paper *Learning under Incomplete Information with Applications to Auctions*, I show that bidders may adopt highly rational bidding strategies over time, even if they employ simple adaptive learning rules. This result allows for a reinterpretation of classical equilibrium results in the following sense. While such approaches view the equilibrium of an auction as the result of careful reasoning among highly rational bidders, I argue that a bidding equilibrium can also be viewed as the long-run outcome of interaction between less sophisticated bidders.

There is another important class of learning models, which is generally referred to as the class of *herding models*. The main objective of this literature is to explain phenomena of uniform social behavior, such as fads, fashions or stock market bubbles. This literature studies situation in which rational individuals observe the actions of others and draw conclusions about the private information underlying these actions. The bulk of the literature assumes that payoffs only depend upon the information of others, but not on their actions. Since this seems to be at odds with many real-world situations, the paper *Observational Learning & Strategic Externalities* analyzes a herding model in which there is a direct dependence between individuals' payoffs and the actions of others. I study how this assumption affects the incentives to imitate others and under which conditions uniform behavior occurs.

Contents

Heiner Schumacher†
Alexander Gaigl††

On the Dynamics in a Market for long-term Relationships*

University of Mannheim
December 2009

Abstract

The prisoner's dilemma is played by many pairs simultaneously in a random matching game in which players have the option to maintain or to quit relationships. Hence, the population consists of fixed matches and a "market for long-term relationships". Agents learn their current opponent's strategy in finitely many periods. Further, they update their subjective belief over the aggregate behavior of players in the market. As learning rule we use fictitious play. Analytically, and by simulating the model, we derive conditions under which a significant degree of cooperation can be expected in an infinite population. Then, we extend the model to finite populations and show that the dynamics are similar to the infinite case if there are sufficiently many agents.

JEL Classification: C70, C72, C78

Keywords: Learning, Prisoner's Dilemma, Random Matching, Fictitious Play, Long-term Relationship

---

†Corresponding author, Contacts: Goethe-Universität Frankfurt, Professur fr BWL, insbes. Organisation und Management, Postfach 53, Grüneburgplatz 1, D-60629 Frankfurt, H.Schumacher[at]em.uni-frankfurt.de.

††Contacts: Department of Economics, University of Mannheim, D-68131 Mannheim, Germany, agaigl[at]rumms.uni-mannheim.de, http://www.vwl.uni-mannheim.de/gk/_gaigl.

*We thank Ernst-Ludwig von Thadden, Jörg Oechssler, Jürgen Eichberger, Karl Schlag, Frank Rosar, the session participants of the CDSE Seminar and the doctoral workshops at the University of Heidelberg and the University of Osaka for helpful suggestions. All remaining errors are our own.

1 Introduction

Both the theoretical and experimental literature on cooperation in the prisoner's dilemma (PD) mostly concentrated on settings in which agents play the PD repeatedly against exogenously determined opponents[1]. For many cases in real-life situations, such as labor markets or business relations, this is unnatural: One usually has the option to maintain or to quit the relationship with a certain person. However, a framework in which agents have this option, has rarely been analyzed. For a summary, see Mailath and Samuelson (2006), chapter 5.2.

In this paper, we therefore consider the following setup: After observing the opponent's action choice in the stage game (the PD or a variation of the PD), each agent of an infinite population has the option to maintain or to quit the relationship with her current opponent. If and only if both agents choose the first option, they play against each other in the next period with positive probability. Otherwise they return to a "market for long-term relationships" and are matched randomly to other players in this market. The matching process in the market is global and non-assortative. Furthermore, there are no information flows between pairs.

The most important contribution for this game was made by Gosh and Ray (1996): There are two types of agents in the population—myopic and patient ones. Patient players "test" their opponent in the first periods of a new relationship by increasing slowly the degree of cooperation. While myopic players defect after few periods and the relationship is broken up subsequently, patient agents maintain the relationship and continue to cooperate. With this strategy of "starting small", any gain from defection is wiped out by the subsequent restart of a phase of low payoffs in the new relationship. However, the resulting equilibrium relies on two assumptions:

(1.) Players know the aggregate play of agents in the market.

(2.) There are fixed shares of myopic and patient individuals.

In a paper by Datta (1993), the second assumption is suppressed. The strategies are the same, i.e., in the first periods of a new relationship agents choose

[1]For the theory on infinitely repeated games with fixed opponents, see Mailath and Samuelson (2006) as reference, experiments were conducted by Roth and Murnighan (1978), Murnighan and Roth (1983), Aoyagi and Fréchette (2003), Dal Bó and Fréchette (2007) or Duffy and Ochs (2003). Cooperation in random matching games was analyzed theoretically by Ellison (1994) and Kandori (1992). Experimental evidence on cooperation in random matching games can be found in Duffy and Ochs (2003), and in one-shot games for example in Brosig (2001).

not to cooperate (or only a very small degree of cooperation) and start to cooperate in later periods. With this strategy, cooperation can be established in an homogenous population. Gosh and Ray (1996) note that the considered strategy then does not fulfill the criterion of "bilateral rationality": Given that all other players in the population stick to the described path of play, it is optimal for two agents who meet for the first time in the market, to quit the punishment in the first periods and to start the relationship with cooperation immediately. This would not violate any incentive constraint. However, if all pairs act in this way, we are no longer in equilibrium. Thus, there is no cooperative equilibrium in the above game in a homogenous population in pure strategies which do not violate the refinement of "bilateral rationality". Conventional approaches therefore do not provide a convincing solution.

We solve this dilemma by dropping also the first assumption: Players do not know the aggregate play of agents in the market. They only know that some agents play according to a "cooperative strategy" (for example, a "starting small" strategy as in the papers cited above), which entails a long-term relationship, and some agents play according to a "non-cooperative strategy", which is to defect in each period. Thus, after finite time each agent knows the play of her current opponent. Players choose to maintain the relationship with their current opponent if and only if her play is consistent with the cooperative strategy. Furthermore, they have a subjective belief $\tilde{\mu}$ about the share of agents in the market μ who play according to the cooperative strategy. Agents update this belief based on past experiences. For a given subjective belief, they choose the strategy which maximizes the sum of discounted expected payoffs.

The cooperative strategies we consider start with T^* periods of defection, i.e., we allow for "starting small" strategies. However, we are mainly interested to what extent cooperation can prevail in the population if long-term relationships start with cooperation immediately which means that $T^* = 0$. Note that such a strategy would never support a symmetric Nash equilibrium in pure strategies with cooperation. We will see that cooperation in the population may be a stable outcome even if $T^* = 0$. Therefore, we do not only drop an unrealistic assumption, we also maintain a more plausible solution, as the outcome of a cooperative relationship can not be improved by players, and therefore agents' behavior is "bilateral rational".

As updating rule we take "fictitious play", initially introduced by Brown (1951) as a means of calculating Nash-equilibria and extensively studied thereafter. Under fictitious play, each player assumes that her opponents are playing according to a stationary distribution. In each round, every individual plays a best response to the empirical frequency of her opponent's play—see for example Fudenberg and Levine (1998). In particular, players do not try to influence the future play of their opponents. This assumption may be problematic in small populations, but is very reasonable in large population frameworks, where players will not meet again after a relationship was broken up.

Due to the nature of the game, there is one important difference to the standard fictitious play model: Agents count one opponent's strategy choice as one observation (and not one action choice). Two types of fictitious play will be considered: In the first one, agents calculate the average behavior using all their observations. In the second one, agent's memory is limited to the last n observations. Throughout the paper we assume that agents are symmetric with respect to the updating rule.

The optimal strategy depends on the subjective belief. Assume, for example, that the cooperative strategy prescribes to cooperate in each period. Then we can observe the following:

- If an agent has met mainly players who followed the non-cooperative strategy, then she assumes that the probability of meeting an agent in the market who plays according to the cooperative strategy, is very small and thus chooses the non-cooperative strategy in order to avoid exploitation.

- If an agent has met mainly players who followed the cooperative strategy, then she assumes that the probability of meeting an agent in the market who plays according to the cooperative strategy, is very high and thus chooses the non-cooperative strategy in order to exploit future opponents.

- If an agent has made mixed experiences, then she chooses the cooperative strategy in order to establish a cooperative long-term relationship which is expected to be more beneficial than staying in the market.

The goal of this paper is to analyze the aggregate dynamics in the population, i.e., the evolution of the share of agents who choose the cooperative strategy. We proof that with unlimited memory, beliefs converge to a single value whenever aggregate play in the market converges. This value is consistent with a Nash equilibrium of the game.

Under limited memory, beliefs may remain heterogenous even when aggregate play converges. We show that the state in which all agents defect, can be asymptotically stable—depending on the discount factor and the exogenous rate of breakup. Furthermore, if the cooperative strategy involves sufficiently many periods of non-cooperation at the beginning of a relationship, the state in which all agents play cooperatively, can be asymptotically stable. This result then replicates Datta (1993) without the assumption of knowledge.

In general, the dynamics under both updating rules can not be determined explicitly and we are not aware of any method that allows to proof convergence in our setting. We therefore simulate the model and focus on the cooperative strategy which prescribes to start with cooperation immediately, i.e. $T^* = 0$. We observe that aggregate behavior always converges under both specifications of the updating rule. For a large set of parameter specifications and distributions of initial beliefs, cooperation is a stable outcome. With limited memory, aggregate behavior is in general inconsistent with the Nash equilibrium of the game. Therefore, the results of the present approach may differ substantially from the ones obtained in settings with the assumption of knowledge of aggregate play.

In a further step, we consider the model for large but finite populations, such that the outcome becomes stochastic. The main result is that the dynamics resemble the ones obtained in an infinite population if the population is sufficiently large. However, we also show that for some parameter values, cooperation breaks down in a finite population, while it would be a stable outcome in the infinite case under the same specification.

The rest of the paper is organized as follows: The next section outlines the model. In chapter 3, we present the updating rules and derive analytical results regarding the stability and degree of cooperation. In chapter 4, we summarize the results of the simulation. Most of our intuition for the dynamics of the model will follow from this section. Chapter 5 extends the model to finite populations. Readers who are not interested in determinstic approximation may wish to skip this section. Chapter 6 concludes. All proofs and figures can be found in the appendix.

2 The Basic Model

We consider an infinitely repeated two-player PD which is played simultaneously by a continuum of agents. Time is discrete and denoted by $t \in \{1, 2, ...\}$. Every agent plays the PD in each period with some opponent: an agent has the options "cooperate" (C) and "do not cooperate" (D). Payoffs are shown in the following matrix (where player 1 chooses rows and player 2 chooses columns):

	D	C
D	$1, 1$	$H, 0$
C	$0, H$	G, G

We fix $G, H \in \mathbb{R}$ with $1 < G < H < 2G$, such that the sum of payoffs is maximal at the profile (C, C). After observing the opponent's action choice, each agent has to choose whether to maintain (M) or to quit (Q) the current relationship. If and only if both partners choose action M they play the game together again in the next period with probability $1 - \sigma$. The parameter σ is the exogenous rate of breakup. If and only if an agent plays the PD in period t with the opponent of the previous period, we call the link between those two agents a long-term relationship. Agents who are not in a long-term relationship in t, will be paired up randomly at the beginning of period t. The pool of agents who are not in a long-term relationship at the beginning of a period, will be called the 'market for long-term relationships'.

2.1 Evaluation of the Current Opponent and Strategies

Let $c(t)$ be a counting function. If $c(t) = i$, then the agent plays the stage game in period t against her i'th opponent. Accordingly, we have $c(0) = 1$. Further, denote by $T^i \in \mathbb{N}$ the number of periods in which the agent played the PD with opponent number i until the current period. Let h_{T^i} be the history of actions of opponent i in the T^i periods in which an agent played the PD with this opponent, i.e., each element in h_{T^i} is either D or C. Let $T, T^* \in \mathbb{N}$ and define h^c_T as a history of actions where

- all elements are equal to D if $T \leq T^*$ and
- the first T^* elements are equal to D and the remaining ones are equal to C if $T > T^*$.

Then define the evaluation function as

$$g(h_{T^i}) = \begin{cases} 1 & \text{if } T^i > T^* \text{ and } h_{T^i} = h^c_{T^i} \\ 0 & \text{if } T^i \leq T^* \text{ and } h_{T^i} = h^c_{T^i} \\ -1 & \text{otherwise} \end{cases} .$$

Whenever an agent is in the market, she chooses between two strategies: a cooperative one, $f^c_{T^*}$, and a non-cooperative one, $f^d_{T^*}$. The strategies are specified as follows:

$f^c_{T^*}$: Choose D if $T^i \leq T^*$. Choose C if $T^i > T^*$. As long as the value of g for your current opponent is equal to 0 or 1, choose M, otherwise Q.

$f^d_{T^*}$: Choose D in each period. As long as the value of g for your current opponent is equal to 0 or 1, choose M, otherwise Q.

The interpretation of the evaluation function g is then as follows: As long as for a given opponent the value of g is equal to 0, an agent does not know her strategy. If it is equal to 1, she knows that her opponent plays according to the cooperative strategy, if it is equal to -1, she knows that her opponent plays according to the non-cooperative strategy or any other strategy. If $T^* > 0$, players start a long-term relationship with non-cooperation and switch to cooperation after T^* periods. This sort of strategy decreases the sum of expected discounted payoffs of agents who play according to the non-cooperative strategy relative to cooperative agents.

2.2 Evaluation of the Population

Define by $\mu(t)$ the share of agents in the market in period t who play according to the cooperative strategy. Each player has a subjective belief $\tilde{\mu}(t)$ over $\mu(t)$. Let

$$hs^{c(t)} = \{g(h_{T^{-n+1}}), ..., g(h_{T^{c(t)}})\}$$

be an agent's history of evaluations in period t, where the vector

$$\{g(h_{T^{-n+1}}), ..., g(h_{T^0})\}, \tag{1}$$

for $n \in \mathbb{N}$, and $g(h_{T^i}) \in \{-1, 1\}$ for all $i \in \{-n+1, ..., 0\}$, determines the agent's subjective belief in the first period (the 'preplay-observations'). Denote by

$$HS^c(n) = \big\{\{g_i\}_{i \in \{1,...,n+c\}} \mid g_i \in \{-1, 0, 1\}\big\}$$

the set of all histories of evaluations of length $n + c$ and by

$$HS(n) = \bigcup_{c>0} HS^c(n)$$

the set of all finite histories of evaluations with length of at least $n + 1$. For given n, the belief in a period t then is given by an updating rule

$$\tilde{\mu} : HS(n) \to [0, 1].$$

Let $\tilde{\mu}(t)$ be the abbreviation for $\tilde{\mu}(hs^{c(t)})$.

2.3 Strategy Choice and Sequence of Events

Agents discount future gains with δ and maximize over the sum of discounted expected utility. Each agent in the market chooses the strategy which yields her the highest sum of discounted expected utility for given subjective belief $\tilde{\mu}(t)$. Denote by $E[f^d_{T^*}, \tilde{\mu}(t)]$ the sum of discounted expected utility if in period

t the non-cooperative strategy $f^d_{T^*}$ is chosen, and by $E[f^c_{T^*}, \tilde{\mu}(t)]$ the sum of discounted expected utility if in period t the cooperative strategy $f^c_{T^*}$ is chosen. For the case $E[f^d, \tilde{\mu}(t)] = E[f^c, \tilde{\mu}(t)]$ we assume that agents select the cooperative strategy. After some calculations (given in the appendix), we find that

$$E[f^c_{T^*}, \tilde{\mu}(t)] = \frac{1 + \delta^{T^*}(1-\sigma)^{T^*}(\tilde{\mu}(t)G - 1)}{(1-\delta)(1 - \delta^{T^*+1}(1-\sigma)^{T^*+1}(1 - \tilde{\mu}(t)))} \tag{2}$$

and

$$E[f^d_{T^*}, \tilde{\mu}(t)] = \frac{1 + \delta^{T^*}(1-\sigma)^{T^*}[(1 - \delta(1-\sigma))(\tilde{\mu}(t)H + 1 - \tilde{\mu}(t)) - 1]}{(1-\delta)(1 - \delta^{T^*+1}(1-\sigma)^{T^*+1})}. \tag{3}$$

We summarize the collection of parameters of the game by $\Gamma = \{H, G, \delta, \sigma\}$. With equations (2) and (3), we can show the following result:

Lemma 1 [Cooperative Intervals]

(a) For any payoffs H, G and given T^*, there are values $\bar{\delta} < 1$, $\bar{\sigma} > 0$, such that for $\delta \geq \bar{\delta}$ and $\sigma \leq \bar{\sigma}$ there is an interval $\nabla = [\underline{\mu}_{\Gamma,T^*}, \bar{\mu}_{\Gamma,T^*}]$ with $0 < \underline{\mu}_{\Gamma,T^*} < \bar{\mu}_{\Gamma,T^*} \leq 1$, where we have $E[f^c_{T^*}, \tilde{\mu}(t)] \geq E[f^d_{T^*}, \tilde{\mu}(t)]$ whenever $\tilde{\mu}(t) \in \nabla$, and $E[f^c_{T^*}, \tilde{\mu}(t)] < E[f^d_{T^*}, \tilde{\mu}(t)]$ otherwise.

(b) If $T^* > \frac{H-G}{G-1}$, then for any $\tilde{\mu}^* \in (0, 1]$, there are values $\bar{\delta} < 1$, $\bar{\sigma} > 0$, such that for $\delta \geq \bar{\delta}$ and $\sigma \leq \bar{\sigma}$, we have $E[f^c_{T^*}, \tilde{\mu}(t)] \geq E[f^d_{T^*}, \tilde{\mu}(t)]$ whenever $\tilde{\mu}(t) \in [\tilde{\mu}^*, 1]$.

Proof. See Appendix. □

If $T^* = 0$, then for very small and very high values of $\tilde{\mu}$, an agent in the market chooses the non-cooperative strategy. In between, she chooses the cooperative strategy if δ is sufficiently high and σ is sufficiently small. For an example, consider figure (I). We plot the sum of expected discounted utility for $T^* = 0$, $T^* = 1$ and $T^* = 2$. For $T^* = 0$ and $\Gamma_0 = \{3.5, 2, 0.98, 0.08\}$, there is no subjective belief at which an agent chooses the cooperative strategy. For $T^* = 0$ and $\Gamma_1 = \{3.5, 2, 0.98, 0.04\}$, we get that $\underline{\mu}_{\Gamma_1,0} \approx 0,083$ and $\bar{\mu}_{\Gamma_1,0} \approx 0,305$. If we increase T^*, the cooperative interval also increases. For $T^* = 2$, we have that $\bar{\mu}_{\Gamma_1,2} = 1$. The sequence of play in every period t is then as follows:

(i) Pairs in which both agents have chosen M in the previous period, are matched together with probability $1 - \sigma$.

(ii) Those agents who were not matched in [(i)], are paired up randomly.

(iii) Those agents who are matched to a new opponent, choose a strategy according to their current belief $\tilde{\mu}(t)$. All other agents maintain the strategy from the last period.

(iv) The PD is played according to the respective strategies.

(v) Agents observe the action choice of their opponent and evaluate her according to g.

(vi) Agents choose to maintain or to quit the relationship.

(vii) Agents update their beliefs to $\tilde{\mu}(t+1)$.

The assumption that agents only optimize when they are in the market, rules out that a player starts the relationship with strategy $f^c_{T^*}$ and switches to $f^d_{T^*}$ in the next period, after her opponent has been evaluated as cooperative. Without this assumption, the values of $\tilde{\mu}(t)$ and $\mu(t)$ no longer would be one-dimensional, as there are more than two types of observed behavior.

2.4 Nash Equilibria in Symmetric Strategies

Assume for a moment that agents have common knowledge of the aggregate behavior of the population. Further assume that

$$0 < \underline{\mu}_{\Gamma,T^*} < \bar{\mu}_{\Gamma,T^*} \leq 1.$$

Then one can show that for the considered strategies there are two mixed Nash-equilibria and one strict Nash-equilibrium in the described game:

(i) In the market, all agents play with probability $\underline{\mu}_{\Gamma,T^*}$ according to $f^c_{T^*}$ and with probability $1-\underline{\mu}_{\Gamma,T^*}$ according to $f^d_{T^*}$ in each period.

(ii) In the market, all agents play with probability $\bar{\mu}_{\Gamma,T^*}$ according to $f^c_{T^*}$ and with probability $1-\bar{\mu}_{\Gamma,T^*}$ according to $f^d_{T^*}$ in each period.

(iii) All agents play according to $f^d_{T^*}$ in each period.

At a later stage, we will compare the outcome of the game without knowledge of aggregate play to these equilibria.

2.5 Distribution of States, Beliefs and Histories

Each updating rule $\tilde{\mu}$ gives rise to a set of subjective beliefs $\mathcal{B}$ which potentially are reached. Let

$$Y(t) \in \Delta(\mathcal{B}^{T^*+2}) \tag{4}$$

be the distribution of states and beliefs in period t, i.e., a single element in $Y(t)$ is the share of players at the beginning of period t who have a certain subjective belief and are (or are not) in a long-term relationship since $l \in \{1, ..., T^*\}$ or more periods. Accordingly, denote by

$$Y^{HS}(t) \in \Delta(HS(n)^{T^*+2}) \tag{5}$$

the distribution of states and histories in period t, i.e., a single element in $Y^{HS}(t)$ is the share of players who have the same history of evaluations and are (or are not) in a long-term relationship since $l \in \{1, ..., T^*\}$ or more periods. The sequences

$$\mathcal{Y} = \{Y(t)\}_{t\geq 0}$$

and

$$\mathcal{Y}^{HS} = \left\{Y^{HS}(t)\right\}_{t\geq 0}$$

are implied by the updating rule $\tilde{\mu}$, T^*, n, Γ and the distribution of states and beliefs in the initial period, $Y^{HS}(0)$. We will call such a sequence a "process" without making further reference to the underlying parameters. Define a function

$$\tau_{\tilde{\mu}} : \Delta(HS(n)^{T^*+2}) \rightarrow \Delta(\mathcal{B}^{T^*+2}),$$

which assigns to each element of $\Delta(HS(n)^{T^*+2})$ the corresponding distribution of states and beliefs. Obviously, this mapping depends on the updating rule $\tilde{\mu}$. Therefore, we have

$$\mathcal{Y} = \{\tau_{\tilde{\mu}^{FP}}(Y^{HS}(t))\}_{t\geq 0}.$$

2.6 Definition of the Steady State and Stability

With the definitions of the preceding section, we can introduce the notion of a steady state to our framework:

Definition [Steady state]

A distribution of states and beliefs $Y^* \in \Delta(\mathcal{B}^{T^*+2})$ is called a steady state of $\mathcal{Y}$ if there exists a $Y^{HS} \in \tau_{\tilde{\mu}}^{-1}(Y^*)$, such that $Y^{HS}(t) = Y^{HS}$ implies

$Y^{HS}(t+s) \in \tau_{\tilde{\mu}}^{-1}(Y^*)$ *for all* $s > 0$*.*

We classify the steady states as follows:

Definition [Classification of steady states]

A steady state $Y^* \in \Delta(\mathcal{B}^{T^*+2})$ *of* $\mathcal{Y}$ *is called "non-cooperative" if the corresponding value of* μ^* *equals 0, while it is called "cooperative" if* $\mu^* > 0$*. A steady state* Y^* *with* $\mu^* = 1$ *is called "fully-cooperative".*

It is straightforward to adapt two important definitions of stability to our framework.

Definition [Lyapunov stability]

A state $Y \in \Delta(\mathcal{B}^{T^*+2})$ *is called Lyapunov stable with respect to* $\mathcal{Y}$ *if in every neighborhood* B *of* Y*, there is a neighborhood* B_0 *of* Y *with* $B_0 \subset B$*, such that for all* $t > 0$*, from* $Y(0) \in B_0 \cap \Delta(\mathcal{B}^{T^*+2})$ *it follows that* $Y(t) \in B$*.*

Definition [Asymptotic stability]

An element $Y \in \Delta(\mathcal{B}^{T^*+2})$ *is called asymptotically stable if it is Lyapunov stable, and there is a neighborhood* B *of* Y*, such that from* $Y(0) \in B$*, it follows that* $\lim_{t\to\infty} Y(t) = Y$. Note that we defined stability as a property of distributions of states and strategies, not histories. We therefore require that for an element $Y(t) \in \Delta(\mathcal{B}^{T^*+2})$ which is sufficiently close to an asymptotically stable steady state Y, we get convergence to Y, regardless of the distribution of histories which generates $Y(t)$.

3 On the Dynamics under Fictitious Play

3.1 Average over all Observations

Under fictitious play, every player has two weight functions $\kappa_c(t)$ and $\kappa_d(t)$ from which she calculates the subjective belief about the aggregate play of agents in the market. The weights are updated in the following way:

$$\kappa_c(t) = \sum_{i=-n+1}^{c(t)} \mathbf{1}_{\{g(h_{T^i})=1\}},$$

$$\kappa_d(t) = \sum_{i=-n+1}^{c(t)} \mathbf{1}_{\{g(h_{T^i})=-1\}},$$

where $\mathbf{1}$ is the indicator function. Individual beliefs are given by

$$\tilde{\mu}^{FP}(t) = \frac{\kappa_c(t)}{\kappa_c(t) + \kappa_d(t)}.$$

The subjective belief in the first period, $\tilde{\mu}^{FP}(0)$, therefore is determined by (1): the more elements in (1) take on the value 1, the more optimistic is the agent about the behavior of her opponents at the beginning. Since beliefs are always given by a rational number, we have

$$\mathcal{B} = \mathbb{R} \cap [0, 1].$$

The corresponding process $\mathcal{Y}^{HS}$ is called the fictitious play process. Let $y_{\tilde{\mu}}(t)$ denote the fraction of individuals in the market in period t with subjective belief equal to $\tilde{\mu} \in \mathcal{B}$ and let $y^m_{\tilde{\mu}}(t)$ denote the fraction of individuals in long-term relationships with subjective belief equal to $\tilde{\mu}$. For the case $T^* > 0$, we denote the fraction of individuals which have been in a relationship for $l \in \{1, ..., T^*\}$ periods and hold belief $\tilde{\mu}$ by $y^l_{\tilde{\mu}}(t)$.

Definition [Convergence]

The process $\mathcal{Y}$ converges to $Z \in \Delta(\mathcal{B}^{T^*+2})$ if for all $\epsilon > 0$, all $\tilde{\mu}^* \in \mathcal{B}$ and all $l \in \{1, ..., T^*\}$ it holds that for $t \to \infty$ we have

$$\begin{aligned}
\sum_{\tilde{\mu}\in[\tilde{\mu}^*-\epsilon,\tilde{\mu}^*+\epsilon]} y_{\tilde{\mu}}(t) &\to \sum_{\tilde{\mu}\in[\tilde{\mu}^*-\epsilon,\tilde{\mu}^*+\epsilon]} z_{\tilde{\mu}}, \\
\sum_{\tilde{\mu}\in[\tilde{\mu}^*-\epsilon,\tilde{\mu}^*+\epsilon]} y^l_{\tilde{\mu}}(t) &\to \sum_{\tilde{\mu}\in[\tilde{\mu}^*-\epsilon,\tilde{\mu}^*+\epsilon]} z^l_{\tilde{\mu}}, \\
\sum_{\tilde{\mu}\in[\tilde{\mu}^*-\epsilon,\tilde{\mu}^*+\epsilon]} y^m_{\tilde{\mu}}(t) &\to \sum_{\tilde{\mu}\in[\tilde{\mu}^*-\epsilon,\tilde{\mu}^*+\epsilon]} z^m_{\tilde{\mu}}.
\end{aligned}$$

For some value $\tilde{\mu}^* \in \mathcal{B}$, let $Y_{\tilde{\mu}^*} \subset \Delta(\mathcal{B}^{T^*+2})$ denote the collection of distributions of states and beliefs which assign their entire mass to belief $\tilde{\mu}^*$. Thus, $Y_0 \in \Delta(\mathcal{B}^{T^*+2})$ is the distribution of states and beliefs where every agent has a subjective belief of 0 and does not cooperate. With this, we can state:

Lemma 2 [Beliefs under convergence]

Let $Y^* \in \Delta(\mathcal{B}^{T^*+2})$ and μ^* the corresponding market share of cooperators. If the fictitious play process converges to Y^* or if Y^* is a steady state (or both), then $Y^* \in Y_{\mu^*}$.

Proof. See Appendix. □

For the stationarity assumption of fictitious play to make sense, we are particularly interested in the average behavior of the population. Since players are randomly and anonymously assigned to each other, every individual behaves as if she was always assigned to the same opponent who plays a mixed strategy given by the subjective belief $\tilde{\mu}(t)$. The crucial question is whether the average play in the population converges, rather than convergence of individual play. However, it can easily be checked from lemma 2 that convergence of individual play (as defined before) is in fact equivalent to convergence of aggregate play (convergence of μ). In particular, our model prevents deterministic cycles on the individual level which can arise from correlated play between players. Fudenberg and Kreps (1993) for example show that such cycles may persist under fictitious play even if the empirical distribution of actions converges. This is not possible in an anonymous random matching scheme, where players can only observe the actions of their particular opponents but ignore what the entire population is doing—compare Hopkins (1995).

Proposition 1 [Limit points and steady states]

Assume that $0 < \underline{\mu}_{\Gamma,T^*} < \bar{\mu}_{\Gamma,T^*} \leq 1$.

(a) The state Y_0 is a steady state of the fictitious play process. There exists a steady state $Y \in Y_{\bar{\mu}_{\Gamma,T^*}}$ if and only if $\bar{\mu}_{\Gamma,T^*} = 1$. There are no other steady states.

(b) If fictitious play converges to some $Y \in \Delta(\mathcal{B}^{T^*+2})$, then either $Y = Y_0$, $Y \in Y_{\underline{\mu}_{\Gamma,T^*}}$ or $Y \in Y_{\bar{\mu}_{\Gamma,T^*}}$.

Proof. See Appendix. □

One of the standard results about fictitious play states that every strict Nash equilibrium is an absorbing state—see for example Fudenberg and Levine (1998). The first part of proposition 1 shows that this result also holds in our framework. Another well known result about fictitious play is that if the empirical distribution over player's choices converges, then the corresponding strategy profile is a Nash equilibrium. This is what is stated in the second part of proposition 1. With the above learning rule, individuals asymptotically learn the true parameter μ—given that μ converges—and the limit sets support homogeneous beliefs. This property assures that we get a very clear prediction about the limit of the learning process.

3.2 Average over finitely many Observations

Under the previous updating rule, the speed of belief updating converges to 0. This is implausible whenever μ changes over time. The speed of updating remains constant if in a given period t, only the last n observations determine the subjective belief. This is also more appropriate if the agents' memory is finite and, more importantly, players account for the dynamic structure of the market. A subjective belief $\tilde{\mu}(t)$ then can take on only finitely many values, such that

$$\mathcal{B} = \left\{0, \frac{1}{n}, ..., \frac{n-1}{n}, 1\right\}. \tag{6}$$

Note that there is a trade-off in the size of n: If μ remains constant over time, the agent has a more precise estimate of μ if n is large. However, as long as μ varies over time, agents should replace very old observations quickly by new ones such that n should be limited. Define

$$k(t) = \arg_{k\in\mathbb{N}}\left\{\sum_{i=k}^{c(t)} \mathbf{1}_{\{g(h_{T^i})\neq 0\}} = n\right\}.$$

In words: $k(t)$ is the oldest of n observations with $g(h_{T^i}) \in \{-1, 1\}$, i.e., observations with an evaluation equal to 0 are not taken into account. As all of the considered observations have the same weight, the required updating rule is given by

$$\tilde{\mu}^A(t) = \frac{1}{n}\sum_{i=k(t)}^{c(t)} \mathbf{1}_{\{g(h_{T^i})=1\}}.$$

The subjective belief in the first period is again determined by the vector given in (1). As $\tilde{\mu}(t)$ can take on only finitely many values given in (6), we introduce the following notation related to the boundaries $\underline{\mu}_{\Gamma,T^*}$ and $\bar{\mu}_{\Gamma,T^*}$:

$$\begin{aligned} \underline{k}_{\Gamma,T^*,n} &= [n\underline{\mu}_{\Gamma,T^*}]_+, \\ \bar{k}_{\Gamma,T^*,n} &= [n\bar{\mu}_{\Gamma,T^*} - 1]_+, \end{aligned} \tag{7}$$

where $[\cdot]_+$ denotes the smallest integer which is larger or equal than the expression in the brackets. In order to keep notation tractable, we drop the subscripts $\{\Gamma, T^*, n\}$ in the following. If δ and n are sufficiently large and σ is sufficiently small, then $\underline{k} < \bar{k}$ and strategy $f^c_{T^*}$ is chosen in t if and only if $\tilde{\mu}(t) \in \left[\frac{\underline{k}}{n}, \frac{\bar{k}}{n}\right]$.

Consider first the cooperative strategy with $T^* = 0$. The set of distributions of beliefs and strategies is then given by $\Delta(\mathcal{B}^2)$. Define the elements of $Y(t) \in \Delta(\mathcal{B}^2)$ as follows: $y_i(t)$ is the share of players in period t who are in the market and have the belief $\tilde{\mu}(t) = \frac{i}{n}$, $i \in \{0, 1, ...n\}$. Furthermore, let $y^m_i(t)$ be the share of players who are in a long-term relationship and have the belief $\tilde{\mu}(t) = \frac{i}{n}$. The share of agents in the market with strategy f^c_0 in period t therefore is given by

$$y_C(t) = \sum_{i=\underline{k}}^{\bar{k}} y_i(t),$$

where the share of agents in the market with strategy f^d_0 is

$$y_D(t) = \sum_{i=0}^{\underline{k}} y_i(t) + \sum_{i=\bar{k}}^{n} y_i(t).$$

With these specifications, we get

$$\mu(t) = \frac{y_C(t)}{y_D(t) + y_C(t)}$$

as the share of agents in the market who behave cooperatively.

Obviously, for each T^* the element of $\Delta(\mathcal{B}^{T^*+2})$ with $y_0 = 1$ is a steady state.

Proposition 2 [**Non-cooperative steady state**]

Assume that for all agents $\tilde{\mu}$ is given by $\tilde{\mu}^A$, $T^* = 0$ and $n \geq 2$, such that n and Γ imply $\underline{k} \geq 2$. Then, the element $Y \in \Delta(\mathcal{B}^2)$ with $y_0 = 1$ is an asymptotically stable steady state.

Proof. See Appendix. □

Now assume that $T^* > 0$. Then, there is a share of agents in a long-term relationship whose opponent's evaluation still is equal to zero. Denote the share of players in period t who have the subjective belief $\frac{i}{n}$ and are in a long-term relationship since $l \in \{1, ..., T^*\}$ periods, by $y_i^l(t)$. The function $\tilde{\mu}^A$ is specified such that the belief does not change if the relationship is broken up by chance before agents start to cooperate. We have

$$y_i^{l+1}(t) = (1-\sigma)y_i^l(t-1) \tag{8}$$

for $i \in \{0, ..., n\}$ and $l \in \{1, ..., T^* - 1\}$. From Lemma 1(b) it follows that for appropriate values of δ and σ, we have $\underline{k} = 1$ and $\bar{k} = n$ if T is chosen sufficiently large. Then we get the following result.

Proposition 3 [Fully cooperative steady state]

Assume that for all agents $\tilde{\mu}$ is given by $\tilde{\mu}^A$, $n > 2$ and that for given payoffs H, G, we have $T^* > \frac{H-G}{G-1}$. Then, there are values $\bar{\delta}$, $\bar{\sigma}$, such that for $\delta \geq \bar{\delta}$, $\sigma \leq \bar{\sigma}$ and $y_0(0) < 1$, we have $\lim_{t\to\infty} \mu(t) = 1$, i.e., the fully cooperative steady state is asymptotically stable.

Proof. See Appendix. □

Thus, if the cooperative strategy involves sufficiently many periods of non-cooperation at the beginning of each relationship, we get almost global convergence to the fully cooperative steady state. Note that in this case each agent's subjective belief converges to 1 and behavior again converges to a Nash equilibrium. We therefore obtain the same results as in models with the assumption of common knowledge. However, for smaller values of T^*, it remains subject to simulations of the model, under which conditions convergence occurs and cooperation is a stable outcome.

4 Simulation

To complement on the analytical results of the last section, we simulated[2] the model for many parameter specifications Γ. We are mainly interested in the question whether μ converges to a limit point μ^* or not. In this section, we present the results for some illustrative examples. The statements below are valid for all specifications we ever considered. We spend most efforts on scenarios with $T^* = 0$, as

- for $T^* > 0$ and $\bar{\mu}_{\Gamma,T^*} < 1$, we make the same observations,
- for $T^* > 0$ and $\bar{\mu}_{\Gamma,T^*} = 1$, we have an analytical result on the outcome (for δ sufficiently high and σ sufficiently small) in the propositions 1 and 3

[2]The code can be downloaded from http://www.vwl.uni-mannheim.de/gk/_gaigl.

- for $T^* = 0$ the cooperative strategy is robust against communication: players can not improve the outcome of a cooperative long-term relationship.

For each of the considered scenarios Γ, we vary $n \in \{5, 10, 20\}$, and the distribution of initial beliefs. Note that for $\tilde{\mu}^{FP}$ the number n has only meaning for the distribution of beliefs in the initial period. The distributions of beliefs for each n are given table (I).

For $\tilde{\mu}^{A}$ we specify that in the first period, all histories which give rise to the same belief, have the same relative frequency. The following scenarios, Γ_1, Γ_2, Γ_3 and Γ_4, imply different cooperative intervals (for $T^* = 0$):

	H	G	δ	σ	$\underline{\mu}_{\Gamma,0}$	$\bar{\mu}_{\Gamma,0}$
Γ_1	3.5	2.0	0.98	0.04	0.083	0.305
Γ_2	3.5	3.0	0.98	0.04	0.032	0.780
Γ_3	2.0	1.5	0.99	0.01	0.044	0.467
Γ_4	2.0	1.9	0.99	0.01	0.023	0.899

Whether a limit point μ^* is implied by a steady-state or not, we know for $\tilde{\mu}^{FP}$ from the proposition 1. For $\tilde{\mu}^{A}$ we can conclude that this is the case if the distribution of beliefs converges to a constant $Y^* \in \Delta(\mathcal{B}^2)$.

The results of the presented examples are summarized in table (II). We immediately recognize:

Observation 1a *Under $\tilde{\mu}^{FP}$, μ converges under all scenarios, values of n and initial beliefs to either $\mu^* = 0$ or $\mu^* = \bar{\mu}_{\Gamma,0}$.*

The intuitive reason for this observation is that only the players that hold beliefs in the neighborhood of $\underline{\mu}_{\Gamma,0}$ and $\bar{\mu}_{\Gamma,0}$ change their actions as a result of changing belief. Since the size of the change in individual beliefs goes to zero over time, either the rate in which these players switch between actions, goes to zero, or the agents' beliefs become closer to 0, $\underline{\mu}_{\Gamma,0}$ or $\bar{\mu}_{\Gamma,0}$. Therefore, the distribution of beliefs in the population converges to a single value.

We find that the cooperative limit point is always given by $\bar{\mu}_{\Gamma,0}$. This can be explained intuitively as follows: any belief distribution concentrated around $\bar{\mu}_{\Gamma,0}$ that assigns more than mass $\bar{\mu}_{\Gamma,0}$ to cooperative beliefs, satisfies $\mu > \bar{\mu}_{\Gamma,0}$. Since individual's beliefs approach μ over time, more and more individuals switch to defection and hence, μ decreases. The reverse argument shows that the share of cooperators in the market increases if $\mu < \bar{\mu}_{\Gamma,0}$. For

$\underline{\mu}_{\Gamma,0}$ this argument does not work: If $\mu > \underline{\mu}_{\Gamma,0}$ holds for several periods, more and more agents switch to the cooperative strategy such that μ increases even further.

Observation 1b *Under $\tilde{\mu}^A$, μ converges under all scenarios, values of n and initial beliefs.*

To the best of the author's knowledge, there is no setting in which μ does not converge to a single value. However, among the different updating rules, the dynamics in the market in the first periods and the number of periods, until μ is close to the respective limit point, may vary substantially—see figure (II), [TOP]. We also can observe that the smaller σ is, the slower μ convergences, as agents are less often in the market.

Observation 2 *Under the updating rules $\tilde{\mu}^{FP}$ and $\tilde{\mu}^A$, we have $\mu^* > 0$ if the distribution of initial beliefs is not too pessimistic.*

We see from the table (II) that under $\tilde{\mu}^A$, there is no case in which μ converges to 0 although under some distributions of initial beliefs, agents are quite pessimistic in the first periods. Under $\tilde{\mu}^{FP}$, a cooperative outcome is reached if the distribution of initial beliefs is not too pessimistic.

This observation is of course dependent on the fact that there are some cooperative agents in the initial period. If $\mu(0) = 0$, then μ stays at this level forever. Very pessimistic beliefs in the first period are less harmful for cooperation if memory is finite, as preplay-observations will be substituted by more recent ones. This is not the case under $\tilde{\mu}^{FP}$, where agents recall the entire history of preplay-observations: the non-cooperative steady state is reached if the subjective beliefs in the initial period are very pessimistic—see figure (II), [Top-left].

Observation 3 *The limit points of the updating rules can differ substantially.*

We can observe that the conditions in the market for long-term relationships vary among the different updating rules: Under $\tilde{\mu}^A$, the cooperative limit point μ^* is bounded away from 0 and 1. The limit points under $\tilde{\mu}^{FP}$ and $\tilde{\mu}^A$ can differ substantially—see for example in the scenarios Γ_2 and Γ_4. In general, aggregate play does not converge to the Nash equilibrium under the updating rule with limited memory.

Finally, we display in figure (II), [BOTTOM], the distribution of beliefs in the market for all updating rules, when μ is close to its limit point in scenario Γ_1: Under $\tilde{\mu}^{FP}$, beliefs are distributed closely around $\bar{\mu}_{\Gamma_1,0}$ and will converge even further in the next periods. For $\tilde{\mu}^{A}$, we get significant variation in the distribution subjective beliefs in the steady state.

5 The Finite Population Process and its Deterministic Approximation

Up to now we used deterministic processes in order to analyze the population dynamics. We understood them as approximations to the stochastic population processes that occur when finitely many agents are randomly matched for interaction. In this section we explore whether the deterministic process is in fact a good approximation of the stochastic population process as the population size goes to infinity. In particular, we are interested in the relationship between the long run behavior of the deterministic and the stochastic process.[3]

In order to be consistent with the infinite population case, we must therefore exclude that players hold different beliefs about every particular opponent. We achieve this by assuming that players are matched anonymously. Moreover, we assume that players do not carry out any strategic reasoning, but simply play best responses to their current beliefs. This is a weak assumption in finite, but large population because every action has a very small effect on the overall dynamics.

In order to provide the first result, we look at some arbitrary period t and therefore skip time indices. We abstract from the set of matched players and only analyze the matching procedure in the market. Let M denote the number of individuals in the market. Consider the standard matching procedure that assigns one partner to every player, such that all pairs are equally likely. For any $hs \in HS(n)$, let the random variable $m^M(hs, f^c_{T^*})$ denote the fraction of players with history of evaluations hs that are matched to a partner

[3]This deterministic approximation approach is frequently used by biologists and economists for analyzing interaction in large populations where individuals are matched randomly. Many approximation results have been established so far—see for example Boylan (1995), Corradi and Sarin (1999) or Benaïm and Weibull (2003). These models assume that the time between two matches as well as the fraction of the population which is matched each time, are diminishing over time. Typically, this results in a differential equation in the limit. In our model, the entire population is matched at fixed points in time. Since we are not aware of an approximation results for such a framework, we are going to provide one in this section.

which plays according to strategy $f^c_{T^*}$. As before, μ denotes the fraction of individuals in the market which play according to strategy $f^c_{T^*}$. The following lemma provides a version of the law of large numbers which is adapted to our needs.

Lemma 3 [A law of large numbers]

For all $hs \in HS(n)$ and any $\epsilon > 0$:

$$\lim_{M\to\infty} \Pr\left[|m^M(hs, f^c_{T^*}) - \mu| > \epsilon\right] = 0.$$

Proof. See Appendix. □

The definition of the stochastic population process

$$\mathcal{X}^{HS}_N = \left\{X^{HS}(t)\right\}_{t\geq 0}$$

is straightforward from our model. The state space of $\mathcal{X}^{HS}_N$ is given by $\Delta(HS(n)^{T^*+2})$. Moreover, for any updating rule $\tilde{\mu}$, we denote by $\mathcal{X}_N$ the process $\tau_{\tilde{\mu}}(\mathcal{X}^{HS}_N)$ induced by $\mathcal{X}^{HS}_N$. The deterministic approximation process $\mathcal{Y}^{HS}$ is derived from $\mathcal{X}^{HS}_N$ in the following way. First, assume a continuum population and denote by y_{hs} the share of individuals in the market with history of evaluations $hs \in HS(n)$. Now, probabilities are replaced by shares: for all $hs \in HS(n)$, the fraction $y_{hs}m^M(hs, f^c_{T^*})$ of individuals in the market meets an individual which plays strategy $f^c_{T^*}$ and the fraction $y_{hs}(1-m^M(hs, f^c_{T^*}))$ meets an individual which plays strategy $f^d_{T^*}$. Accordingly, individuals switch to new individual histories (and possibly get matched or divorced). Moreover, for every $hs \in HS(n)$ and $t \leq T^*$, the share σ of the matched individuals gets divorced. Let $\mathcal{Y}$ denote the process $\tau_{\tilde{\mu}}(\mathcal{Y}^{HS})$ induced by $\mathcal{Y}^{HS}$. The following result shows that $\mathcal{Y}$ can be considered as the limiting case of the random process $\mathcal{X}_N$.

Proposition 4 [Finite population process]

For a given population of size N, consider the deterministic process $\mathcal{Y}^{HS}$ derived from $\mathcal{X}^{HS}_N$ as described above. Moreover, consider the corresponding processes $\mathcal{Y}$ and $\mathcal{X}_N$. If $X^{HS}_N(0) = Y^{HS}(0)$, then [4]

$$\Pr\left[d\left(X_N(t), Y(t)\right) > \epsilon\right] \to 0 \text{ as } N \to \infty,$$

for all $\epsilon > 0$ and all $t > 0$.

[4]For any two distributions of states and beliefs Y and Z, we choose the distance function in a way that it is well-defined whenever the supporting beliefs of Y and Z are discrete

Proof. See Appendix. □

Consider any asymptotically stable steady state $Y^* \in \Delta(\mathcal{B}^{T^*+2})$. We define its basin of attraction as the set of all initial states in $\Delta(HS(n)^{T^*+2})$ from which $\mathcal{Y}$ approaches Y^*.[5] Formally,

$$\begin{aligned} B(Y^*) &= \{Y^{HS} \in \Delta(HS(n)^{T^*+2}) : Y^{HS}(0) = Y^{HS} \Rightarrow \\ &\quad d(Y(t), Y^*) \to 0 \text{ as } t \to \infty\}. \end{aligned}$$

The following corollary is straightforward.

Corollary 1

Let $Y^* \in \Delta(\mathcal{B}^{T^*+2})$ be an asymptotically stable steady state. Whenever

$$X_N^{HS}(0) \in B(Y^*),$$

then

$$\lim_{t\to\infty} \lim_{N\to\infty} \Pr\left[d(X_N(t), Y^*) > \epsilon\right] = 0,$$

for all $\epsilon > 0$.

The corollary assures that $\mathcal{X}_N$ is likely to be close to Y^* for a very long time, given that the population is sufficiently large and that $\mathcal{X}_N^{HS}$ starts in the basin of attraction of Y^*. In most of our simulations we found two asymptotically stable steady states of $\mathcal{Y}$, namely a cooperative and a non-cooperative one. It follows from the corollary that they can be used as predictors of the outcome of the finite population process in the medium and long run.

This result may be satisfactory in many frameworks. However, note that for finite N, the outcome of $\mathcal{X}_N$ in the very long run might still be far away from Y^*, even if $\mathcal{X}_N^{HS}$ starts in the basin of attraction of Y^*.

points in the interval $[0,1]$. Let $\widetilde{\mu} \in [0,1]$ and $L = \{1, ..., T^*\}$ (assume that $L = \emptyset$ if T^*=0).

$$d(Y,Z) = \sum_{\widetilde{\mu}\,:\, y_{\widetilde{\mu}}>0\,,\, z_{\widetilde{\mu}}>0} |y_{\widetilde{\mu}} - z_{\widetilde{\mu}}| + \sum_{l\in L} \sum_{\widetilde{\mu}\,:\, y^l_{\widetilde{\mu}}>0\,,\, z^l_{\widetilde{\mu}}>0} |y^l_{\widetilde{\mu}} - z^l_{\widetilde{\mu}}| + \sum_{\widetilde{\mu}\,:\, y^m_{\widetilde{\mu}}>0\,,\, z^m_{\widetilde{\mu}}>0} |y^m_{\widetilde{\mu}} - z^m_{\widetilde{\mu}}|$$

[5]Note that in case of fictitious play, Y^* has discrete support in $[0,1]$ (lemma 2) and therefore Y^* is a feasible argument of the distance function d.

Proposition 5 [Long run outcome in a finite population]

Consider some finite N (not too small) and the updating rule $\tilde{\mu}^A$. Then

$$\lim_{t\to\infty} \Pr\left[X_N(t) = Y_0\right] = 1, \text{ if } \underline{k} \geq 2, \text{ and}$$

$$\lim_{t\to\infty} \Pr\left[X_N(t) = Y_0 | X_N(0) \neq Y_0\right] = 0, \text{ if } \underline{k} = 1.$$

Proof. See Appendix. □

6 Conclusion

In this paper we analyzed the dynamics in a market for long-term relationships when agents do not observe aggregate play but rather form beliefs through their private observations. We saw that for a large measure of initial distributions of beliefs, aggregate play converges to a cooperative outcome if agents update their beliefs based on past experiences. This remains true if strategies are very simple and punishment within a relationship is not possible. The steady state has the property that, even though aggregate play remains constant, different agents make different experiences in the market and therefore act differently. We observed from the simulations of the model that aggregate play converges in many (if not all) cases. However, if agents base their subjective belief on finitely many observations, aggregate play in a steady state may not be consistent with a Nash equilibrium of the game.

We experienced that the analytical tools for analyzing the considered model are fairly limited. Future research may concentrate on finding methods from non-linear dynamics in order to make the model tractable. The benefit could be the identification of tools for predicting behavior in more complex games without the assumption of common knowledge.

Appendix

Calculation of (2) and (3)

We calculate these two terms from the expressions

$$
\begin{aligned}
E[f^c_{T^*}\tilde{\mu}(t)] \;\; = \;\; & \tilde{\mu}(t)\left[\sigma \sum_{T=0}^{T^*-1}\left[(1-\sigma)^T \sum_{\tau=0}^{T}\delta^\tau\right] + \sigma \sum_{T=T^*}^{\infty}\left[(1-\sigma)^T\left(\sum_{\tau=0}^{T^*-1}\delta^\tau + \sum_{\tau=T^*}^{T}\delta^\tau G\right)\right]\right. \\
& \left. + \sigma \sum_{\tau=0}^{\infty}\delta^{\tau+1}(1-\sigma)^\tau E[f^c_{T^*},\tilde{\mu}(t)]\right] + (1-\tilde{\mu}(t))\left[\sigma \sum_{T=0}^{T^*-1}\left[(1-\sigma)^T\sum_{\tau=0}^{T}\delta^\tau\right]\right. \\
& \left. + \sigma \sum_{\tau=0}^{T^*-1}\delta^{\tau+1}(1-\sigma)^\tau E[f^c_{T^*},\tilde{\mu}(t)] + \delta^{T^*+1}(1-\sigma)^{T^*}E[f^c_{T^*}\tilde{\mu}(t)]\right],
\end{aligned}
$$

and

$$
\begin{aligned}
E[f^d_{T^*}\tilde{\mu}(t)] \;\; = \;\; & \sigma \sum_{T=0}^{T^*-1}\left[(1-\sigma)^T\sum_{\tau=0}^{T}\delta^\tau\right] + \tilde{\mu}(t)\delta^{T^*}(1-\sigma)^{T^*}H + (1-\tilde{\mu}(t))\delta^{T^*}(1-\sigma)^{T^*} \\
& + \sigma \sum_{\tau=0}^{T^*-1}\delta^{\tau+1}(1-\sigma)^\tau E[f^d_{T^*},\tilde{\mu}(t)] + \delta^{T^*+1}(1-\sigma)^{T^*}E[f^d_{T^*},\tilde{\mu}(t)].
\end{aligned}
$$

Proof of Lemma 1

First consider the ratio

$$
E_{T^*}(\tilde{\mu}) = \frac{E[f^c_{T^*},\tilde{\mu}]}{E[f^d_{T^*},\tilde{\mu}]}. \tag{9}
$$

Taking the limit yields us

$$
\lim_{\sigma\to 0} E_{T^*}(\tilde{\mu}) = \frac{[1+\delta^{T^*}(\tilde{\mu}G-1)][1-\delta^{T^*+1}]}{[1-(1-\tilde{\mu})\delta^{T^*+1}][1+\delta^{T^*}((1-\delta)(\tilde{\mu}H+1-\tilde{\mu})-1)]},
$$

and with l'Hospitals rule

$$
\lim_{\delta\to 1}\lim_{\sigma\to 0} E_{T^*}(\tilde{\mu}) = \frac{G(T^*+1)}{(T^*+1)+\tilde{\mu}(H-1)}, \tag{10}
$$

which is larger than 1 for all $\tilde{\mu} \in [0,1]$ if T^* is chosen sufficiently high. Further, it follows that the right-hand side of (10) is larger than 1 if

$$
\tilde{\mu} < (T^*+1)\frac{G-1}{H-1}.
$$

As δ and σ enter (9) continuously, part (b) of the result follows from (10).

Now fix a $\tilde{\mu}^*$ such that the right-hand side of (10) is strictly larger than 1 if $\tilde{\mu} = \tilde{\mu}^*$. As $E_{T^*}(\tilde{\mu})$ is continuous in δ and σ, there are values $\bar{\delta}$ and $\bar{\sigma}$, such that

$$
E[f^c_{T^*},\tilde{\mu}^*] > E[f^d_{T^*},\tilde{\mu}^*] \tag{11}
$$

whenever $\delta \geq \bar{\delta}$ and $\sigma \leq \bar{\sigma}$. We can calculate that

$$
\lim_{\tilde{\mu}\to 0} E_{T^*}(\tilde{\mu}) = \frac{1-\delta^{T^*}(1-\sigma)^{T^*}}{1-\delta^{T^*+1}(1-\sigma)^{T^*+1}} < 1. \tag{12}
$$

Define

$$
\nabla E_{T^*}(\tilde{\mu}) = E[f^c_{T^*},\tilde{\mu}] - E[f^d_{T^*},\tilde{\mu}].
$$

From (11) and (12) we know that $\nabla E_{T^*}(\tilde{\mu}^*) > 0$ and $\nabla E_{T^*}(0) < 0$. Further from

$$\frac{\partial \nabla E_{T^*}(\tilde{\mu})^2}{\partial^2 \tilde{\mu}} = \frac{-2G(1-\delta)^2 \delta^{T^*+1}(1-\sigma)^{T^*+1}[A - \delta^{T^*+1}(1-\sigma)^{T^*+1}(1-\delta^{T^*}(1-\sigma)^{T^*})]}{[(1-\delta)(1-\delta^{T^*+1}(1-\sigma)^{T^*+1}) + (1-\delta)\delta^{T^*+1}(1-\sigma)^{T^*+1}\tilde{\mu}]^3} < 0$$

with

$$A = \delta^{T^*}(1-\sigma)^{T^*} G(1-\delta^{T^*+1}(1-\sigma)^{T^*+1})$$

we get that $\nabla E_{T^*}(\tilde{\mu})$ is a concave function. With this, part (a) of the result follows.
Q.E.D.

Proof of Lemma 2

We assume that Y^* is the limit point of the fictitious play process (in case Y^* is a steady state, the proof is analogue). Suppose, by contradiction, that $Y^* \in \Delta(\mathcal{B}^{T^*+2})$ assigns positive mass to beliefs different from μ^*. Since $\mu(t) \to \mu^*$ as $t \to \infty$, it follows from the properties of fictitious play that every individual belief $\tilde{\mu}$ converges to μ^* as well, a contradiction.
Q.E.D.

Proof of Proposition 1

(a) Y_0 is a steady state because $\forall t > 0 \,:\, Y(t) = Y_0 \Rightarrow \mu(t) = 0 \Rightarrow Y(t+1) = Y_0$. If $\bar{\mu} = 1$, it can easily be checked that the state $Y \in Y_1$ with $y_1 = \sigma$, $y_1^l = \sigma(1-\sigma)^l$ for all $l \in \{1, .., T^*\}$ and $y_1^m = 1 - \sum_{i \geq 0}^{T^*} \sigma(1-\sigma)^i$ is a steady state. On the other hand, if $\bar{\mu} < 1$, no element of the set Y_1 can be a steady state, because:

$$\forall t > 0 \,:\, Y(t) \in Y_1 \Rightarrow \mu(t) = 0 \Rightarrow Y(t+1) \notin Y_1.$$

From Lemma 2 we know that every steady state belongs to some set Y_x. No state $Y \in Y_x$ with $x \notin \{\{0\}, \{1\}\}$ can be a steady state which can be seen as follows. Let all individuals hold belief x at time t. It is easy to see that at least at one of the dates t and $t+1$ the market is non-empty. Hence, at time $t+2$, there will be some individuals with either one of the beliefs $\frac{(n+t)x+1}{n+t+1}$ or $\frac{(n+t)x}{n+t+1}$.

(b) Assume that fictitious play converges to some $Y \in Y_x$ and assume that $x \in (\underline{\mu}, \bar{\mu})$. Since individual beliefs converge to x, for any $b \in (0,1)$, there exists a $t > 0$ such that $\mu(t) > b$, a contradiction. The similar argument holds for the cases $x \in (0, \underline{\mu})$ and $x \in (\bar{\mu}, 1)$.
Q.E.D.

Proof of Proposition 2

To economize on notation, let $y_i(t)$ be also the "set" of agents with belief $\frac{i}{n}$ in period t. Consider two periods t and $t+n$. We compare the flow $\nabla^1(t)$ from the set $1 - y_0(t) - y_1(t)$ to the set $y_0(t+n) + y_1(t+n)$ and the flow $\nabla^2(t)$ of agents from the set $y_0(t) + y_1(t)$ to the set $1 - y_0(t+n) - y_1(t+n)$. As $y_C(t) \leq 1 - y_0(t) - y_1(t)$, we have $\mu(t) \leq 1 - y_0(t) - y_1(t)$. Further, we can estimate

$$\max\{\mu(\tau) \mid \tau \in [t, t+n]\} \leq 2^n(1 - y_0(t) - y_1(t)),$$

as in each period, every cooperator can only meet one defector in the market who himself becomes a cooperator. Thus, the probability that an agent who is in the set $1 - y_0(t) - y_1(t)$, is also in the set $y_0(t+n)$, is at least $\sigma(1 - 2^n(1 - y_0(t) - y_0(t)))^n$. Then, we have

$$\nabla^1(t) > \sigma(1 - 2^n(1 - y_0(t) - y_1(t)))^n(1 - y_0(t) - y_1(t)).$$

An agent in the set $y_0(t)$ [$y_1(t)$] must meet at least two [one] cooperators in the market to become a cooperator himself between the periods t and $t+n$. Therefore we can estimate that

$$\nabla^2(t) < n\left(2^{2n}(1 - y_0(t) - y_1(t))^2 y_0(t) + 2^n(1 - y_0(t) - y_1(t))y_1(t)\right).$$

We therefore get

$$\frac{\nabla^1(t)}{\nabla^2(t)} > \frac{\sigma(1 - y_0(t) - y_1(t))}{n(2^{2n}(1 - y_0(t) - y_1(t))y_0(t) + 2^n y_1(t))}, \tag{13}$$

which is larger than 1 if $y_0(t)$ is sufficiently close to 1. In this case, we have $\mu(t+n) < \mu(t)$ and $y_0(t+n) > y_0(t)$. By going through the same steps, one can also show that for sufficiently large $y_0(t)$

$$\mu(t + n - 1) < \mu(t). \tag{14}$$

With

$$y_1(t + 1) < (1 - \mu(t))y_2(t) + \mu(t) < (1 - \mu(t))\mu(t) + \mu(t) < 2\mu(t)$$

and (14) we get $y_1(t + n) < 2\mu(t + n - 1) < 2\mu(t)$. Thus, we have

$$\frac{\nabla^1(t+n)}{\nabla^2(t+n)} > \frac{\sigma(1 - y_0(t) - y_1(t))}{n(2^{2n}(1 - y_0(t) - y_1(t))y_0(t) + 2^{n+1}\mu(t))}. \tag{15}$$

If the expressions on the right-hand side of (13) and (15) are larger than 1, this yields us

$$\frac{\nabla^1(t+in)}{\nabla^2(t+in)} > \frac{\sigma}{n2^n\mu(t)} > 1$$

for all $i \geq 2$. Thus, it follows that if $Y(t)$ is sufficiently close to Y_0, we get

$$\lim_{t\to\infty} y_0(t) = 1,$$

which completes the proof.
Q.E.D.

Proof of Proposition 3

Assume that $\underline{k} = 1$ and $\bar{k} = n$ where $n \geq 2$. Consider the set $\nabla_1(t)$ of agents who are both in $y_1(t)$ and $y_0(t + T^* + 1)$ and the set $\nabla_2(t)$ of agents who are both in $y_0(t)$ and $y_1(t + T^* + 1)$. We then have

$$\begin{aligned} \nabla_1(t) &\leq (1-\sigma)^{T^*}(1-\mu(t))y_1(t), \\ \nabla_2(t) &= (1-\sigma)^{T^*}\mu(t)y_0(t). \end{aligned}$$

From the definition of $\mu(t)$ it follows that $\nabla_2(t) > \nabla_1(t)$ if $\sum_{i=2}^{n} y_i(t) > 0$. Furthermore, it follows from the assumption $y_0(0) < 1$ that $y_0(t) > 0$ implies $y_1(t + T^* + 1) > 0$ and $y_2(t + 2T^* + 2 + s) > 0$ for all $s \geq 0$. Therefore, we get

$$\lim_{t\to\infty} y_0(t) = 0,$$

which implies that

$$\lim_{t\to\infty} \mu(t) = 1,$$

as

$$y_C(t) > \sigma(1 - y_0(t - 1)).$$

The result then follows from lemma 1(b).
Q.E.D.

Proof of Lemma 3

We proof statement (9) by showing that the expected value of $m^M(hs, f^c_{T^*})$ goes to μ and that its variance goes to 0 (this implies convergence in probability, which is equivalent to (9)). Consider a market of size $M \in \mathbb{N}$ and any $hs \in HS(n)$ such that the best response to belief $\tilde{\mu}(h)$ is $f^c_{T^*}$. Let L denote the number of individuals in the market with history of evaluations hs and let y_{hs} denote the corresponding fraction of the population. Moreover, we denote by C the number of individuals which play $f^c_{T^*}$. Hence,

$$\begin{aligned} L &= y_{hs}M \\ C &= \mu M. \end{aligned}$$

Let L_c denote the number of individuals of class hs that are matched to an individual which plays $f^c_{T^*}$. We decompose L_c into the number $L_{c,1}$ of individuals that are matched to group hs and the number $L_{c,2}$ of individuals that are matched to individuals which play $f^c_{T^*}$ but not to group hs:

$$L_c = L_{c,1} + L_{c,2}.$$

It will be convenient to decompose the matching procedure into two steps. In the first step, all individuals from the population are assigned randomly to two equally large subsets A and B. In the second step, each of the individuals in A is assigned randomly to one of the individuals in B. Clearly, this matching procedure is equivalent to the one-step random matching procedure. Let L_A denote the number of individuals from class hs that are assigned to set A. Given the first step of this matching procedure, L_A follows a hypergeometric distribution with parameters $M/2$, L and M. It is well known that $\mathbf{E}(L_A) = (M/2) \cdot (L/M) = L/2$. The second moment of L_A can be calculated as follows:

$$\begin{aligned} \mathbf{E}(L_A^2) &= \mathbf{E}(L_A)(L_A - 1) + \mathbf{E}(L_A) \\ &= \sum_{k=0}^{\min\{L,M/2\}} \frac{k(k-1)\binom{L}{k}\binom{M-L}{M/2-k}}{\binom{M}{M/2}} + \mathbf{E}(L_A) \\ &= L(L-1)\frac{\binom{M-2}{M/2-2}}{\binom{M}{M/2}} \sum_{k=2}^{\min\{L-2,M/2-2\}} \frac{\binom{L-2}{k-2}\binom{M-2-(L-2)}{M/2-2-(k-2)}}{\binom{M-2}{M/2-2}} + \mathbf{E}(L_A) \end{aligned}$$

Since the sum is equal to one, one gets

$$\mathbf{E}(L_A^2) = \frac{L(L-1)M/2(M/2-1)}{M(M-1)} + \frac{M/2L}{M} = \frac{L(L-1)(M-2)}{4 \cdot (M-1)} + \frac{L}{2} \tag{16}$$

Note that $L_{c,1}$ is always an even number. For given L_A, $L_{c,1}/2$ follows a hypergeometric distribution with parameters L_A, $L - L_A$ and $M/2$, hence

$$\begin{aligned} \mathbf{E}(L_{c,1}) &= \mathbf{E}(\mathbf{E}(L_{c,1}|L_A)) \\ &= 2 \cdot \mathbf{E}\left(\frac{L_A(L-L_A)}{M/2}\right) \\ &= \frac{4}{M}\left(L\mathbf{E}(L_A) - \mathbf{E}(L_A^2)\right) \\ &= \frac{L(L-1)}{M-1} \end{aligned}$$

Moreover, for given $L_{c,1} = c$, $L_{c,2}$ follows a hypergeometric distribution with parameters $L - c$, $C - L$ and $M - L$. Hence,

$$\begin{aligned} \mathbf{E}(L_{c,2}) &= \mathbf{E}(\mathbf{E}(L_{c,2}|L_{c,1})) \\ &= \mathbf{E}\left(\frac{(L-L_{c,1})(C-L)}{M-L}\right) \\ &= \frac{\left(L - \frac{L(L-1)}{M-1}\right)(C-L)}{M-L} \end{aligned}$$

Thus, we can establish the first part of the proof:

$$\begin{aligned}\mathbf{E}\left(m^M(hs, f^c_{T^*})\right) &= \frac{1}{L}\left(\mathbf{E}(L_{c,1}) + \mathbf{E}(L_{c,2})\right) \\ &= \frac{1}{L}\left(\frac{L(C-1)}{M-1}\right) \\ &= \frac{\mu M - 1}{M-1} \to \mu \text{ as } M \to \infty \end{aligned} \tag{17}$$

It remains to show that the following expression goes to zero:

$$\begin{aligned}\operatorname{Var}\left(m^M(hs, f^c_{T^*})\right) &= \frac{\mathbf{E}(L_c^2) - \mathbf{E}^2(L_c)}{L^2} \\ &= \frac{1}{L^2}\left[\mathbf{E}(L_{c,1}^2) + \mathbf{E}(L_{c,2}^2) + 2\mathbf{E}(L_{c,1}L_{c,2}) - \mathbf{E}^2(L_c)\right] \end{aligned} \tag{18}$$

Using equation (16) we get

$$\begin{aligned}\mathbf{E}(L_{c,1}^2) &= 4 \cdot \mathbf{E}\left(\frac{(L-L_A)(L-L_A-1)L_A(L_A-1)}{\frac{M}{2}(\frac{M}{2}-1)}\right) + 2\frac{L(L-1)}{M-1} \\ &= \frac{16 \cdot \left[(L(1-L))\mathbf{E}(L_A) + (L^2+L-1)\mathbf{E}(L_A^2) - 2L\mathbf{E}(L_A^3) + \mathbf{E}(L_A^4)\right]}{M(M-2)} + 2\frac{L(L-1)}{M-1},\end{aligned}$$

The third and the fourth moment of the hypergeometric distribution can be calculated in the same way as the second moment, namely

$$\begin{aligned}\mathbf{E}(L_A^3) &= \frac{L(L-1)(L-2)M/2(M/2-1)(M/2-2)}{M(M-1)(M-2)} + 3 \cdot \mathbf{E}(L_A^2) - 2 \cdot \mathbf{E}(L_A) \\ &= \frac{(M-4)L(L-1)(L-2)}{8(M-1)} + 3 \cdot \mathbf{E}(L_A^2) - 2 \cdot \mathbf{E}(L_A) \end{aligned} \tag{19}$$

and

$$\begin{aligned}\mathbf{E}(L_A^4) &= \frac{L(L-1)(L-2)(L-3)M/2(M/2-1)(M/2-2)(M/2-3)}{M(M-1)(M-2)(M-3)} \\ &\quad + 6 \cdot \mathbf{E}(L_A^3) - 11 \cdot \mathbf{E}(L_A^2) + 6 \cdot \mathbf{E}(L_A) \\ &= \frac{(M-4)(M-6)L(L-1)(L-2)(L-3)}{16(M-1)(M-3)} + 6 \cdot \mathbf{E}(L_A^3) - 11 \cdot \mathbf{E}(L_A^2) + 6 \cdot \mathbf{E}(L_A) \end{aligned} \tag{20}$$

Plugging equations (16), (19) and (20) into the expression for $\mathbf{E}(L_{c,1}^2)$, dividing by L^2 and replacing L by $y_{hs}M$ yields

$$\frac{\mathbf{E}(L_{c,1}^2)}{L^2} \to 16 \cdot \left[0 + \frac{y_{hs}^2}{4} - \frac{2y_{hs}^2}{8} + \frac{y_{hs}^2}{16}\right] + 0 = y_{hs}^2, \text{ as } M \to \infty. \tag{21}$$

Furthermore,

$$\begin{aligned}\mathbf{E}(L_{c,2}^2) &= \mathbf{E}\left(\frac{(C-L)(C-L-1)(L-L_{c,1})(L-L_{c,1}-1)}{(M-L)(M-L-1)} + \frac{(L-L_{c,1})(C-L)}{M-L}\right) \\ &= \frac{(C-L)(C-L-1)}{(M-L)(M-L-1)}\left[\mathbf{E}(L_{c,1}^2) + (1-2L)\mathbf{E}(L_{c,1}) + L(L-1)\right] + \frac{C-L}{M-L}(L - \mathbf{E}(L_{c,1}))\end{aligned}$$

Using $\mathbf{E}(L_{c,1})/L \to y_{hs}$ and (21), we get from the previous equation

$$\frac{\mathbf{E}(L_{c,2}^2)}{L^2} \to \frac{(\mu - y_{hs})^2}{(1-y_{hs})^2}\left[y_{hs}^2 - 2y_{hs} + 1\right] + 0 = (\mu - y_{hs})^2, \text{ as } M \to \infty.$$

Finally, the third term of equation (18) can be evaluated as follows:

$$\begin{aligned}\mathbf{E}(L_{c,1}L_{c,2}) &= \mathbf{E}(\mathbf{E}(L_{c,1}L_{c,2}|L_{c,1})) \\ &= \mathbf{E}\left(L_{c,1}\frac{(L-L_{c,1})(C-L)}{M-L}\right) \\ &= \frac{C-L}{M-L}\left[L\mathbf{E}(L_{c,1})-\mathbf{E}(L_{c,1}^2)\right].\end{aligned}$$

It has the following limit:

$$\frac{\mathbf{E}(L_{c,1}L_{c,2})}{L^2} \to \frac{\mu-y_{hs}}{1-y_{hs}}\left[y_{hs}-y_{hs}^2\right] = y_{hs}(\mu-y_{hs}), \text{ as } M\to\infty.$$

Since all limits exist, sums and limits and interchangeable and we finally get from equation (18)

$$\mathrm{Var}\left(m^M(hs, f_{T^*}^c)\right) \to y_{hs}^2 + (\mu-y_{hs})^2 + 2y_{hs}(\mu-y_{hs}) - \mu^2 = 0, \text{ as } M\to\infty. \tag{22}$$

It follows from (17) and (22) that $m^M(hs, f_{T^*}^c) \xrightarrow{p} \mu$. The same proof applies to $f_{T^*}^d$ (by just renaming $f_{T^*}^c$ by $f_{T^*}^d$). It follows that $m^M(hs, f_{T^*}^c) \xrightarrow{p} \mu$ holds for all $hs \in HS$. This accomplishes the proof.
Q.E.D.

Proof of Proposition 4

For any $Y^{HS} \in \Delta(HS(n)^{T^*+2})$, define

$$||Y^{HS}|| = \sum_{hs\in HS(n)}\sum_{s\in\{0,..,T^*+1\}} |Y_{hs,s}^{HS}|.$$

Moreover, for any $hs \in HS(n)^{T^*+2}$, let $\mathcal{Y}_{hs}^{HS}$ denote the process $\mathcal{Y}^{HS}$ with initial condition $Y^{HS}(0) = hs$ and define $U_t^N = X_N^{HS}(t+1) - Y_{X_N^{HS}(t)}(1)$. By Lemma 3 and by construction of the process $\mathcal{Y}^{HS}$, we know that for all population shares $hs \in HS(n)$ and states $s \leq T^*$:

$$|X_{N,i,s}^{HS}(t+1) - Y_{X_N^{HS}(t),i,s}(1)| \xrightarrow{p} 0.$$

Hence, $||U_t^N|| =\xrightarrow{p} 0$, and therefore $\sum_{s=0}^t ||U_s^N|| \xrightarrow{p} 0$.

From $Y_{X_N^{HS}(0)}^{HS}(1) = Y^{HS}(1)$ and continuity of $\mathcal{Y}^{HS}$ it follows that

$$||Y_{X_N^{HS}(1)}(1) - Y^{HS}(2)|| \xrightarrow{p} 0 \tag{23}$$

$$\vdots$$

$$||Y_{X_N^{HS}(t)}(1) - Y^{HS}(t+1)|| \xrightarrow{p} 0, \tag{24}$$

and hence

$$\sum_{s=0}^t ||Y_{X_N^{HS}(s)}^{HS}(1) - Y^{HS}(s+1)|| \xrightarrow{p} 0.$$

This implies

$$\begin{aligned}&\Pr\left[||X_N^{HS}(t) - Y^{HS}(t)|| > \epsilon\right] \\ = &\Pr\left[||\sum_{s=0}^t Y_{X_N^{HS}(s)}^{HS}(1) - Y^{HS}(s+1) + U_s^N|| > \epsilon\right] \\ \leq &\Pr\left[\sum_{s=0}^t ||Y_{X_N^{HS}(s)}^{HS}(1) - Y^{HS}(s+1)|| + ||\sum_{s=0}^t U_s^N|| > \epsilon\right] \to 0\,\forall\epsilon > 0.\end{aligned}$$

Finally, the claim follows from the fact that the updating function $\tau_{\tilde{\mu}}$ is continuous.
Q.E.D.

Proof of Proposition 5

Since we only consider the three updating rules with finite use of information, we can redefine $\mathcal{X}_N^{HS}$ as a process on the space of finite histories of length n, without changing the dynamics. In this case, $\mathcal{X}_N^{HS}$ clearly satisfies the Markov property and has finite state space. Moreover, the state Y_0 is an absorbing state of the corresponding process $\mathcal{X}_N$ because for each of the three updating rules, it holds that $\mathcal{X}_N(t) = Y_0 \Rightarrow \mathcal{X}_N(t+1) = Y_0$. Now we are going to show that there exists a $s \in \mathbb{N}$ such that for all times $t \in \mathbb{N}$ and all states $Y \in \Delta(\mathcal{B}^{T^*+2})$, the Markov chain $\mathcal{X}_N^{HS}$ satisfies

$$\Pr\left[X_N^{HS}(t+s) \in \tau_{\tilde{\mu}}^{-1}(Y_0) | X_N^{HS}(t) \in \tau_{\tilde{\mu}}^{-1}(Y)\right] > 0. \tag{25}$$

Let N_c denote the number of individuals in the market which play according to $f_{T^*}^c$, let N_d^p denote the number of individuals in the market that hold pessimistic belief and play according to $f_{T^*}^d$ and let N_d^o denote the number of individuals in the market that hold optimistic belief and play according to $f_{T^*}^d$. Consider any distribution of beliefs which is different from Y_0. Now we describe a particular sequence of matches that leads the process $\mathcal{X}_N$ to reach Y_0 and has positive probability. In the first round, all individuals from the classes N_c, N_d^p and N_d^o are matched to partners of the same class (the following argument also holds in case that two of these sets do not contain an even number of individuals, because then the remaining two individuals can be matched until they have the same belief). After T^*+1 rounds, all the individuals from N_d^o except for two are matched to members of the same class until they enter N_c and finally get matched. The two remaining individuals of class N_d^o are repeatedly matched to two individuals from a divorced long-term relationship. Meanwhile, all members from N_d^p are matched to each other and no other couples are separated. Hence, there are always exactly two individuals in N_c and two individuals in N_d^o and these pairs are matched to each other until the pair in N_c enters the set N_d^p. After this, they are matched to each other and a new couple gets divorced to be matched with the two individuals in N_d^o. This procedure is carried on until all matched individuals have moved to N_d^p. Finally, each of the individuals in N_d^p is matched to either one of the two individuals of N_d^o and if $\underline{k} \geq 2$ and $N \geq 2n+2$ it follows that all individuals are members of the set N_d^p and $\mathcal{X}_N$ reaches Y_0 after some finite time. This sequence of matches can be executed in finite time, is time independent and has positive probability which proves statement (25).

Therefore, any state $Y^{HS} \in \Delta(HS(n)^{T^*+2})$ which satisfies $\tau_{\tilde{\mu}}(Y^{HS}) \neq Y_0$ is a transient state of $\mathcal{X}_N^{HS}$ (which means that $\mathcal{X}_N^{HS}$ returns to Y^{HS} with some probability strictly smaller than one, once it leaves Y^{HS}). It it a well known fact from the theory about Markov chains that the limit distribution assigns zero probability to each transient state. Hence, $\mathcal{X}_N$ reaches Y_0 with certainty, which proves the statement in case $\underline{k} \geq 2$.

If $\underline{k} = 1$ and $X_N(t) \neq Y_0$ for some t, then at least one individual cooperates and therefore $X_N(t+1) \neq Y_0$. The claim follows from $X_N(0) \neq Y_0$.
Q.E.D.

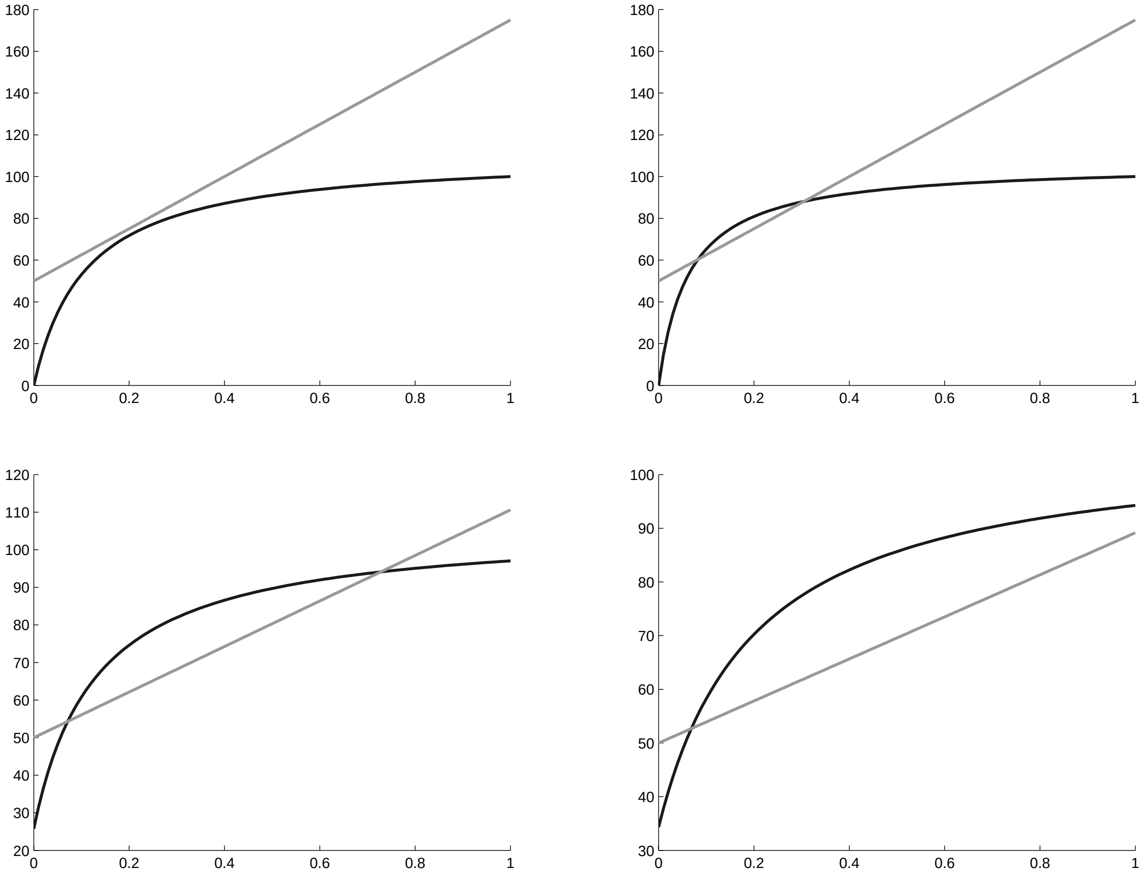


FIGURE I

Caption: [Gray] $E[f^d_{T^*}, \tilde{\mu}]$ **[Black]** $E[f^c_{T^*}, \tilde{\mu}]$ **[Horizontal axis]** $\tilde{\mu}$

[Top-left] Γ_0, $T^* = 0$ **[Top-right]** Γ_1, $T^* = 0$ **[Bottom-left]** Γ_1, $T^* = 1$ **[Bottom-right]** Γ_1, $T^* = 2$

TABLE I: Distribution of initial beliefs for $n \in \{5, 10, 20\}$

$n = 20$	$Y^{1,20}$	$Y^{2,20}$	$Y^{3,20}$	$Y^{4,20}$	$Y^{5,20}$	$Y^{6,20}$	$Y^{7,20}$	$Y^{8,20}$	$Y^{9,20}$	$Y^{10,20}$
$\tilde{\mu} = 0.00$	0.990	0.900	0.330	0.166	0.125	0.048	0.000	0.000	0.125	0.000
$\tilde{\mu} = 0.05$	0.000	0.000	0.340	0.166	0.125	0.048	0.000	0.000	0.125	0.000
$\tilde{\mu} = 0.10$	0.010	0.000	0.330	0.166	0.125	0.048	0.000	0.000	0.125	0.000
$\tilde{\mu} = 0.15$	0.000	0.000	0.000	0.170	0.125	0.048	0.000	0.000	0.125	0.000
$\tilde{\mu} = 0.20$	0.000	0.000	0.000	0.166	0.125	0.048	0.000	0.000	0.000	0.000
$\tilde{\mu} = 0.25$	0.000	0.000	0.000	0.166	0.125	0.048	0.000	0.000	0.000	0.000
$\tilde{\mu} = 0.30$	0.000	0.000	0.000	0.000	0.125	0.047	0.000	0.000	0.000	0.000
$\tilde{\mu} = 0.35$	0.000	0.000	0.000	0.000	0.125	0.047	0.000	0.000	0.000	0.000
$\tilde{\mu} = 0.40$	0.000	0.100	0.000	0.000	0.000	0.047	0.000	0.000	0.000	0.125
$\tilde{\mu} = 0.45$	0.000	0.000	0.000	0.000	0.000	0.047	0.000	0.000	0.000	0.125
$\tilde{\mu} = 0.50$	0.000	0.000	0.000	0.000	0.000	0.047	0.000	0.000	0.000	0.125
$\tilde{\mu} = 0.55$	0.000	0.000	0.000	0.000	0.000	0.047	0.000	0.000	0.000	0.125
$\tilde{\mu} = 0.60$	0.000	0.000	0.000	0.000	0.000	0.047	0.000	0.000	0.000	0.125
$\tilde{\mu} = 0.65$	0.000	0.000	0.000	0.000	0.000	0.047	0.000	0.000	0.000	0.125
$\tilde{\mu} = 0.70$	0.000	0.000	0.000	0.000	0.000	0.048	0.000	0.000	0.000	0.125
$\tilde{\mu} = 0.75$	0.000	0.000	0.000	0.000	0.000	0.048	0.000	0.000	0.000	0.125
$\tilde{\mu} = 0.80$	0.000	0.000	0.000	0.000	0.000	0.048	0.000	0.000	0.000	0.000
$\tilde{\mu} = 0.85$	0.000	0.000	0.000	0.000	0.000	0.048	0.250	0.000	0.125	0.000
$\tilde{\mu} = 0.90$	0.000	0.000	0.000	0.000	0.000	0.048	0.250	0.000	0.125	0.000
$\tilde{\mu} = 0.95$	0.000	0.000	0.000	0.000	0.000	0.048	0.250	0.000	0.125	0.000
$\tilde{\mu} = 1.00$	0.000	0.000	0.000	0.000	0.000	0.048	0.250	1.000	0.125	0.000

$n = 10$	$Y^{1,10}$	$Y^{2,10}$	$Y^{3,10}$	$Y^{4,10}$	$Y^{5,10}$	$Y^{6,10}$	$Y^{7,10}$	$Y^{8,10}$
$\tilde{\mu} = 0.00$	0.990	0.330	0.250	0.090	0.000	0.000	0.250	0.000
$\tilde{\mu} = 0.10$	0.010	0.340	0.250	0.090	0.000	0.000	0.250	0.000
$\tilde{\mu} = 0.20$	0.000	0.330	0.250	0.090	0.000	0.000	0.000	0.000
$\tilde{\mu} = 0.30$	0.000	0.000	0.250	0.090	0.000	0.000	0.000	0.000
$\tilde{\mu} = 0.40$	0.000	0.000	0.000	0.090	0.000	0.000	0.000	0.250
$\tilde{\mu} = 0.50$	0.000	0.000	0.000	0.100	0.000	0.000	0.000	0.250
$\tilde{\mu} = 0.60$	0.000	0.000	0.000	0.090	0.000	0.000	0.000	0.250
$\tilde{\mu} = 0.70$	0.000	0.000	0.000	0.090	0.000	0.000	0.000	0.250
$\tilde{\mu} = 0.80$	0.000	0.000	0.000	0.090	0.000	0.000	0.000	0.000
$\tilde{\mu} = 0.90$	0.000	0.000	0.000	0.090	0.500	0.000	0.250	0.000
$\tilde{\mu} = 1.00$	0.000	0.000	0.000	0.090	0.500	1.000	0.250	0.000

$n = 5$	$Y^{1,5}$	$Y^{2,5}$	$Y^{3,5}$	$Y^{4,5}$	$Y^{5,5}$	$Y^{6,5}$
$\tilde{\mu} = 0.00$	0.990	0.500	0.166	0.000	0.500	0.000
$\tilde{\mu} = 0.20$	0.010	0.500	0.166	0.000	0.000	0.000
$\tilde{\mu} = 0.40$	0.000	0.000	0.166	0.000	0.000	0.500
$\tilde{\mu} = 0.60$	0.000	0.000	0.170	0.000	0.000	0.500
$\tilde{\mu} = 0.80$	0.000	0.000	0.166	0.000	0.000	0.000
$\tilde{\mu} = 1.00$	0.000	0.000	0.166	1.000	0.500	0.000

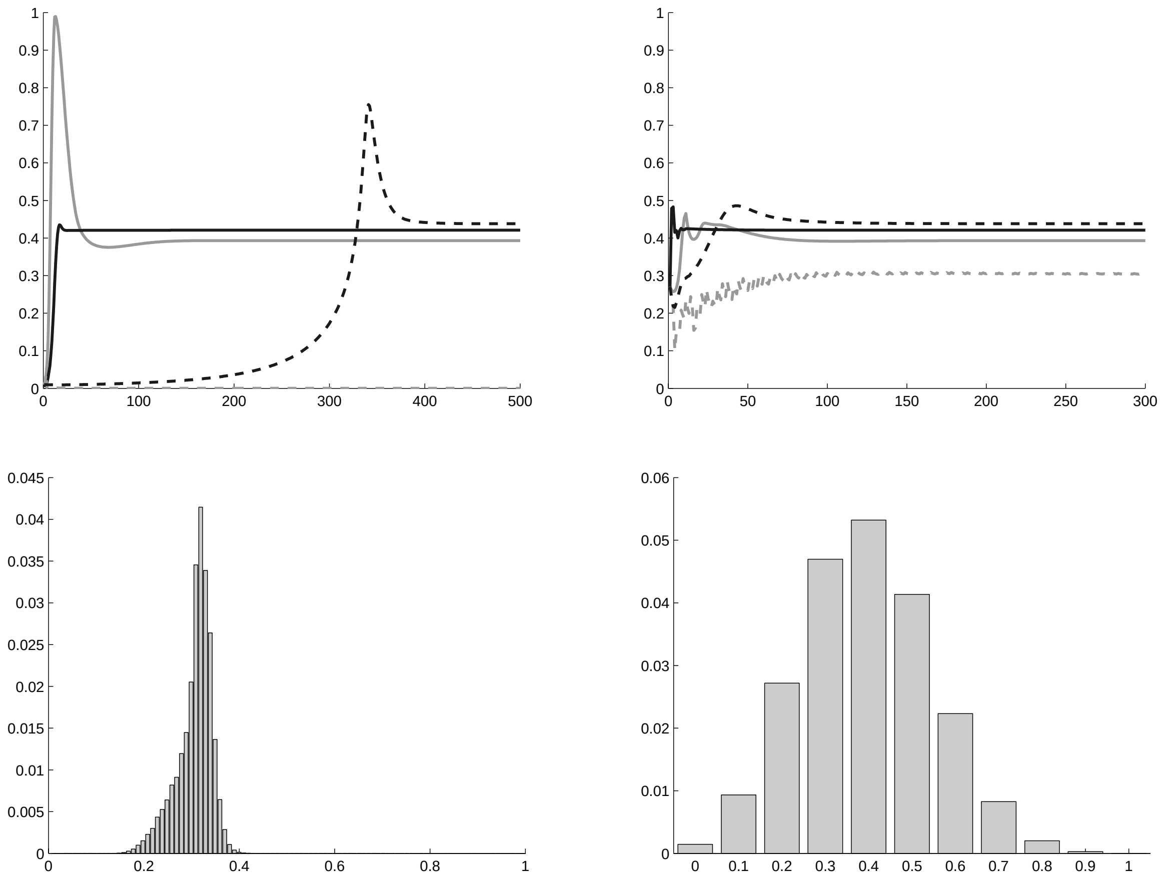


FIGURE II

[TOP]

Caption: [Black-solid] μ under $\tilde{\mu}^{FA}$ **[Gray-solid]** μ under $\tilde{\mu}^{A}$ **[Horizontal axis]** $\tilde{\mu}$

[Top-left] Γ_1, $T^* = 0$, $n = 10$, initial beliefs $Y^{1,10}$ **[Top-right]** Γ_1, $T^* = 0$, $n = 10$, initial beliefs $Y^{4,10}$

[BOTTOM]

Caption: [Horizontal axis] $\tilde{\mu}$ **[Vertical axis]** $y_{\tilde{\mu}}(500)$

Parameters: Γ_1, $n = 10$, $T^* = 0$, initial beliefs are $Y^{4,10}$

[Bottom-left] $\tilde{\mu}^{FP}$ **[Bottom-right]** $\tilde{\mu}^{A}$

TABLE II: Limit points

Γ_1	$n = 5$ [$\underline{k} = 1, \bar{k} = 1$]	$n = 10$ [$\underline{k} = 1, \bar{k} = 3$]	$n = 20$ [$\underline{k} = 2, \bar{k} = 6$]
$\tilde{\mu}^{FP}$	0.305	0.305	0.000 (*), 0.305 (**)
$\tilde{\mu}^{A}$	0.331	0.393	0.364

Γ_2	$n = 5$ [$\underline{k} = 1, \bar{k} = 3$]	$n = 10$ [$\underline{k} = 1, \bar{k} = 7$]	$n = 20$ [$\underline{k} = 1, \bar{k} = 15$]
$\tilde{\mu}^{FP}$	0.780	0.780	0.780
$\tilde{\mu}^{A}$	0.620	0.677	0.715

Γ_3	$n = 5$ [$\underline{k} = 1, \bar{k} = 2$]	$n = 10$ [$\underline{k} = 1, \bar{k} = 4$]	$n = 20$ [$\underline{k} = 1, \bar{k} = 9$]
$\tilde{\mu}^{FP}$	0.467	0.467	0.467
$\tilde{\mu}^{A}$	0.488	0.464	0.481

Γ_4	$n = 5$ [$\underline{k} = 1, \bar{k} = 4$]	$n = 10$ [$\underline{k} = 1, \bar{k} = 8$]	$n = 20$ [$\underline{k} = 1, \bar{k} = 17$]
$\tilde{\mu}^{FP}$	0.896	0.896	0.896
$\tilde{\mu}^{A}$	0.755	0.753	0.810

(*) Under $Y^{1,20}$, (**) Under all other initial beliefs.

References

[1] Aoyagi, M. and Fréchette, G.R., 2003. Collusion in Repeated Games with Imperfect Public Monitoring. Harvard Business School, mimeo.

[2] Benaïm, M. and Weibull, J., 2003. Deterministic Approximation of Stochastic Evolution in Games. *Econometrica* 71, 873-903.

[3] Boylan, R., 1995. Continuous Approximation of Dynamical Systems with Randomly Matched Individuals. *Journal of Economic Theory* 66, 615-625.

[4] Brown, G., 1951. Iterative solution of games by fictitious play. In Koopmanns, T. C. (Ed.). *Activity Analysis of Production and Allocation.* Wiley, New York.

[5] Brosig, J., 2002. Identifying cooperative behavior: some experimental results in a prisoner's dilemma game. *Journal of Economic Behavior and Organization* 47, 275-290.

[6] Corradi, V. and Sarin, R., 1999. Continuous Approximation of Stochastic Evolutionary Game Dynamics. *Journal of Economic Theory*, forthcoming.

[7] Dal Bó, P. and Fréchette, G. R., 2007. The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence. Working Paper, Brown University.

[8] Datta, S., 1993. Building Trust. Mimeo, London School of Economics.

[9] Duffy, J. and Ochs, J., 2003. Cooperative Behavior and the Frequency of Social Interaction. Mimeo, University of Pittsburgh.

[10] Ellison, G., 1994. Cooperation in the prisoner's Dilemma with Anonymous Random Matching. *Review of Economic Studies* 61, 567-588.

[11] Fudenberg, D. and Kreps, D., 1993. Learning Mixed Equilibria. *Games and Economic Behavior* 5, 320-367.

[12] Fudenberg, D. and Levine, D.K., 1998. The theory of learning in games. *MIT Press, Cambridge, Massachusetts.*

[13] Ghosh, P. and Ray, D., 1996. Cooperation in Community Interaction without Information Flows. *Review of Economic Studies* 63, 491-519.

[14] Hopkins, E., 1999. Learning, Matching, and Aggregation. *Games and Economic Behavior* 26, 79-110.

[15] Kandori, M., Mailath, G. and Rob, R., 1993. Learning, Mutation, and Long Run Equilibria in Games. *Econometrica* 61, 29-56.

[16] Kandori, M., 1992. Social Norms and Community Enforcement. *Review of Economic Studies* 59, 63-80.

[17] Mailath, G.J. and Samuelson, L., 2006. Repeated Games and Reputations. *Oxford University Press*, New York.

[18] Murnighan, J.K. and Roth, A.E., 1983. Expecting Continued Play in Prisoner's Dilemma Games .*Journal of Conflict Resolution* 27, 279-300.

[19] Roth, A.E. and Murnighan, J.K., 1978. Equilibrium Behavior and Repeated Play of the Prisoner's Dilemma. *Journal of Mathematical Psychology* 17, 189-197.

Alexander Gaigl[†]

Learning under Incomplete Information with Applications to Auctions[*]

University of Mannheim
December 2009

Abstract

This paper studies learning under incomplete information. Players sometimes reveal their types and actions, which enables other players to learn their strategies (i.e. the mapping from types to actions). Convergence results in terms of best-reply graphs are derived for general $n-$player games. Moreover, these results are applied to discrete first- and second price auctions with either private or common valuations. Attention is restricted to models with two bidders and a uniform distribution of types. It is shown that small perturbations of players' beliefs may lead to unique outcomes of the learning process. These outcomes correspond to equilibria that are of particular interest in the different continuous benchmark models (in most cases the symmetric outcomes are selected). It is argued that these results justify some of the *ad hoc* approaches on equilibrium selection taken by classical auction theory (such as the restriction to symmetric equilibria in second-price auctions with common values).

JEL Classification: C72; C73; D44; D83

Keywords: Auctions, learning, incomplete information, fictitious play, bounded rationality

[†]Contacts: Department of Economics, University of Mannheim, D-68131 Mannheim, Germany, agaigl[at]rumms.uni-mannheim.de, http://www.vwl.uni-mannheim.de/gk/_gaigl.

[*]I thank Philippe Jehiel, Jörg Oechssler and the participants of the CDSE seminar in Mannheim for helpful suggestions. All remaining errors are my own. Moreover, I am grateful to the Economics Department at UCL for their hospitality.

1 Introduction

The most common solution concept employed in game theory is the concept of Nash equilibrium. This approach assumes that players play best responses to their beliefs about their opponents' behavior and that these beliefs are consistent with actual play of their opponents. However, in many situations is in no way clear how players may obtain such consistent beliefs. In particular, if players lack common knowledge of important characteristics of the game, such as common knowledge of rationality, equilibrium approaches are hard to justify. Recently, game theorists have addressed this problem by assuming that players obtain consistent beliefs via learning and repeated interaction. One of the most prominent learning rule is *fictitious play*, which was introduced by Brown (1951). An excellent account of the literature on learning can be found in either Weibull (1997) or Fudenberg and Levine (1998).

While in standard learning models, players are assumed to have fixed types in each encounter, I extend the analysis to the area of Bayesian games in which players do not know each other's types. I show how a learning approach may justify the concept of Bayesian Nash equilibrium in such a setting. More specifically, I assume that a static Bayesian game is played repeatedly and that players form beliefs about opponents' future play from their behavior in the past. It is assumed that the distribution of types is common knowledge and that each players' type (as well as her action) is revealed to all the other players with a certain probability after each round. Moreover, there exist public records about the behavior of each type of each player in the past. Players are not assumed to be highly sophisticated in the sense that they carry out complicated strategic considerations. Instead, they play simple myopic best responses to beliefs implied by the public records. I show that, even with these low requirements on players' rationality and knowledge, behavior approaches Nash equilibrium play in the long-run. Results are obtained by using specific properties of best-reply graphs, a technique introduced by Young (1993). The approach taken in Young (1993) is extended to games of incomplete information and several tools are developed which facilitate the analysis of best-reply graphs for such games.

The second part of this paper applies the learning approach under incomplete information to auctions, in particular to first- and second price auction with either private or interdependent valuations. These auction formats have been extensively studied over the past decades. Most of the literature has focused on the Bayesian equilibria of the corresponding one-shot games, see Krishna (2002) for an extensive account of the theory. There exist very few approaches that apply models of learning and evolution to auctions. Most notably, Hon-Snir, Monderer and Sela (1998) study a model in which bidders' valuations (i.e. their types) are determined before the first round. Then, a fixed set of players repeatedly plays a first-price auction and players observe the profile of bids at the end of each auction. Since players interact with a fixed set of opponents whose types are fixed, bidders eventually learn their opponent's types and play converges to the Nash equilibrium of the one-shot auction in which types are common knowledge (the highest valuation bidder wins and pays a price equal to the second-highest valuation). Their approach imitates standard learning models of complete information and is therefore very different from the type of learning considered here.

In order to integrate learning and auction models, the action space as well as the type space have to be discretized. The resulting discrete one-shot auction model turns out to have a variety of equilibria. I show for different auction settings that if small perturbations are introduced on players' beliefs, the set of equilibria reduces to a unique, in most cases symmetric, equilibrium and I show that the learning process almost surely converges to these equilibria. Note that, in contrast to Kandori, Mailath and Rob (1993) and Young (1993), equilibrium selection here is not directly a result of learning itself but rather relies on these perturbations of beliefs.

Classical auction theory sometimes takes *ad hoc* approaches when it comes to equilibrium selection. For example, the second-price auction with common values is known to have many asymmetric equilibria, but theorists usually focus on the symmetric ones for the simple reason that they are more tractable. The equilibrium selection results obtained here can be very useful in justifying some of these approaches. Moreover, some auction models reject the symmetric equilibria and predict a unique asymmetric outcome, such as the second-price auction model by Klemperer (1998). He shows that if one introduces (arbitrarily) small asymmetries between the players, the unique equilibrium of the auction game is such that the favored player always wins. I show that in the presence of small asymmetries, the learning process selects this extreme outcome. In other words, the equilibrium of Klemperer (1998) can be viewed as the outcome of learning.

My approach to learning under incomplete information is similar to the one adopted in Jensen, Sloth and Whitta-Jacobsen (2005). They, however, focus rather on equilibrium selection and efficiency by using the concept of stochastic stability. Moreover, they make restricting assumptions on the way beliefs are formed and they only consider two player games. Other approaches focus on equilibrium selection in dynamic games under incomplete information. Examples of equilibrium selection in signalling games include Canning (1992), Nöldeke and Samuelson (1997) and Jacobsen, Jensen and Sloth (2001). Finally, see Agastya (2004) for an application of the concept of stochastic stability to double auctions with complete information.

The paper is structured as follows. In section 2, I introduce a model of learning under incomplete information in a n-player game and derive some general results with respect to convergence. Section 3 introduces a basic auction model which is studied in detail in sections 4 and 5. In section 4, convergence of the learning process in case of private and common values is studied for a second-price auction. Section 5 analyzes first-price auctions. Since no general results can be obtained in this case, convergence is studied for two examples (one of private and one of common values).

2 An Incomplete Information Learning Model

Let Γ be a finite n-person game of incomplete information, and let N be the set of players. For each player $i \in N$, let $\mathcal{S}_i$ denote the finite set of player i's types (or signals) and let $\mathcal{B}_i$ denote the set of actions available to player i. Let f_i be the distribution of types. Assume that all f_i are common knowledge. For any profile of actions $b = (b_1, b_2, ..., b_n)$ and any profile of types $s = (s_1, s_2, ..., s_n)$, player i's utility is given by the von Neumann utility function $u_i(b, s)$.

Let $t = 0, 1, 2, ...$ denote successive time periods. Γ is played once in each period by the n players. In every period t, player $i \in N$ observes her private signal $s_i(t)$, where signals are drawn independently over time and players. Thereafter, player i chooses an action $b_i(t)$ from her action space according to a rule described below. At the end of each period t, player i's data set $(b_i(t), s_i(t))$ is revealed to all the other players with probability $\gamma > 0$. The assumption that types are revealed to other players makes particular sense in situations in which payoffs have enough structure so that players can infer their opponents' types from their own payoffs. Note that players do not have a strategic incentive to hide their types in this model, because new types are drawn independently in each round and the distribution of types is common knowledge.

Let me describe the decision-making rule. Each player i forms beliefs about her opponents' future behavior based on their action choices in past periods. Player i's actions are best responses to these beliefs. I assume that players form beliefs about each opponent's entire *strategy* (i.e. about the mapping from types to actions) and not just about aggregate play. Specifically, for every player $i \in N$ and for all types $s \in \mathcal{S}$, there exists a (public) record (or *history*) $h_i^s = (h_i^{s,1}, h_i^{s,2}, ..., h_i^{s,\ell})$ which contains her actions in the $\ell \geq 1$ most recent periods in which she has been of type s, and in which her type-strategy profile was publicly observable. Let $h_i = (h_i^1, h_i^2, ..., h_i^{T_i})$ denote player i's personal history and let $h = (h_1, h_2, ..., h_n)$ denote the entire *history of play*. As usual, $h_{-i} = (h_1, h_2, ..., h_{i-1}, h_{i+1}, ..., h_n)$ denotes the history of play of all players other than i. Let the corresponding collection of histories be denoted by $\mathbb{H}_i$, $\mathbb{H}$ and $\mathbb{H}_{-i}$. Note that this finite-history approach is similar to fictitious play. Under fictitious play, however, agents base their decisions on the *entire* history of actions by other agents. Here, I assume that agents only consider limited information from the recent past. This seems to be more realistic[1] than fictitious play.

The belief of players $j \neq i$ about the behavior of player i's s-type is given by a *belief function* $\beta_i^s : \mathbb{H}_i \mapsto \Delta(\mathcal{B})$ (for any finite set X, Δ(X) denotes the collection of probability distributions over X). Note that this specification implicitly contains the assumptions that belief functions are time-independent and that the belief about player i's behavior is independent of other player's previous play. However, beliefs β_i^s may well depend on past observations of other (neighboring) types of the same player i. Furthermore, the belief functions might incorporate some sort of weighting in the sense that more recent observations have a bigger impact on current beliefs than old observations.

For any history of play $h \in \mathbb{H}$, let $BR_i^s(h) \subset \mathcal{B}$ be the set of best responses of player i's s-type to the family of beliefs $\{\beta_j^s(h_j)\}_{j \neq i}$. More specifically, $b \in BR_i^s(h)$ if and only if

$$b \in \arg\max_{b' \in \mathcal{B}} \sum_{s_{-i} \in \mathcal{S}_{-i}} \sum_{b_{-i} \in \mathcal{B}_{-i}} \left[\prod_{j \neq i} f_j(s_j) \beta_j^{s_j}(h_j)(b_j) \right] u_i(b', b_{-i}, s, s_{-i}),$$

where s_{-i}, b_{-i}, $\mathcal{S}_{-i}$ and $\mathcal{B}_{-i}$ are defined in the usual way.[2] If player i is of type s, she takes an action from the set BR_i^s. In case BR_i^s is not a singleton, player i's action is determined by a tie-breaking rule that assigns positive probability to all elements in BR_i^s. Let me denote the family $\{\beta_i^s\}_s$ of beliefs about player i by β_i.

Assume that the first actions are taken in round 1 and that before round 1, there exists an initial history of actions $h_0 \in \mathbb{H}$. The initial history might, for example, result from a (sufficiently long) initial experimentation phase. The sequence of histories $\{h(t)\}$ with initial state $h(0) = h_0$ forms a Markov chain on the state space $\mathbb{H}$. Let me refer to $\{h(t)\}$ simply as the *learning process*. The transition probabilities follow from the distribution of types, the tie-breaking rule and γ.

In order to introduce some more stochastic variability into the learning process let me assume that the probability γ that players reveal their private

[1] or 'less fictitious', as Young (1993) puts it

[2] $s_{-i} = (s_1, s_2, ..., s_{i-1}, s_{i+1}, ..., s_n)$, $b_{-i} = (b_1, b_2, ..., b_{i-1}, b_{i+1}, ..., b_n)$, $\mathcal{S}_{-i} = \mathcal{S}_1 \times \mathcal{S}_2 \times ... \times \mathcal{S}_{i-1} \times \mathcal{S}_{i+1} \times ... \times \mathcal{S}_n$ and $\mathcal{B}_{-i} = \mathcal{B}_1 \times \mathcal{B}_2 \times ... \times \mathcal{B}_{i-1} \times \mathcal{B}_{i+1} \times ... \times \mathcal{B}_n$.

information is strictly smaller than one. This additional degree of variability will turn out to be crucial in showing convergence of the learning process. If not stated otherwise, this assumption is maintained throughout the paper.

Assumption 1. $\gamma < 1$ *(i.e. players do not always reveal their private information)*

Define the *best-reply graph* G_Γ of Γ as follows. Each vertex b of G_Γ is a profile of bids that assigns to each type s of each player i some bid $b_i^s \in \mathcal{B}$. Let $\mathbb{B}$ denote the collection of all vertices. It is straightforward to extend the definition of the best response correspondence to the space of vertices. For any vertex $b \in \mathbb{B}$, the set $BR_i^s(b)$ is defined by the set $BR_i^s(h)$, where h is such that $h_i^s = (b_i^s, b_i^s, ..., b_i^s)$, $\forall i \in N$, $s \in \mathcal{S}$. For any two vertices b and $\widehat{b}$, G_Γ exhibits a directed edge from b to $\widehat{b}$ if and only if $b \neq \widehat{b}$ and there is exactly one type s of one player i such that $\widehat{b}_i^s \in BR_i^s(b)$ and $b_j^{s'} = \widehat{b}_j^{s'}$, for all $(j, s') \neq (i, s)$. A *sink* b^* of G_Γ is defined as a vertex from which there exists no directed edge to some other vertex. Let $B^* \subset \mathbb{B} \cup \emptyset$ be the set of all sinks of G_Γ. Note that any sink b^* satisfies $BR_i^s(b^*) = \{b_i^{*,s}\}$ for every type s of every player i. From this, it is easy to see that every sink corresponds to an absorbing state of the learning process and to a *strict* pure Bayesian Nash equilibrium of Γ. Moreover, Γ is *weakly acyclic* if from any vertex $b \in \mathbb{B}$ of G_Γ, there exists a directed path (i.e. a succession of zero or more directed vertices) to some sink b^*. This concept is due to Following Young (1993).

Proposition 1. *If Γ is weakly acyclic and assumption 1 holds, then the learning process converges almost surely to an absorbing state. In case each player has a (weakly) dominant strategy for at least one of her types, then assumption 1 is not required.*

Proof. Let H denote the set of all states $h \in \mathbb{H}$ for which there exists some vertex $b \in \mathbb{B}$ such that $h_i^s = (b_i^s, b_i^s, ..., b_i^s)$, $\forall i \in N$ and $s \in \mathcal{S}$. Moreover, let $H^* \subset H$ be the subset of states $h \in \mathbb{H}$ for which there exists some vertex $b^* \in \mathbb{B}^*$ such that $h_i^s = (b_i^{*,s}, b_i^{*,s}, ..., b_i^{*,s})$, $\forall i \in N$ and $s \in \mathcal{S}$. In the following, I show that there exists a finite number $\overline{m}$ and a strictly positive probability $\overline{p}$ such that from any state $h \in \mathbb{H}$, the probability that the learning process reaches H^* in at most $\overline{m}$ periods is at least $\overline{p}$. Therefore, the probability of not reaching H^* after at least $r\overline{m}$ periods is at most $(1 - \overline{p})^r$, which goes to zero as r goes to infinity.

Let $T = \sum_{i=1}^{N} T_i$ and consider any state $h_1 \in \mathbb{H}$. I show that there exists a sequence of ℓT transitions of the learning process (starting at period 1) that satisfies $h(1) = h_1$ and $h(\ell T) \in H$. The sequence is subdivided into n subsequences, where each of them is denoted by i and has length ℓT_i.

Assume that in each subsequence, player i is of each type $s \in S_i$ for ℓ rounds. Assumption 1 assures that, with positive probability, all beliefs β_j about players $j \neq i$ are constant in the entire subsequence i and therefore all sets of best responses of player i, BR_i^s, remain constant with positive probability. It is easy to see that in case all players have dominant strategies, beliefs about players $j \neq i$ are constant with positive probability, even if assumption 1 is not satisfied. Therefore, with positive probability, each s-type of player i takes constant actions throughout subsequence i. Assume that agent i's private information is revealed in each period of subsequence i (which has positive probability according to assumption 1). It follows that in period ℓT, all vectors h_i^s contain ℓ identical observations, which implies that $h(\ell T) \in H$ is satisfied. Let $p(h_1)$ denote the probability of the sequence and define $p_1 = \min_{h_1 \in \mathbb{H}} p(h_1)$.

Let $b \in \mathbb{B}$ denote the vertex that corresponds to the end-point $h(\ell T)$ of the above sequence and fix some sink $b^* \in B^*$, where h^* denotes the corresponding state of $\{h_t\}$. Let $m(b)$ be the length of the shortest directed path from b to b^* (which exists because G_Γ is weakly acyclic). By constructing a sequence in the spirit of the first part of the proof, one finds that there exists a transition path of length $\ell m(b)$ from h to h^*. Let $p(b)$ be the corresponding transition probability. By putting together the two sequences, it follows that $\{h(t)\}$ reaches H^* from any initial condition in at most $\overline{m} = \ell(T + \max_b m(b))$ periods with probability of at least $\overline{p} = p_1 \min_b p(b) > 0$. Both, $\overline{m}$ and $\overline{p}$ exist since the state space $\mathbb{H}$ is finite. □

A vertex b of G_Γ is *undominated* if there does not exist a player $i \in N$ and a type $s \in S_i$ such that b_i^s is weakly dominated, i.e. there exists some action $k \neq b_i^s$ such that for all opponents' types and actions, k is always weakly better than b_i^s. It is easy to see that for any dominated vertex b', there exists a directed path of G_Γ to an undominated vertex. Therefore, if the sub-graph of G_Γ which only contains undominated vertices is weakly acyclic, it follows that G_Γ is weakly acyclic as well. This property will turn out to be very useful in the auction context studied in sections 4 and 5.

For games Γ in which players' types and actions can be ordered, lemma 1 provides a way of further reducing the size of the best-reply graph. An additional condition on players' utility functions is needed, which requires that higher types have a higher incentive to take higher actions. Assume that $\mathcal{B}, \mathcal{S}_i \subseteq \mathbb{N}^0$. For any vertex $b \in \mathbb{B}$, denote the expected utility of player i's s-type from $k \in \mathcal{B}_i$ by $u_i^s(k) = \mathrm{E}\left[u_i(k, b_{-i}, s, s_{-i})\right]$, where expectations are taken with respect to f_{-i}.

Condition 1. *For all players $i \in N$, types $s, s' \in \mathcal{S}_i$ and actions $k, k' \in \mathcal{B}_i$, Γ satisfies $u_i^s(k') - u_i^s(k) \leq u_i^{s'}(k') - u_i^{s'}(k)$ whenever $s' \geq s$ and $k' \geq k$.*

Lemma 1 ensures that higher types never (i.e. for any distribution of opponents' types and actions) take actions below the actions of all the lower types. This implies that G_Γ is weakly acyclic if from all vertices b in which b_i^s is weakly increasing, there exists a directed path to some sink. Since condition 1 is satisfied in most standard auction frameworks (see lemma 2), lemma 1 will turn out to be an important tool for showing convergence in the next two sections.

Lemma 1. *If there exists a directed path from any undominated vertex b to some sink b^*, then Γ is weakly acyclic. Moreover, if Γ satisfies condition 1 and there exists a directed path from any undominated vertex b with $b_i^s \leq b_i^{s'}$ for all $s, s' \in S_i$, $s \leq s'$, to some sink b^*, then Γ is weakly acyclic.*

Proof. Let $B' \subset \mathbb{B}$ denote the subset of all vertices $b \in \mathbb{B}$ that satisfy $b_i^s \leq b_i^{s'}$ for all $s, s' \in S_i$, where $s \leq s'$. Consider any vertex $b \in \mathbb{B}$. In the same way as in the proof of proposition 1, consider a sequence consisting of n subsequences in which each h_i^s is updated exactly ℓ times. With positive probability, only undominated actions are played in each of the ℓT periods and therefore, the final vertex is undominated. In order to show that there exists a sequence with end point in B', it remains to be shown that for all $i \in N$ and $s, s' \in \mathcal{S}_i$, $s < s'$, $k' \in BR_i^s(b)$ implies that $u_i^{s'}(k) \leq u_i^{s'}(k')$ for all $k < k'$. Assume (by contradiction) that the claim does not hold, i.e. there exists some $k < k'$ with $u_i^{s'}(k) > u_i^{s'}(k')$. By assumption, type s (weakly) prefers k' to k and type s' strictly prefers k to k'. Hence,

$$u_i^s(k') - u_i^s(k) \geq 0$$
$$u_i^{s'}(k) - u_i^{s'}(k') > 0.$$

By adding the two inequalities, one obtains a contradiction to condition 1. □

3 A Discrete Auction Model

In this section, I provide a discrete auction game Γ^A which is a special case of the general n-person game introduced in the previous section. There are $n = 2$ identical bidders that play the auction repeatedly. Before each round of play, each bidder $i \in N = \{1, 2\}$ observes a signal s_i from the set $\mathcal{S}_i = \mathcal{S} = \{0, 1, ..., T\}$ and proposes bids from the set $\mathcal{B}_i = \mathcal{B} = \{0, 1, ...M\}$, where M is some sufficiently large number. Assume that belief functions are identical for both bidders, i.e. $\beta_1^s \equiv \beta_2^s \equiv \beta^s$. Signals are assumed to be independently and equally distributed and the probability that bidder i is of type $s_i \in \mathcal{S}$ is denoted by p_i. For any given profile of signals $s \in \mathcal{S}^2$, bidder i's valuation for the object is given by $v_i(s_i, s_{-i}) = s_i + \alpha_i s_{-i}$, where $\alpha_i \in [0, 1]$. This model of interdependent valuations includes the cases of private values ($\alpha_i = 0$) and common values ($\alpha_i = 1$) as special cases. The seller has a zero valuation for the object. Except for the realization of the opponent's signal, all components of the models are common knowledge among the players. I consider the two most common auction formats, namely the first-price and second-price sealed bid auctions. I assume that bidders are risk-neutral and that each bidder gets the object with probability $\frac{1}{2}$ if both bidders propose the same bid. For ease of presentation, I assume that in case of a tie, each bidder receives her expected utility. This assumption is innocuous since bidders are risk-neutral. Hence, bidder i's utility in the first-price auction is given by

$$u_i^I(b_i, b_{-i}, s_i, s_{-i}) = \begin{cases} v(s_i, s_{-i}) - b_i & \text{if } b_i > b_{-i} \\ \frac{1}{2}(v(s_i, s_{-i}) - b_i) & \text{if } b_i = b_{-i} \\ 0 & \text{otherwise.} \end{cases}$$

Likewise, bidder i's utility in the second-price auction is given by

$$u_i^{II}(b_i, b_{-i}, s_i, s_{-i}) = \begin{cases} v(s_i, s_{-i}) - b_{-i} & \text{if } b_i > b_{-i} \\ \frac{1}{2}(v(s_i, s_{-i}) - b_{-i}) & \text{if } b_i = b_{-i} \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 2. *The auction game Γ^A satisfies condition 1.*

Proof. It needs to be shown that for all $i \in N$, $s, s' \in \mathcal{S}_i$ and $k, k' \in \mathcal{B}_i$, Γ^A satisfies $u_i^s(k') - u_i^s(k) \leq u_i^{s'}(k') - u_i^{s'}(k)$ whenever $s' \geq s$ and $k' \geq k$. Let me drop the explicit dependence of h_{-i} and let p_{-i}^b denote the probability that bidder i bids b (i.e. $p_{-i}^b = \sum_{s \in S} f_s \beta^s(h_{-i})(b)$). Bidder i's expected utility conditional on her type s and bid k can be written as follows.

$$u_i^s(k) = s\left[\frac{1}{2}p_{-i}^k + \sum_{m=0}^{k-1} p_{-i}^m\right] + g_2(k) = s g_1(k) + g_2(k),$$

where $g_2(\cdot)$ is an arbitrary function that is independent of s. Hence, $u_i^s(k') - u_i^s(k) - u_i^{s'}(k') + u_i^{s'}(k) = (s - s')(g_1(k') - g_1(k)) \leq 0$.

□

4 Learning in a Second-price Auction

4.1 Common Values and Locally Perturbed Beliefs

It is well known that in a continuous second-price auction with common values, there exists one symmetric equilibrium and an infinite amount of asymmetric equilibria, see Milgrom (1981) or Klemperer (1998). It is, for example, an equilibrium for player 1 to bid ks_1, and for player 2 to bid $\frac{k-1}{k}s_2$, where $k \in \mathbb{N}$ and $k \geq 2$ (the symmetric equilibrium is obtained by setting $k = 2$). In particular, there is a symmetric equilibrium in which both bidders follow the strategies $b_i(s_i) = 2s_i$. I show that learning in the discrete auction model converges almost surely to a state that corresponds to the symmetric equilibrium.

In this section, I assume that bidders form beliefs about the behavior of a specific type of their opponent exclusively from observations of this particular type's past actions, i.e. beliefs are not influenced by past behavior of neighboring types. Moreover, it is assumed that belief functions never assign probability one to some specific bid. Instead, bidders' reasoning incorporates the possibility that the opponent's future action may slightly depart from what she did in the past. Formally, the assumption looks as follows.

Assumption 2. *For any given player i and type s: whenever $h_i^s = (\bar{b}, \bar{b}, ..., \bar{b})$, $\beta_i^s(h)$ assigns probability $1 - 2\varepsilon$ to $\bar{b}$ and probability $\varepsilon > 0$ to both $\bar{b} - 1$ and $\bar{b} + 1$ (in case $\bar{b} = 0$ or $\bar{b} = T$, $\beta_i^s(h)$ assigns probability $1 - \varepsilon$ to $\bar{b}$ and probability ε to 1, resp. $T - 1$).*

This assumption is necessary because there exist many equilibria (symmetric and asymmetric ones) in the discrete auction model. It is easy to check that all such equilibria are *weak*, in the sense that there exists at least one type of one player who is indifferent between two alternatives. This implies that if beliefs are formed by attributing to each player's type her past average behavior, the learning process never settles down and therefore clear predictions about its limit behavior are impossible to make. By adding assumption 2, convergence can be obtained. This is the case because learning may settle down due to the fact that all types of both players have unique best responses.

Let $\widehat{h}$ denote the state h of the learning process that satisfies $h_i^s = (2s, 2s, ..., 2s)$ for all players i and types s, i.e. $\widehat{h}$ corresponds to the symmetric equilibrium. Lemma 3 shows that $\widehat{h}$ is an absorbing state of the learning process under the class of belief functions defined in this section. This result relies on the fact that beliefs are perturbed. In case there are no such perturbations ($\varepsilon = 0$), any type $s > 0$ would be indifferent between bidding

$2s-1$, $2s$ or $2s+1$. If ε is positive however, it is easy to see that $2s$ is a unique best response for player i's s-type.

Lemma 3. *Under assumption 2, $\widehat{h}$ is an absorbing state of the learning process (i.e. $BR_i^s = \{2s\}$ for all $i \in N$ and $s \in \mathcal{S}$).*

Proof. For any given history of bidder i, let p_i^b denote the (unconditional) probability that player i bids b and let A_i^b denote the expected type of player i, conditional on player i bidding b and provided that $p_i^b > 0$. In case $p_i^b = 0$, I define $A_i^b = 0$. Define $\pi_i^s(b \to b') = u_i^s(b') - u_i^s(b)$. For all $k \in \mathcal{B}$, one can write

$$\pi_i^s(k \to k+1) = \frac{1}{2} p_{-i}^k \left[A_i^k + s - k\right] + \frac{1}{2} p_{-i}^{k+1} \left[A_i^{k+1} + s - k - 1\right].$$

$A_i^k = \frac{k}{2}$ is satisfied in case k is even and in case k is odd, one obtains $A_i^k = (f_a a + f_b b)/(f_a + f_b)$, where $a = \frac{k-1}{2}$ and $b = \frac{k+1}{2}$. Hence, $A_i^k \in (\frac{k-1}{2}\frac{k+1}{2})$ is satisfied for all $k \leq 2T$. For all $s \in \mathcal{S}$, it follows that $\pi_i^s(k \to k+1)$ is *strictly* positive for all $k \leq 2s-1$, *strictly* negative for $k = 2s$ and non-positive for $k > 2s$. Hence, $b_i^s = 2s$ is the *unique* best response for every type s of every bidder i. □

Note that lemma 3 states that the (symmetric) equilibrium in which everybody bids twice her signal (let me denote it by b^*) is a trembling-hand perfect equilibrium of Γ^A. This is the case because restricting beliefs to interior points corresponds directly to Selten's definition of perturbed games. In this sense, b^* is the unique trembling hand perfect equilibrium, if one restricts the sequence of perturbed games in Selten's definition according to the local trembles considered here.

Proposition 2 demonstrates that the learning process eventually converges to $\widehat{h}$ from any initial condition. This implies that $\widehat{h}$ is the *only* absorbing state of the learning process in the class of belief functions defined by assumption 2. Therefore it is worth noting that the equilibrium selection argument contained in proposition 2 is not directly a result of learning itself but rather a result of the specification of the belief function. Moreover, the proposition assumes that the distribution of types is uniform in order to avoid technical difficulties. The proof shows that Γ^A is weakly acyclic with respect to any belief function of the class defined by assumption 2. Therefore, proposition 1 and lemma 1 together imply the result.

Proposition 2. *Consider the second-price auction game Γ^A with common values and a uniform distribution of types. If the belief function satisfies assumption 2, the learning process converges almost surely to $\widehat{h}$.*

Proof. Consider the best-reply graph of Γ^A and let me show that it is weakly acyclic, i.e. that from every vertex, there exists a directed path to a unique sink which satisfies $b_i^s = 2s$, for all i and s. This fact in combination with lemma 1 and lemma 2 imply that it is sufficient to restrict attention to vertices b that satisfy $b_i^s \leq b_i^{s'}$ for all $s, s' \in S_i$, $s \leq s'$. For any given vertex, let p_i^n, A_i^n and $\pi_i^s(m \to n) = u_i(n|s) - u_i(m|s)$ be defined as in the proof of lemma 3. The assertion is established via a set of claims.

Claim 1. *Consider any vertex $b \in \mathcal{B}^{2(T+1)}$. For all $i \in N$ and $s \in \mathcal{S}$, $u_i^s(s) \geq u_i^s(k)$ is satisfied for all $k \in \mathcal{B}$ that satisfy $k < s$ (i.e. with positive probability, type s proposes a bid no smaller than s).*

Proof. Consider any type $s \in \mathcal{S}$ and let $k \in \{0, 1, ..., s-1\}$. If player i of type s raises her bid from k to $k+1$, she gains

$$\pi_i(k \to k+1) = \frac{1}{2}p_{-i}^k \left[A_{-i}^k + s - k\right] + \frac{1}{2}p_{-i}^{k+1}\left[A_{-i}^{k+1} + s - k - 1\right].$$

$\pi_i(k \to k+1) \geq 0$ follows directly from $s \geq k+1$, $A_{-i}^k \geq 0$ and $A_{-i}^{k+1} \geq 0$. □

Claim 2. *Let $i \in N$ and $\widehat{s} \in \mathcal{S} \cup \{T+1\}$, where $\widehat{s} \geq 1$. For all $b \in \mathcal{B}^{2(T+1)}$ that satisfy $b_{-i}^{s'} = 2s' \, \forall s' \in \mathcal{S}$ s.t. $s' \geq \widehat{s}$, it holds that $u_i^s(s + \widehat{s} - 1) \geq u_i^s(k)$ for all $s \in \mathcal{S}$, $s \leq \widehat{s} - 1$, and $k > s + \widehat{s} - 1$ (i.e. with positive probability, any type s below $\widehat{s}$ bids no higher than $s + \widehat{s} - 1$).*

Proof. Consider any type $s \in \mathcal{S}$ of bidder i that satisfies $s \leq \widehat{s} - 1$. Moreover, let $k \in \{0, 1, ..., M - s - \widehat{s}\}$. Bidder i's utility from raising her bid from $s + \widehat{s} - 1 + k$ to $s + \widehat{s} + k$ is given by

$$\begin{aligned}\pi_i^s(s + \widehat{s} - 1 + k \to s + \widehat{s} + k) &= \frac{1}{2}p_{-i}^{s+\widehat{s}-1+k}\left[A_{-i}^{s+\widehat{s}-1+k} - \widehat{s} + 1 - k\right] \\ &\quad + \frac{1}{2}p_{-i}^{s+\widehat{s}+k}\left[A_{-i}^{s+\widehat{s}+k} - \widehat{s} - k\right].\end{aligned}$$

In case $\widehat{s} = T + 1$, $\pi_i^s(s + \widehat{s} - 1 + k \to s + \widehat{s} + k) \leq 0$ follows directly from $A_{-i}^{s+\widehat{s}-1+k} \leq T$ and $A_{-i}^{s+\widehat{s}+k} \leq T$. In case $\widehat{s} \leq T$, for all $s \leq \widehat{s} - 1$ and all $0 \leq n \leq T - \widehat{s}$, b satisfies $b_{-i}^{\widehat{s}+n} > 2\widehat{s} - 1 + n \geq s + \widehat{s} + n$, which implies $A_{-i}^{s+\widehat{s}-1+n} \leq \widehat{s} + n - 1$ (and therefore $A_{-i}^{s+\widehat{s}+n} \leq \widehat{s} + n$). Hence, $\pi_i^s(s + \widehat{s} - 1 + k \to s + \widehat{s} + k) \leq 0$ for all k.

□

Claim 3. *Let $i \in N$ and $\widehat{s} \in \mathcal{S} \cup \{T+1\}$, where $\widehat{s} \geq 2$. For all $b \in \mathcal{B}^{2(T+1)}$ that satisfy $b_i^s = b_{-i}^s = 2s$, $\forall s \in \mathcal{S}$ s.t. $s \geq \widehat{s}$, there exists some $\widehat{b} \in \mathcal{B}^{2(T+1)}$ that is connected to b through a finite number of edges and satisfies $\widehat{b}_i^s = \widehat{b}_{-i}^s = 2s, \forall s \in \mathcal{S}$ s.t. $s \geq \widehat{s} - 1$.*

Proof. See Appendix. □

Let $\widehat{b}$ denote the unique vertex that satisfies $b_i^s = 2s$ for all $i \in N$ and all $s \in \mathcal{S}$ and let me show that the above claims imply that from any vertex there exists a path to $\widehat{b}$. Consider any vertex $b^{(0)} \in \mathcal{B}^{2(T+1)}$ and let me construct a finite sequence of vertices $(b^{(0)}, b^{(1)}, ..., b^{(T)}, b^{(T+1)})$ that leads to $\widehat{b}$ through a finite number of best responses. Apply claim 3 with $\widehat{s} = T + 1$ for both $i \in N$. It follows that there exists a finite number of best responses that leads to some other vertex $b^{(1)}$ which satisfies $b_i^{T,(1)} = b_{-i}^{T,(1)} = 2T$. This, in turn, enables the application of claim 3 with $\widehat{s} = T$ and so on. After T such iterations, claim 2 can be applied for $\widehat{s} = 1$ and both $i \in N$. Hence, the final vertex of this sequence satisfies $b^{(T+1)} = \widehat{b}$. Hence, Γ^A is weakly acyclic and proposition 1 implies the assertion. □

Let me summarize the main intuitions behind this proof. In order to establish the assertion, I relate the best-reply graph to the learning process, which is a Markov chain on $\mathbb{H}^2$. Each vertex of the best-reply graph naturally corresponds to a state of the Markov chain. Moreover, each edge of the best-reply graph corresponds to a sequence of state transitions of the Markov chain. This is the case because assumption 1 assures that the history of bids of one player remains constant for a very long time, which implies a positive probability that a given type of her opponent plays the same best response in many consecutive rounds. Hence, $\widehat{h}$ can be reached with positive probability from any initial state $h^0 \in \mathbb{H}^2$ within a finite number of steps. Moreover, $\widehat{h}$ is an absorbing state of the Markov chain. It is easy to see that the probability of reaching the unique absorbing state approaches one as time goes to infinity.

The following proposition relaxes the assumption of a uniform distribution of types to any strictly positive distribution. The number of types, however, is restricted to two.

Proposition 3. *Consider the second-price auction game Γ^A with common values, $T = 1$ (i.e. two types) and an arbitrary, but strictly positive, distribution of types. If the belief function satisfies assumption 2 (with ε sufficiently small), the learning process converges almost surely to $\widehat{h}$.*

Proof. Let $\widehat{b}$ denote the vertex of the best-reply graph that corresponds to $\widehat{h}$. It is easy to check that any vertex b which satisfies $b_i^0 > 1$ or $b_i^1 > 2$ for some i, is (weakly) dominated according to the definition in section 2. Therefore, according to lemma 1 and lemma 2, it is sufficient to show that there exists a directed path of the best-reply graph from any vertex of the set

$$B' = \{(b_1^0, b_1^1, b_2^0, b_2^1) : b_1^0, b_2^0 \in \{0,1\}; b_1^1, b_2^1 \in \{0,1,2\}; b_1^0 \leq b_1^1, b_2^0 \leq b_2^1\},$$

to some sink. Let $F = \{f_0, f_1\}$ denote the distribution of types, where $f_0 > 0$ and $f_1 > 0$, and let $(k^0, k^1)_i$ denote the subset of vertices $b \in B'$ that satisfy $b_i^0 = k^0$ and $b_i^1 = k^1$. For any vertex in $(0,0)_i$ there exists a directed path to a vertex in $(1,2)_{-i}$, because $\pi_{-i}^0(0 \to 1) = \frac{1-\varepsilon}{2} f_1 - \frac{\varepsilon}{2} f_0 > 0$ and $\pi_{-i}^1(1 \to 2) = \frac{\varepsilon}{2} f_1 > 0$. For any vertex in $(1,1)_i$ there exists a directed path to a vertex in $(0,2)_{-i}$, because $\pi_{-i}^0(0 \to 1) = \frac{\varepsilon}{2} f_1 - \frac{1-2\varepsilon}{2} f_0 < 0$, $\pi_{-i}^1(0 \to 1) = \frac{\varepsilon}{2} f_0 + \frac{\varepsilon+1}{2} f_1 > 0$ and $\pi_{-i}^1(1 \to 2) = \frac{1-2\varepsilon}{2} f_1 - \frac{\varepsilon}{2} f_0 > 0$.

Consider the set of vertices $(0,1)_i$. In this case, $\pi_{-i}^0(0 \to 1) = \frac{\varepsilon}{2}(f_1 - f_0)$, $\pi_{-i}^1(0 \to 1) = \frac{1-\varepsilon}{2} f_0 + \frac{1}{2} f_1 > 0$ and $\pi_{-i}^1(1 \to 2) = \frac{1-2\varepsilon}{2} f_1 > 0$. Hence, if $f_1 \geq f_0$, there exists a directed path to a vertex in $(1,2)_{-i}$ and if $f_1 \leq f_0$, there exists a directed path to a vertex in $(0,2)_{-i}$. Moreover, consider the set of vertices $(1,2)_i$. In this case, $\pi_{-i}^0(0 \to 1) = -\frac{1-2\varepsilon}{2} f_0 < 0$, $\pi_{-i}^1(0 \to 1) = \frac{\varepsilon}{2}(f_0 + f_1) > 0$ and $\pi_{-i}^1(1 \to 2) = \frac{\varepsilon}{2}(f_1 - f_0)$. Hence, if $f_1 \geq f_0$, there exists a directed path to a vertex in $(0,2)_{-i}$ and if $f_1 \leq f_0$, there exists a directed path to a vertex in $(0,1)_{-i}$. By putting together these transitions (separately for both cases $f_1 \geq f_0$ and $f_1 < f_0$), it easily follows that for each vertex $b \in B'$, there exists a directed path to $\widehat{b}$. This completes the proof since $\widehat{b}$ is absorbing according to lemma 3.

□

4.2 Uniform Trembles and Common Values

It is well known that different ways of modeling perturbations and trembles in economic models may result in surprisingly different outcomes, even if perturbations are very small in scale (see, for example, Bergin and Lipman (1996) for the effect of state-dependent mutations on the outcome of evolutionary models). Therefore, I study a different type of perturbations in this section in order to check whether the results obtained in section 4.1 are robust to modifications in the way perturbations are modelled. In particular, I use a model of *uniform trembles* which has been frequently used in the literature on learning and evolution. No analytical results can be obtained in this case, so I will restrict attention to an example which I will implement in a computer simulation.

In order to define the uniform trembles, let $U[0, M']$ denote the uniform distribution on the set $\{0, 1, ..., M'\}$ and let $\widehat{\beta}$ be the (standard) belief function which maps histories to relative frequencies of past play.

Assumption 3. *Beliefs about each player i's s-types are given by* $\beta_i^s = (1-\varepsilon)\widehat{\beta}_i^s + \varepsilon U[0, M']$*, whenever* ε *is arbitrarily small and* $M' < M$ *is sufficiently large (the composition of the two discrete probability distributions is defined in the obvious way).*

Example 1. Consider the second-price auction game Γ^A with a uniform distribution of types, a uniform tie-breaking rule and let $T = 8$. Moreover, let the belief function satisfy assumption 3. Let any $b \in \mathbb{B}$ be denoted by $[b_1, b_2]$ and define

$$\begin{aligned} b^{(1)} &= [(1,3,5,7,8,10,12,14,16),(0,2,4,6,8,9,11,13,15)] \\ b^{(2)} &= [(1,3,5,7,8,9,11,13,15),(0,2,4,6,8,10,12,14,16)] \\ b^{(3)} &= [(0,2,4,6,8,9,11,13,15),(1,3,5,7,8,10,12,14,16)] \\ b^{(4)} &= [(0,2,4,6,8,10,12,14,16),(1,3,5,7,8,9,11,13,15)], \end{aligned}$$

The simulation is carried out in the following way. In simulation step 1, I select a random vertex $b(1)$ from the set $\mathbb{B}$ of vertices of the best-reply graph, according to a uniform distribution over $\mathbb{B}$. In each further simulation step t, types are randomly determined for both player and players choose best responses to $b(t-1)$ as described in section 3. The best responses determine the new state $b(t)$. In other words, I generate directed paths through the best-reply graph. As soon as a sink is reached (i.e. $b(t^*) = b(t^*+1) = ...$), the absorbing state is saved and a new simulation begins. Figure 1 summarizes the result of 1000 such simulations. The result suggests that the best-reply graph is weakly acyclic, which implies that learning converges almost surely to one of the four states in which players exhibit the behavior described by $b^{(1)}, b^{(2)}, b^{(3)}$ and $b^{(4)}$, according to proposition 1. These limit state are such that each type s proposes a bid which is equal to $s-1$, s or $s+1$. Hence, the result of this numerical example is consistent with the analytical result of section 4.1.[3]

	$b^{(1)}$	$b^{(2)}$	$b^{(3)}$	$b^{(4)}$	other	total
abs. freq.	255	241	268	236	0	1000

Figure 1: Absolute frequencies of the limit states for simulated transitions in the best-reply graph with random initial conditions.

4.3 Common Values and Asymmetries

In the previous specifications of the auction model, the learning process always converged to a symmetric absorbing state in which each bidder of type

[3]The resemblance of the limit states becomes stronger if the number of types in the present example is increased. The matlab-code of the simulation can be downloaded on the author's web-page.

s proposed a bid equal to $2s$. These absorbing states correspond to symmetric equilibria of the baseline model. Klemperer (1998) argues that small asymmetries between the bidders may result in unique equilibria which are asymmetric in a very strong sense. Let me introduce a (arbitrarily) small asymmetry between the bidders in order to check whether symmetry is a crucial assumption for the symmetric outcome to be obtained. Consider the following modification of the model. Whenever bidder one wins the auction, she receives a (small) bonus $\eta > 0$ in addition to the common value of the object. As before, bidder two only receives the common value of the object. Alternatively, assume that bidder one receives a small fraction $\eta > 0$ of the revenues after each auction. As the following proposition shows, both of these asymmetries have the same effect on the equilibrium and the outcome of the learning process.

The following assumptions on perturbations is similar to assumption 2. For technical reasons, I allow for trembles by more than one bid.

Assumption 4. *For any given player i and type s: whenever $h_i^s = (\bar{b}, \bar{b}, ..., \bar{b})$, $\beta_i^s(h)$ assigns probability $\varepsilon^{|b-\bar{b}|}$ to any bid $b \neq \bar{b}$ and the remaining probability mass $1 - \sum_{b \neq \bar{k}} \varepsilon^{|b-\bar{b}|}$ to $\bar{b}$, where $\varepsilon > 0$.*

Proposition 4. *Consider the second-price auction game Γ^A with common values and a uniform distribution of types. Let $T = 2$ (i.e. three types) and let η be a sufficiently small but positive number. Either assume that bidder one receives a fixed bonus of η whenever she wins the object or assume that bidder one receives a fraction η of the revenues after each auction. If the belief function satisfies assumption 3 and the perturbation $\varepsilon > 0$ is sufficiently small, the learning process converges almost surely to a state in which bidder one always wins.*

Proof. See Appendix. □

These results are consistent with those in Klemperer (1998). I showed that the introduction of small asymmetries in favor of bidder one (in combination with small perturbations of the belief function) implies convergence to a unique asymmetric state in which bidder one always wins, regardless of bidders' types.

While this result relies on the specific assumption on perturbations, figure 2 shows that for uniform perturbations (as defined in assumption 3), results are similar, but less extreme. Two Monte Carlo simulations runs with 10000 rounds each have been carried out. Despite the small degree of asymmetries, player one wins significantly more often than player two in both runs (66.2%, resp. 59.3%).

	pl.1 wins	pl.2 wins	tie	total
$\ell = 4, T = 6$	6620	3299	81	10000
$\ell = 5, T = 9$	5931	3965	104	10000

Figure 2: Asymmetries, uniform trembles, 10000 rounds, η=0.1, ε=0.01, M'=2T+2

4.4 Private Values

It is well known that if valuations are private (i.e. $\alpha_1 = \alpha_2 = 0$), each bidder has a weak incentive to propose a bid which is equal to her private signal (irrespectively of whether signals are continuous or discrete). Let $\widehat{h}$ denote the corresponding state of the learning process. If the learning process is currently in state $\widehat{h}$, the highest type T is indifferent between bidding T and any bid equal or larger than $T+1$. Eventually, she will propose a bid larger than T which, in turn, provides an incentive for type $T-1$ to increase her bid as well. In other words, $\widehat{h}$ is not an absorbing state of the learning process and convergence can not be obtained without further assumptions. Therefore, I assume as before that belief functions are slightly perturbed. This allows for the following result.

Proposition 5. *Consider the second-price auction game Γ^A with private valuations. If the perturbations of the belief function satisfy either assumption 2, assumption 3 or assumption 4 (while $\varepsilon > 0$ has to be sufficiently small in each case), the learning process converges almost surely to $\widehat{h}$.*

Proof. For any initial state, each type of each player has a weak incentive to bid according to $\widehat{h}$. Hence, $\widehat{h}$ is reached from any initial state after $2\ell(T+1)$ rounds with strictly positive probability. In $\widehat{h}$, each bidder has a strict incentive to bid according to $\widehat{h}$. □

5 First-price Auction: Some Examples

It is well known that in a first-price auction with private values (and continuous signals), there do not exist asymmetric equilibria. Under common values, asymmetric equilibria may exists but they are hard to study and not much is known about their properties, see Milgrom and Weber (1982). The symmetric equilibria of the first-price auction models under private and common valuations are given by the strategies $b_i(s_i) = s_i/2$ and $b_i(s_i) = s_i$, respectively. By employing a simple example with only 6 types, I show in section 5.1 that if valuations are private, learning in a discrete version of this model may converge to a state in which bidders follow strategies similar to the equilibrium strategies of the continuous benchmark model. Moreover, I show in section 5.2 that a similar result holds under common values.

As before, the distribution of types is assumed to be uniform. Since no general results can be obtained in this framework, I merely focus on examples. The following assumption on the belief function is maintained throughout this section.

Assumption 5. *For any given player i and type s: whenever $h_i^s = (\overline{b}, \overline{b}, ..., \overline{b})$, $\beta_i^s(h)$ assigns probability one to $\overline{b}$.*

5.1 First-price Auction with Independent Values

The following example suggests that, if the continuous auction model with private valuations is discretized and the number of types is limited, learning may converge to a state which is consistent with the prediction obtained from the equilibrium analysis. Note that in this example, the lowest types dispose of weakly dominant strategies (proposing a bid of zero) and therefore, according to proposition 1, the convergence result holds even without assumption 1.

Claim 4. *Consider the first-price auction game Γ^A with private valuations, $T = 5$ (i.e. six types) and a uniform distribution of types. The learning process converges almost surely to a symmetric state $\widehat{h}$, which satisfies $\widehat{h}_i^s = \{\lfloor \frac{s}{2} \rfloor, \lfloor \frac{s}{2} \rfloor, ..., \lfloor \frac{s}{2} \rfloor\}$, for all $i \in N$ and $s \in \mathcal{S}$ (where $\lfloor x \rfloor$ denotes the largest integer which is smaller or equal than x).*

Proof. Let $\widehat{b}$ denote the vertex of the best-reply graph that corresponds to $\widehat{h}$, i.e. $b_i^0 = b_i^1 = 0$, $b_i^2 = b_i^3 = 1$ and $b_i^4 = b_i^5 = 2$ (for $i \in N$). Clearly, any vertex b which satisfies $b_i^0 > 0$, for some i, is (weakly) dominated according to the definition in section 2. The same is true for any vertex b which satisfies $b_i^s \geq s$, where $s > 0$. Moreover, any vertex with $b_i^0 = 0$, $b_i^1 = 0$ and $b_{-i}^2 \neq 1$ is weakly dominated as well, because $\pi_{-i}^2(0 \to 1) \geq \frac{2}{6}(2-1) - \frac{1}{2}\frac{2}{6}(2-0) = 0$. Therefore, according to lemma 1 and lemma 2, it is sufficient to show that there exists a directed path of the best-reply graph from any vertex of the set

$$B' = \{b \in \mathbb{B} : b_i^0 = b_i^1 = 0,\ b_i^2 = 1,\ b_i^3 \leq 2,\ b_i^4 \leq 3,\ b_i^5 \leq 4, \text{ and } b_i^2 \leq b_i^3 \leq b_i^4 \leq b_i^5,\ i \in N\}$$

to some sink. Let $(k^3, k^4, k^5)_i$ denote the subset of vertices $b \in B'$ that satisfy $b_i^3 = k^3$, $b_i^4 = k^4$ and $b_i^5 = k^5$. For any vertex in $(1, 1, 1)_i$, there exists a directed path to a vertex in $(1,2,2)_{-i}$, because $\pi^3_{-i}(1 \to 2) = -\frac{1}{3} < 0$, $\pi^4_{-i}(1 \to 2) = 0$, $\pi^4_{-i}(2 \to 3) = -1 < 0$, $\pi^5_{-i}(1 \to 2) = \frac{1}{3} > 0$, $\pi^5_{-i}(2 \to 3) = \pi^5_{-i}(3 \to 4) = -1 < 0$. Similarly, it can be shown that there exists a directed path to a vertex in $(1,2,2)_{-i}$ from any vertex contained in either of the sets $(1,1,2)_i$, $(1,2,2)_i$, $(1,2,3)_i$, $(1,2,4)_i$, $(2,2,2)_i$, $(2,2,3)_i$ and $(2,2,4)_i$. Moreover, there exists a directed path to a vertex in $(1,1,2)_{-i}$ from any vertex contained in either of the sets $(1,1,3)_i$, $(1,1,4)_i$, $(1,3,3)_i$, $(1,3,4)_i$, $(2,3,3)_i$ and $(2,3,4)_i$. By putting together these transitions, it easily follows that there exists a directed path to $\widehat{b}$ from any vertex $b \in B'$. Finally, $\widehat{h}$ is absorbing, i.e. $BR^s_i(\widehat{b}) = \{\widehat{b^s_i}\}$, because $\pi^3_i(1 \to 2) = -\frac{1}{6} < 0$, $\pi^4_i(1 \to 2) = \frac{1}{6} > 0$, $\pi^4_i(2 \to 3) = -\frac{2}{3} < 0$, $\pi^5_i(1 \to 2) = \frac{1}{2} > 0$, $\pi^5_i(2 \to 3) = -\frac{1}{2} < 0$ and $\pi^5_i(3 \to 4) = -1 < 0$. □

5.2 First-price Auction with Common Values

In the case of common values, the equilibrium analysis in the continuous benchmark model assures the existence of a symmetric equilibrium, in which strategies are given by $b_i(s_i) = s_i$. Not much is known about asymmetric equilibria. The following example suggests that learning in a discrete model leads to symmetric outcomes, that correspond to the symmetric equilibrium of the continuous model.

Claim 5. *Consider the first-price auction game Γ^A with common valuations, $T = 2$ (i.e. three types) and a uniform distribution of types. Moreover, let assumption 1 hold. The learning process converges almost surely to a symmetric state $\widehat{h}$, which satisfies $\widehat{h}^s_i = \{s, s, ..., s\}$, for $s \in \{0,1\}$ and $\widehat{h}^2_i = \{1, 1, ..., 1\}$, for all $i \in N$.*

Proof. Let $\widehat{b}$ denote the vertex of the best-reply graph that corresponds to $\widehat{h}$, i.e. $b^0_i = 0$ and $b^1_i = b^2_i = 1$ (for $i \in N$). According to lemma 1, I restrict attention to vertices of the best-reply graph that satisfy $b^0_i \le b^1_i \le b^2_i$. This implies that players of type s can, in expectation, never gain more than their own signal s plus the average type of the opponent, i.e. they can never gain more than $s+1$. Therefore, according to lemma 1 and lemma 2, it is sufficient to show that there exists a directed path of the best-reply graph from any vertex of the set

$$B' = \{b \in \mathbb{B} : b^0_i \le 1,\, b^1_i \le 2,\, b^2_i \le 3,\, b^0_i \le b^1_i \le b^2_i,\, i \in N\}$$

to some sink. Let $(k^0, k^1, k^2)_i$ denote the subset of vertices $b \in B'$ that satisfy $b^0_i = k^0$, $b^1_i = k^1$ and $b^2_i = k^2$. For any vertex in $(0,0,0)_i$, there exists a directed path to a vertex in $(0,1,1)_{-i}$, because $\pi^0_{-i}(0 \to 1) = -\frac{1}{2} < 0$,

$\pi^1_{-i}(0 \to 1) = 0$, $\pi^1_{-i}(1 \to 2) = -1 < 0$, $\pi^2_{-i}(0 \to 1) = \frac{1}{2} > 0$ and $\pi^2_{-i}(1 \to 2) = \pi^2_{-i}(2 \to 3) = -1 < 0$. Similarly, it can be shown that there exists a directed path to a vertex in $(0,1,1)_{-i}$ from any vertex contained in either of the sets $(0,0,1)_i$, $(0,1,1)_i$, $(0,1,2)_i$, $(0,1,3)_i$, $(1,1,1)_i$, $(1,1,3)_i$ and $(1,2,3)_i$. Moreover, there exists a directed path to a vertex in $(0,0,1)_{-i}$ from any vertex contained in either of the sets $(0,0,2)_i$, $(0,0,3)_i$ and $(0,2,3)_i$ and there exists a directed path to a vertex in $(0,1,2)_{-i}$ from any vertex contained in either of the sets $(1,1,2)_i$ and $(1,2,2)_i$. Finally, it can be shown that there exists a directed path to a vertex in $(0,0,2)_{-i}$ from any vertex on the set $(0,2,2)_i$. By putting together these transitions, it easily follows that there exists a directed path to $\widehat{b}$ from any vertex $b \in B'$. Finally, $\widehat{h}$ is absorbing, i.e. $BR^s_i(\widehat{b}) = \{\widehat{b}^s_i\}$, because $\pi^0_i(0 \to 1) = -\frac{1}{6} < 0$, $\pi^1_i(0 \to 1) = \frac{1}{3} > 0$, $\pi^1_i(1 \to 2) = -\frac{1}{2} < 0$, $\pi^2_i(0 \to 1) = \frac{5}{6} > 0$, $\pi^2_i(1 \to 2) = -\frac{1}{6} < 0$ and $\pi^2_i(2 \to 3) = -1 < 0$.

$\square$

By undertaking considerably more numerical effort, the example can be extended to four and five types. It can be shown that in the case of four types, learning converges almost surely to a state in which the two low types bid 0 and the two high types bid 2. In the case of five types, the process converges almost surely to a state in which the lowest type bids 0, the next two types bid 1 and the highest two types bid 3. Moreover, in all three examples, there exist equilibria in which both players propose bids equal to their signals. These equilibria, however, are not strict and therefore not absorbing.

6 Conclusion

In the first part of this paper, I presented a general $n-$player model of learning under incomplete information and I introduced a framework based on best-reply graphs which allows to study its outcome. In particular, I showed that the problem of convergence can be reduced if the utility functions meet a simple monotonicity condition.

In the second part, I introduced a class of discrete $2-$player auction models which fits the general model. In particular, it covers first- and second-price auctions under private and common valuations. For these different specifications of the auction model, I showed that the learning process converges to states which correspond to the Nash equilibria of the respective continuous benchmark models. I argued that this justifies the *ad hoc* equilibrium selection approach taken by classical auction theory. Furthermore, I argued that such a learning approach justifies the concept of Nash equilibrium as a solution concept more generally. Since the above mentioned condition is satisfied by the considered class of auctions, I used it throughout the remainder of the paper in order to proof all main results.

I showed that in the second-price auction with common valuations and slightly perturbed beliefs, learning converges to a unique absorbing state of the learning process, provided either that types follow a uniform distribution or that the number of types is equal to two. The absorbing state was shown to be consistent with the equilibrium of the corresponding continuous benchmark model. Through the use of Monte Carlo simulations, I generated evidence which suggests that the obtained convergence results do not hinge on the specific characteristics of the perturbations of beliefs. Moreover, I showed that if small asymmetries are introduced into the model, the outcome of learning remains consistent with the equilibrium of the continuous benchmark case. Finally, I showed that in the case of private valuations eventually every bidder proposes a bid equal to her own valuation. As before, this is consistent with the equilibrium of the continuous benchmark model.

Moreover, I provided two examples of a first-price auction model with independent and common valuations. These examples suggest that, under certain conditions, also the well-known equilibria of continuous first-price auctions can be regarded as the outcome of a learning process.

Appendix: Proofs

Proof of Claim 3

Consider any vertex b that satisfies the above condition and assume, in addition, that for all s with $1 \leq s < \widehat{s}$, b satisfies $b_i^s \leq 2(\widehat{s}-1)$, $\forall i \in N$, and $\max\{s; b_i^{s-1}\} \leq b_i^s \leq s+\widehat{s}-1$. This assumption on b is without loss of generality because lemma 1, claim 1 and claim 2 imply that some vertex b' that satisfies these requirements is connected to b through a finite number of directed edges. Consider this vertex b' and call it b (with slight abuse of notation). Let me construct a sequence of transitions which reaches some vertex in which both bidders bid $2(\widehat{s}-1)$ when they are of type $\widehat{s}-1$. To this end, I need to show that

$$\pi_i^{\widehat{s}-1}(k \rightarrow k+1) = \frac{1}{2}p_{-i}^k \left[A_{-i}^k + \widehat{s} - 1 - k\right] + \frac{1}{2}p_{-i}^{k+1} \left[A_{-i}^{k+1} + \widehat{s} - 1 - k - 1\right]$$

is non-negative for all $i \in N$ and $k \in \mathcal{B}$ with $k \leq 2\widehat{s}-3$. Let me first show that $A_{-i}^k \geq k - \widehat{s} + 1$ for all $k \leq 2\widehat{s}-3$ with $p_{-i}^k > 0$ (this implies that $\pi_i^{\widehat{s}-1}(k \rightarrow k+1) \geq 0$ for all $k \leq 2\widehat{s}-4$). Remember that $k \geq \widehat{s}-1$ and consider the following two (exhaustive) cases.

case A: $b_{-i}^{k-\widehat{s}+1} = k$
It follows from lemma 1 that $b_{-i}^{k-\widehat{s}+2} \geq k$. Since, $b_{-i}^{k-\widehat{s}} \leq k-1$ and $b_{-i}^{k-\widehat{s}+2} \leq k+1$, it follows that A_{-i}^k is minimized if $b_{-i}^{k-\widehat{s}} = k-1$ and $b_{-i}^{k-\widehat{s}+2} = k+1$. Hence,

$$A_{-i}^k \geq \varepsilon\left(k-\widehat{s}\right) + (1-2\varepsilon)\left(k-\widehat{s}+1\right) + \varepsilon\left(k-\widehat{s}+2\right) = k-\widehat{s}+1.$$

case B: $b_{-i}(b-\widehat{s}+1) < b$ and $p_{-i}^k > 0$
Lemma 1 and claim 2 imply that A_{-i}^b is minimized if $b_{-i}^{k-\widehat{s}} = b_{-i}^{k-\widehat{s}+1} = b_{-i}^{k-\widehat{s}+2}$. Hence,

$$A_{-i}^k \geq \frac{1}{3}\left(k-\widehat{s}+k-\widehat{s}+1+k-\widehat{s}+2\right) = k-\widehat{s}+1.$$

It remains to be shown that $\pi_i^{\widehat{s}-1}(2\widehat{s}-3 \rightarrow 2\widehat{s}-2) \geq 0$. This is clearly the case if $A_{-i}^{2\widehat{s}-2} \geq \widehat{s}-1$ or if $p_{-i}^{2\widehat{s}-2} = 0$. There are two cases that satisfy neither of these two conditions.

case I: $b_{-i}^{\widehat{s}-1} = b_{-i}^{\widehat{s}-2} = 2\widehat{s}-3$
In this case, $A_{-i}^{2\widehat{s}-2} = \frac{1}{2}(\widehat{s}-1) + \frac{1}{2}(\widehat{s}-2) = \widehat{s}-\frac{3}{2}$. If $\widehat{s}=2$, one gets $A_{-i}^1 = \frac{1}{2}$

and with $p^1_{-i} = \frac{2-4\varepsilon}{T+1}$ and $p^2_{-i} = \frac{2\varepsilon}{T+1}$, it follows that $\pi^1_{-i}(1 \to 2) = \frac{1-\varepsilon}{T+1} \geq 0$ if $\varepsilon \leq 1$. In case $\widehat{s} > 2$, $A^{2\widehat{s}-3}_{-i}$ is minimized if $b^{\widehat{s}-3}_{-i} = 2\widehat{s} - 4$, hence

$$A^{2\widehat{s}-3}_{-i} \geq \frac{(1-2\varepsilon)(\widehat{s}-1) + (1-2\varepsilon)(\widehat{s}-2) + \varepsilon(\widehat{s}-3)}{2-3\varepsilon}.$$

Since $p^{2\widehat{s}-3}_{-i} = \frac{2-3\varepsilon}{T+1}$ and $p^{2\widehat{s}-2}_{-i} = \frac{2\varepsilon}{T+1}$ it follows after a few transformations that

$$\pi^{\widehat{s}-1}_i(2\widehat{s}-3 \to 2\widehat{s}-2) \geq \frac{1-4\varepsilon}{2(T+1)} \geq 0 \text{ if } \varepsilon \leq \frac{1}{4}.$$

case II: $b^{\widehat{s}-1}_{-i} = 2\widehat{s} - 2$ and $b^{\widehat{s}-2}_{-i} = 2\widehat{s} - 3$

First, consider the case $\widehat{s} \geq 3$ and $b^{\widehat{s}-3}_{-i} = 2\widehat{s}-4$. It follows from the first part of this proof that there exists a best response of bidder i when she is of type $\widehat{s}-1$ which is either $2\widehat{s}-3$ or $2\widehat{s}-2$. Moreover, it is not hard to check that $\pi^{\widehat{s}-1}_i(2\widehat{s}-3 \to 2\widehat{s}-2) < 0$. Therefore, in case $b_{-i}(\widehat{s}-3) = 2\widehat{s}-4$ (and $\widehat{s} \geq 3$), the claim does not follow directly and I need to construct a finite sequence of best responses that induces bidder $-i$ to shade her bid when she is of type $\widehat{s}-3$. Let $\widehat{b} \in \mathcal{B}^{2(T+1)}$ denote the end point of this sequence. Clearly, I can assume that $\widehat{b}^{\widehat{s}-1}_i = 2\widehat{s}-3$. From claim 2 it follows that there exists a finite sequence of best responses that induces $\widehat{b}^s_i \leq \widehat{s} + s - 2$ for all $s \leq \widehat{s} - 2$. It remains to show that $\widehat{b}$ as specified above satisfies $\pi^{\widehat{s}-3}_{-i}(2\widehat{s}-5 \to 2\widehat{s}-4) \leq 0$. In case $\widehat{s} = 3$, $\pi^0_{-i}(1 \to 2)$ is maximized if $\widehat{b}^1_i = \widehat{b}^0_i = 0$, which implies $A^2_i = 2$ and $A^1_i = 0.5$. Hence, $\pi^0_{-i}(1 \to 2) < 0$ and therefore $\widehat{b}^0_{-i} = 1$. Now consider the case $\widehat{s} \geq 4$. Given the restrictions on $\widehat{b}_i$, it is not hard to see that this expression is maximized if $\widehat{b}^{\widehat{s}-2}_i \leq 2\widehat{s} - 6$ and $\widehat{b}^{\widehat{s}-q}_i \leq 2\widehat{s} - 7$ for all $q \in \{3, 4.., \widehat{s}\}$. Hence, $A^{2\widehat{s}-4}_i = \widehat{s} - 1$ and $A^{2\widehat{s}-5}_i = \widehat{s} - 2$, if $p^{2\widehat{s}-5}_i > 0$. It follows that $\pi^{\widehat{s}-2}_{-i}(2\widehat{s}-4 \to 2\widehat{s}-5) \leq 0$ and therefore $\widehat{b}^{\widehat{s}-3}_{-i} \leq 2\widehat{s}-5$. Finally, either assume $\widehat{s} \geq 2$ and consider any vertex b or let $\widehat{s} \geq 3$ and consider the state $\widehat{b}$ derived above. It follows that $A^{2\widehat{s}-3}_{-i} = \frac{\varepsilon(\widehat{s}-1)+(1-2\varepsilon)(\widehat{s}-2)}{1-\varepsilon} = \widehat{s} - 2 + \frac{\varepsilon}{1-\varepsilon}$ and $A^{2\widehat{s}-2}_{-i} = \frac{\varepsilon(\widehat{s}-2)+(1-2\varepsilon)(\widehat{s}-1)}{1-\varepsilon} = \widehat{s} - \frac{1}{1-\varepsilon}$. Using $A^{2\widehat{s}-2}_{-i} + A^{2\widehat{s}-3}_{-i} = 2\widehat{s} - 3$ it finally follows that $\pi^{\widehat{s}-1}_i(2\widehat{s}-3 \to 2\widehat{s}-2) = 0$.

Q.E.D.

Proof of Proposition 4

Let $\widehat{b}$ denote the vertex of the best-reply graph that satisfies $b_1^0 = 0$, $b_1^1 = 1$, $b_1^2 = 2$, $b_2^0 = 3$, $b_2^1 = 4$ and $b_2^2 = 5$. Clearly, any vertex b which satisfies $b_2^0 > 2$, $b_2^1 > 3$ or $b_2^2 > 4$ and any vertex b which satisfies $b_1^0 > 3$, $b_1^1 > 4$ or $b_1^2 > 5$ is (weakly) dominated according to the definition in section 2. Therefore, according to lemma 1 and proposition 2, it is sufficient to show that there exists a directed path of the best-reply graph from any vertex of the set

$$\begin{aligned} B' \;=\; & \{b \in \mathbb{B} : b_1^0 \leq 3,\, b_1^1 \leq 4,\, b_1^2 \leq 5,\, b_2^0 \leq 2,\, b_2^1 \leq 3,\, b_2^2 \leq 4 \\ & \text{and } b_1^0 \leq b_1^1 \leq b_1^2,\, b_2^0 \leq b_2^1 \leq b_2^2\} \end{aligned}$$

to some sink. Let $(k^0k^1k^2)_i$ denote the subset of vertices $b \in B'$ that satisfy $b_i^0 = k^0$, $b_i^1 = k^1$ and $b_i^2 = k^2$. For any vertex in $(000)_2$, there exists a directed path to a vertex in $(234)_1$, because

$$\pi_1^0(0 \to 1) = \frac{1}{2}\left[(1-\varepsilon')\left(\frac{0+1+2}{3} + \eta - 0\right) + \varepsilon\left(\frac{0+1+2}{3} + \eta - 1\right)\right] > 0$$

(where ε' represents an expression which is larger than ε and goes to zero as ε goes to zero), $\pi_1^0(1 \to 2) = \frac{1}{2}\left[\varepsilon^2(\eta-1) + \varepsilon\eta\right] > 0$, $\pi_1^0(2 \to 3) = \frac{1}{2}\left[\varepsilon^3(\eta-2) + \varepsilon^2(\eta-1)\right] < 0$, $\pi_1^1(2 \to 3) = \frac{1}{2}\left[\varepsilon^3(\eta-1) + \varepsilon^2\eta\right] > 0$, $\pi_1^1(3 \to 4) = \frac{1}{2}\left[\varepsilon^4(\eta-2) + \varepsilon^3(\eta-1)\right] < 0$, $\pi_1^2(3 \to 4) = \frac{1}{2}\left[\varepsilon^4(\eta-1) + \varepsilon^3\eta\right] > 0$ and $\pi_1^2(4 \to 5) = \frac{1}{2}\left[\varepsilon^5(\eta-2) + \varepsilon^4(\eta-1)\right] < 0$.

In the same way, in can be shown that there exist directed paths from the following vertices to the following sets. From any vertex in $(011)_2$ and $(111)_2$ to $(234)_1$, from any vertex in $(001)_2$ to $(344)_1$, from any vertex in $(002)_2$, $(012)_2$, $(013)_2$ to $(345)_1$, from any vertex in $(003)_2$, $(022)_2$, $(023)_2$ to $(145)_1$, from any vertex in $(004)_2$, $(024)_2$ to $(135)_1$, from any vertex in $(014)_2$ to $(235)_1$, from any vertex in $(022)_2$, $(222)_2$ to $(134)_1$, from any vertex in $(033)_2$ to $(124)_1$, from any vertex in $(034)_2$ to $(125)_1$, from any vertex in $(112)_2$, $(113)_2$, $(123)_2$ to $(045)_1$, from any vertex in $(114)_2$, $(134)_2$ to $(025)_1$, from any vertex in $(112)_2$ to $(034)_1$, from any vertex in $(124)_2$ to $(035)_1$, from any vertex in $(133)_2$ to $(024)_1$, from any vertex in $(223)_2$, $(224)_2$, $(234)_2$ to $(015)_1$ and from any vertex in $(233)_2$ to $(014)_1$.

It is sufficient to show that from any vertex of the above sets, there exists a directed path to a sink. For any vertex in $(234)_1$, there exists a directed path to a vertex in $(013)_2$, because $\pi_2^0(0 \to 1) = \frac{1}{6}\left[-\varepsilon + 2\varepsilon^3 + \varepsilon^4\right] < 0$, $\pi_2^1(0 \to 1) = \frac{1}{6}\left[2\varepsilon^2 + 4\varepsilon^3 + 3\varepsilon^4\right] > 0$, $\pi_2^1(1 \to 2) = \frac{1}{6}\left[-(1-\varepsilon') + 2\varepsilon^2 + 2\varepsilon^3\right] < 0$, $\pi_2^2(1 \to 2) = \frac{1}{6}\left[2\varepsilon + 4\varepsilon^2 + 3\varepsilon^3\right] > 0$, $\pi_2^2(2 \to 3) = \frac{1}{6}\left[\varepsilon + 2\varepsilon^2\right] > 0$ and $\pi_2^2(3 \to 4) = \frac{1}{6}\left[-\varepsilon - 2\varepsilon^2\right] < 0$. Similarly, it can be shown that there exists a directed path to $(013)_2$ from any vertex in $(045)_1$, $(134)_1$, $(135)_1$, $(145)_1$

and $(235)_1$, a directed path to $(012)_2$ from any vertex in $(344)_1$ and $(345)_1$, a directed path to $(023)_2$ from any vertex in $(025)_1$, $(034)_1$, $(035)_1$ and $(125)_1$ and a directed path to $(024)_2$ from any vertex in $(014)_1$, $(015)_1$, $(024)_1$ and $(124)_1$.

As shown before, from any vertex in $(012)_2$, $(013)_2$, $(023)_2$ and $(024)_2$ vertices in $(135)_1$, $(145)_1$ and $(345)_1$ can be reached, from which, in turn, vertices in $(012)_2$, $(013)_2$ can be reached, from where $(345)_1$ can be reached. Hence, from any vertex in B', there exists a directed path to a vertex in $(345)_1$ and therefore a connected path to $\widehat{b}$.

Finally, it remains to be shown that $\widehat{h}$ is absorbing, i.e. $BR_i^s(\widehat{b}) = \{\widehat{b}_i^s\}$. First, consider a state in $(012)_2$. In this case,

$$\begin{aligned}
\pi_1^0(0 \to 1) &= \frac{1}{6}\left[2(1-\varepsilon')\eta + \varepsilon(3\eta+1) + \varepsilon^2(\eta+2)\right] > 0 \\
\pi_1^0(1 \to 2) &= \frac{1}{6}\left[2(1-\varepsilon')\eta + \varepsilon(3\eta-1) + \varepsilon^2(\eta-2)\right] > 0 \\
\pi_1^0(2 \to 3) &= \frac{1}{6}\left[(1-\varepsilon')\eta + 2\varepsilon(\eta-1) + 2\varepsilon^2(\eta-2) + \varepsilon^3(\eta-3)\right] > 0 \\
\pi_1^0(3 \to 4) &= \frac{1}{6}\left[\varepsilon(\eta-1) + 2\varepsilon^2(\eta-2) + 2\varepsilon^3(\eta-3) + \varepsilon^4(\eta-4)\right] < 0 \\
\pi_1^1(3 \to 4) &= \frac{1}{6}\left[\varepsilon\eta + 2\varepsilon^2(\eta-1) + 2\varepsilon^3(\eta-2) + \varepsilon^4(\eta-3)\right] > 0 \\
\pi_1^1(4 \to 5) &= \frac{1}{6}\left[\varepsilon^2(\eta-1) + 2\varepsilon^3(\eta-2) + 2\varepsilon^4(\eta-3) + \varepsilon^5(\eta-4)\right] < 0 \\
\pi_1^2(4 \to 5) &= \frac{1}{6}\left[\varepsilon^2\eta + 2\varepsilon^3(\eta-1) + 2\varepsilon^4(\eta-2) + \varepsilon^5(\eta-3)\right] > 0 \\
\pi_1^2(5 \to 6) &= \frac{1}{6}\left[\varepsilon^3(\eta-1) + 2\varepsilon^4(\eta-2) + 2\varepsilon^5(\eta-3) + \varepsilon^6(\eta-4)\right] < 0,
\end{aligned}$$

which implies that $BR_1^0(\widehat{b}) = \{3\}$, $BR_1^1(\widehat{b}) = \{4\}$ and $BR_1^2(\widehat{b}) = \{5\}$. Second, consider a state in $(345)_1$. In this case, $\pi_2^0(0 \to 1) = \frac{1}{6}\left[-\varepsilon^2 + 2\varepsilon^4 + 2\varepsilon^5\right] < 0$, $\pi_2^1(0 \to 1) = \frac{1}{6}\left[2\varepsilon^3 + 4\varepsilon^4 + 3\varepsilon^5\right] > 0$, $\pi_2^1(1 \to 2) = \frac{1}{6}\left[-\varepsilon + 2\varepsilon^3 + 2\varepsilon^4\right] < 0$, $\pi_2^2(1 \to 2) = \frac{1}{6}\left[2\varepsilon^2 + 4\varepsilon^3 + 3\varepsilon^4\right] > 0$ and $\pi_2^2(2 \to 3) = \frac{1}{6}\left[-(1-\varepsilon') + 2\varepsilon^2 + 2\varepsilon^3\right] < 0$, which implies $BR_2^0 = \{0\}$, $BR_2^1 = \{1\}$ and $BR_2^2 = \{2\}$.
Q.E.D.

References

[1] Agastya, M., 2004. Stochastic Stability in a Double Auction. *Games and Economic Behavior* 48, 203-222.

[2] Bergin, J. and Lipman, B.L., 1996. Evolution with State-Dependent Mutations. *Econometrica* 64, 943-956.

[3] Brown, G.W., 1951. Iterative Solutions of Games by Fictitious Play. *Activity Analysis of Production and Allocation.* Wiley, New York.

[4] Canning, D., 1992. Language Conventions in Common Interest Signaling Games. Discussion Paper #607, Columbia University.

[5] Fudenberg, D. and Levine, D.K., 1998. The theory of learning in games. *MIT Press, Cambridge, Massachusetts.*

[6] Hon-Snir, S., Monderer, D. and Sela, A., 1998. A Learning Approach to Auctions. *Journal of Economic Theory* 82, 65-88.

[7] Jacobsen, H.J., Jensen, M. and Sloth, B., 2001. Evolutionary Learning in Signalling Games. *Games and Economic Behavior* 34, 34-63.

[8] Jensen, M., Sloth, B. and Whitta-Jacobsen, H.J., 2005. The Evolution of Conventions Under Incomplete Information. *Economic Theory* 25, 171-185.

[9] Kandori, M., Mailath, G. and Rob, R., 1993. Learning, Mutation, and long-run Equilibria in Games. *Econometrica* 61, 29-56.

[10] Klemperer, P., 1998. Auctions with Almost Common Values: The 'Wallet Game' and its Applications. *European Economic Review* 42, 757-769.

[11] Krishna, V., 2002. Auction Theory. *New York: Academic Press.*

[12] Milgrom, P.R., 1981, Rational Expectations, Information Acquisition, and Competitive Bidding. *Econometrica* 49. 921-943.

[13] Milgrom, P.R. and Weber, R.J., 1982. A Theory of Auctions and Competitive Bidding. *Econometrica* 50, 1089-1122.

[14] Nöldeke, G. and Samuelson L., 1997. A Dynamic Model of Equilibrium Selection in Signaling Markets. *Journal of Economic Theory* 73, 118-156.

[15] Weibull, J.W., 1997. Evolutionary game theory. *MIT Press, Cambridge, Massachusetts.*

[16] Young, H.P., 1993. The Evolution of Conventions. *Econometrica* 61, 57-84.

Alexander Gaigl[†]

Observational Learning & Strategic Externalities[*]

University of Mannheim
December 2009

Abstract

I introduce strategic externalities into a standard herding model. It is assumed that such externalities only affect successors. I study the interplay of informational and strategic externalities and I determine how their relative magnitudes affects the occurrence of herds and informational cascades. If strategic externalities (measured by a parameter σ) are negative and sufficiently strong, an informational cascades arises but there is *no* herding ('*anti-herding*' occurs). This contrasts with the existing literature which generally finds that an informational cascade implies herding. In a continuous-signal version of the model, I show that there exists an interval of σ in which learning is more efficient than in the $\sigma = 0$ case. Moreover, there always exists one value of σ such that *every* individual reveals her signal. I make different assumptions on the observability of actions and I show that agents may engage in either imitative or contrarian behavior, depending on the value of σ. It is shown that some previous results on herding and informational cascades are not robust. Finally, I study the model under binary signals. It is shown that negative strategic externalities always prevent herding and may lead to a considerable increase in the efficiency of learning.

JEL Classification: D82, D83

Keywords: Social Learning, Strategic externalities, Herding, Informational Cascades

[†]Contacts: Department of Economics, University of Mannheim, D-68131 Mannheim, Germany, agaigl[at]rumms.uni-mannheim.de, http://www.vwl.uni-mannheim.de/gk/_gaigl.

[*]I thank Martin Cripps, Jörg Oechssler, Philippe Jehiel, Ran Spiegler and Heiner Schumacher for helpful suggestions. All remaining errors are my own. Moreover, I am grateful to the Economics Department at UCL for their hospitality.

1 Introduction

There is a large number of social and economic situations in which individuals are influenced by the actions of others. Common examples include consumer purchase decisions, the choice of a restaurant, the adoption of a new technology and asset market decisions. It is well known that individuals may underestimate their private information in such situations and that this can lead to inefficient behavior such as fads, booms, financial market bubbles and busts, bank runs, or the failure of firms to coordinate on the adoption the best technology. A comprehensive literature pioneered by Banerjee (1992) and Bikhchandani, Hirshleifer and Welch (1992) has used herd behaviour and informational cascades to explain a variety of such phenomena. This literature analyzes sequential action models with incomplete and asymmetric information, where agents can observe actions taken by their predecessors but not the information upon which these actions were taken. These models generally predict that agents eventually herd on one action, i.e. that rational agents imitate their predecessors' behavior even if it conflicts with their own information.

A central assumption of most herding models is that agents' payoffs are independent of the actions of others. Therefore, agents are concerned about the actions of others only to the extent that they reveal private information about the unknown state of the world. In other words, the only externality present in these models is an informational one. In many real-world situations, however, social learning is significantly affected by the direct dependence of payoffs on actions. This dependence may have a crucial impact on the occurrence of herds and informational cascades. A long waiting line in front of a night club, for example, may well indicate high quality of the club, but herding nevertheless only occurs to a certain extent because eventually newcomers are deterred by the long line. This paper contributes to the literature on social learning by relaxing the assumption of pure informational externalities and by allowing for *strategic externalities*. Strategic externalities will be referred to as strategic *substitutabilities (complementarities)* if they are negative (positive).

In models with pure informational externalities, there is no incentive for agents to behave strategically. Therefore, they exhibit pure backward-looking behaviour, which is relatively easy to study. In the presence of strategic externalities, however, agents also need to consider the impact of their actions on the actions of successors, i.e. they need to exhibit forward-looking behavior. In the technology adoption problem, for example, the choices of predecessors are just as important as the choices of successors because each firm attempts to adopt the technology used by the majority of *all* firms. This may lead to complex types of behavior, involving strategic signalling of private information. Agents' behavior in such models is very hard to study and there is generally a multiplicity of equilibria, see for example Dasgupta (2000) or Drehmann, Oechssler and Roider (2007).

In this paper, I analyze strategic externalities in standard herding models under the restricting assumption that externalities are imposed only on successors. This implies that agents are entirely backward-looking. They do not behave strategically but are merely interested in learning their predecessors' private information from previous actions, while taking into account the strategic externality effects of these actions. In appendix A, I briefly introduce a model with allows individuals to be backward-looking as well as forward-looking. Since externalities cancel each other out in this model, there exists a simple and efficient equilibrium in which agents rely entirely on their own signal.

There are many real-world situations in which individuals impose externalities only on their successors. Examples include waiting lists for a new product or the choice of a parking lot (if availability is not observable before entering the parking lot). More generally, all first-come-first-served queuing systems with waiting costs and uncertainty about the quality of service are valid examples, see Debo, Parlour and Rajan (2005). While these examples involve strategic substitutabilities, there are other situations in which actions impose strategic complementarities on successors. Consider, for example, an individual's decision on whether or not to commit a crime. If a potential criminal observes that many peers have committed the same crime, she may infer that the probability of gain is high, see Kahan (1997). In addition to this informational externality, there is a strategic complementarity (imposed only on successors) due to reduced law enforcement in case many crimes have been committed in the past.

There is another way of interpreting the backward-looking approach. The model adopted here is equivalent to a model in which payoffs are fully dependent on actions of predecessors *and* successors, but in which individuals are boundedly rational in the sense that they do not consider the impact of their choice on the actions of successors. Experimental evidence suggests that individuals may indeed behave in this fashion, see Drehmann, Oechssler and Roider (2007).

The general model is introduced in section 2. I consider a standard setting of social learning, in which agents choose binary actions in a chronological order. Before deciding on a particular action, each agent receives some information (this will be specified below) about the behavior of her predecessors and a private signal that is correlated with the unknown state of the world. The action space is discrete throughout this paper, and I study the model under both a continuous and a discrete signal space in order to make it comparable to large parts of the social learning literature. While the bulk of this literature is of the discrete-signal-discrete-action type, some interesting models of the continuous-signal-discrete-action type have been proposed as well, most notably Smith and Sørensen (2000).

The main focus of my paper is on the efficiency of learning. While sufficiently strong strategic substitutabilities always imply inefficient and strongly non-conformist behavior (so-called *anti-herding*), I show that moderate strategic substitutabilities may lead to an increase in efficiency since neither herding nor anti-herding arise. In particular, I show that there exists a special case in which the informational externality of each action and the corresponding strategic externality cancel each other out. This induces agents to rely entirely on their private information and therefore learning is efficient. Moreover, I show that strategic complementarities have a negative effect on the efficiency of learning since they increase agents' tendency towards conformity and therefore reinforce inefficient herding.

In section 3, I study the model under a continuous signal space. Continuous signals have the feature that they allow a distinction between strong and weak signals, which affects agents' tendency towards incorporating predecessor's decisions into their own action choice. It will turn out that analyzing this tendency subject to agents' confidence in their private information reveals some interesting insights. I show how agents' signals influence their decisions and under what conditions (anti-)herding and cascade behavior arises. Most analytical expressions are derived under a uniform distribution. However, at the end of the section, I show that the results do not change qualitatively if one allows for a general distribution.

The two central concepts studied here are herds and informational cascades. While in discrete models such as Banerjee (1992) and Bikhchandani, Hirshleifer and Welch (1992), these notions are used to address the same phenomenon, Smith and Sørensen (2000) introduce them as two separate concepts. According to their definitions, an informational cascade is an infinite sequence of agents who all neglect their private information and base their actions entirely on the observations of the behavior of others. Hence, once a cascade starts, public information stops accumulating. Herd behavior, on the other hand, describes a situation in which all agents take the same action, not necessarily ignoring their private information. In this case, public information can in principle still be accumulated.

The social learning literature generally finds that informational cascades arise in situations in which agents imitate their predecessors, i.e. they are due to conformist behavior. This paper shows that in the presence of strategic externalities, extreme forms of non-conformist behavior may provide another explanation for informational cascades. It is shown that if strategic substitutabilities are sufficiently strong, agents neglect their private information and engage in anti-herding behavior. In other words, I show that in the presence of strategic externalities, informational cascades are neither necessary nor sufficient for the occurrence of herding behavior. This result contrasts with previous results obtained in models of pure informational externalities.

A common feature of most models of social learning is that agents are able to observe all actions taken by their predecessors (which I refer to as 'perfect observability'). In reality, however, individuals can usually only observe a limited number of others. Several models of social learning take this into account by making restrictions on the observability of other players. For example, Smith and Sørensen (1996) study a model in which recent predecessors are more likely to be observed. In section 3, I study two specifications of the continuous-signal version of my model; one with perfect observability and one with imperfect observability. For the sake of simplicity, I assume that under imperfect observability, agents only observe the actions taken by their immediate predecessors. Both of these informational scenarios are treated by Çelen and Kariv (2004). They find that the informational structure matters for the occurrence of herds. I show that this result only holds true in the presence of pure informational externalities. In the presence of strategic externalities, herding and cascade behavior occur independently of the informational scenario. Moreover, I show that some other results of Celen and Kariv (2004) are not robust since they strongly depend on the absence of strategic externalities.

In addition to greater realism, the imperfect observability scenario has another motivation. Since each agent's information structure is relatively simple in this case (it is two-dimensional: agents observe their private signal and their predecessor's action), it is possible to study directly how agents respond to behavior of their predecessors depending on the degree of strategic externalities. In particular, imitative and contrarian behavior can be elicited (contrarian behavior is defined as taking the action opposite to the predecessor's action). In the perfect observability scenario, this is analytically impossible due to the large and time-varying sets of possible observational histories. I find that if agents only observes their predecessors' actions, weak signals induce them to (rationally) base their action choice on the observation of their predecessor's action, by entirely neglecting their own signal. This either happens through imitating or through contrarian behavior. If strategic externalities are strong, even agents with relatively strong private signals engage into imitation or contrarian behavior and, at the extreme, this may lead to herding or anti-herding. This analysis produces (recursive) analytical expressions describing the relationship between individual tendencies towards imitative or contrarian behavior and the overall social tendency towards herding, anti-herding or cascade behavior.

In section 4, I study the model under the assumption of binary signals (and perfect observability). I show that in the presence of strategic complementarities, herding always occurs, as in Banerjee (1992) and Bikhchandani, Hirshleifer and Welch (1992). Under strategic substitutabilities, herding never occurs but there always is an informational cascade with anti-herding. It is shown that anti-herding typically occurs after a substantial degree of learning has taken place and I show that strategic complementarities may significantly increase the efficiency of learning. However, it is shown that there always exist two possible outcomes of social learning, one in which the public belief is close to the true state of the world and one, in which the public belief assigns a probability close zero to it. It is shown that the problem of convergence to either one of these two states can be transformed into a version of the Gambler's Ruin Problem. Specifically, the probability that social learning is correct is equivalent to the probability that the gambler doubles her fortune, while the probability that social learning is wrong is equal to the probability that the gambler goes broke. Using the solution to the Gambler's Ruin Problem, it is shown that social learning becomes arbitrarily accurate (i.e. the public belief is arbitrarily close to the true state of the world, with arbitrarily high probability), provided that either the precision of the signal goes to one or the magnitude of the strategic externalities goes to zero. However, there may be very long spells of imitative behavior between two agents who reveal their signals. I show via Monte Carlo simulation that learning may indeed take a very long time to reach the absorbing state.

Some alternative studies of the interplay between informational and strategic externalities have been proposed. Choi (1997) analyzes this relationship in a technology adoption framework. In his model, firms do not observe private signals about the quality of alternative technologies and therefore herding is merely due to strategic externalities. Dasgupta (2000) introduces strategic complementarities into a sequential choice model with continuous private signals. Agents receive a positive return from investing only if all the other agents also choose to invest. Moreover, Frisell (2003) studies the interplay between strategic and informational externalities in a waiting game. Finally, there exists a recent literature that combines models of social learning with queueing models, see Debo, Parlour and Rajan (2005) and Debo and Veeraraghavan (2008). This literature analyzes the relationship between the (positive) informational externalities of long lines and the (negative) strategic externalities induced by the corresponding waiting costs.

2 The Basic Model

Consider a countable set of identical agents, indexed by $n = 1, 2, ..., N$, where $N \in \mathbb{N} \cup \infty$. Each agent n has to make a once-in-a-lifetime decision a_n between two alternatives from the set $\mathcal{A} \equiv \{-1, 1\}$. Decisions are made sequentially in an exogenously determined order. Let $\mathcal{N}_a(n)$ denote the number of predecessors of agent n who have chosen alternative a, i.e. $\mathcal{N}_a(n) = \sum_{i=1}^{n-1} \mathbf{1}_{\{a_i=a\}}$, where $\mathbf{1}_{\{\cdot\}}$ represents the indicator function. The payoff of alternative a depends on $\mathcal{N}_a$ as well as on the payoff-relevant *quality* $Q_a \in \mathbb{R}$ of alternative a. Since *relative* qualities are sufficient to describe agent's choices, the state of the world is defined as $\theta_0 = Q_1 - Q_{-1}$. Assume that agents are risk-neutral with utility function

$$u_n(a) = \sigma \mathcal{N}_a(n) + Q_a(\theta_0),$$

where $\sigma \in (-\infty, +\infty)$ measures the size of strategic externalities.

The state of the world is unknown to all agents but each agent n observes a private signal $\theta_n \in \Theta_S \subset \mathbb{R}$ which is correlated with θ_0. Assume that, conditional on θ_0, all θ_n are iid and follow a commonly known distribution.

Furthermore, let $\mathcal{I}_n$ denote the payoff-relevant information available to agent n. I study two different informational scenarios. In the *perfect observability* (PO) scenario, agents observe the actions of all their predecessors, whereas in the *imperfect observability* (IO) scenario only the immediate predecessor's action is observable. Formally, $\mathcal{I}_n^{PO} = (\theta_n, (a_i)_{i=1}^{n-1})$ and $\mathcal{I}_n^{IO} = (\theta_n, a_{n-1})$. It follows that for given $\mathcal{I}_n$, agent n's optimal decision rule is given by:

$$\begin{aligned} a_n = 1 \text{ if and only if } \qquad & \mathrm{E}\left[\sigma \mathcal{N}_1(n) - \sigma \mathcal{N}_{-1}(n) + Q_1(\theta_0) - Q_{-1}(\theta_0) \,|\, \mathcal{I}_n\right] \\ = \; & \mathrm{E}\left[\sigma \sum_{i=1}^{n-1} a_i + \theta_0 \,|\, \mathcal{I}_n\right] \geq 0. \qquad (1) \end{aligned}$$

The decision rule contains the implicit assumption that agents always chose alternative $a = 1$ in case of indifference. This assumption does not affect the results since indifference is an event of zero-measure for each of the specifications of the model.

The focus of this paper is mainly on the occurrence of informational cascades, herds and antiherds (these concepts will be defined below), i.e. on agents' behavior as the number N of agents goes to infinity. Therefore, I either consider the limit of a sequence of economies indexed by N, where N tends to infinity (as in section 3) or I set N directly equal to infinity (as in section 4). The finite-economy approach is used in the continuous signal case in section 3 because under this specification of the model, an economy of finite (but possibly very large) size is required. Note that this approach is analytically straightforward since agents' behavior exclusively depends on the actions of their predecessors and therefore the size of the economy is irrelevant to the decision of any agent. For notational convenience, I occasionally consider only the limit as n goes to infinity without explicitly stating the double limit as both N and n tend to infinity.

Definition 1. An *informational cascade* occurs if there is some agent n_0 and some alternative $a \in \mathcal{A}$ such that for all $n \geq n_0$: $\theta_n \in \Theta_S \Rightarrow a_n = a$ (i.e. all agents $n \geq n_0$ disregard their private signals). In case $N < \infty$, this property has to be satisfied in the limit as N goes to infinity.

Definition 2. A *run*[1] (*anti-run*) at position n_0 of length $m_{n_0}^N$ satisfies $a_n = a_{n-1}$ ($a_n = -a_{n-1}$) for all $n = n_0+1, n_0+2, ..., n_0+m_{n_0}^N$. If $N = \infty$, a *herd* (*anti-herd*) is said to occur if there is some n_0 such that $m_{n_0}^N = \infty$. If $N < \infty$ a *herd* (*anti-herd*) is said to occur if there exists some n_0 such that $\lim_{N\to\infty} m_{n_0}^N = \infty$ (i.e. a run (anti-run) that grows arbitrarily large as the size of the economy goes to infinity). Herding (anti-herding) is *complete* if $n_0 = 1$.

Definition 3. Agent n is *self-reliant* (for given θ_n), if a_n is independent of $(a_i)_{i=1}^{n-1}$. Agent n is *purely* self-reliant if she is self-reliant for all $\theta_n \in \Theta_S$.

3 Continuous signals

In this section, private signals θ_n follow a symmetric distribution with c.d.f. F over the support $\Theta_S \equiv [-b, b]$, where $b < \infty$. For simplicity, I assume that F is continuous and differentiable. The set of agents is finite, i.e. $N < \infty$, and the state of the world is given by the sum of individual signals, i.e. $\theta_0 = \sum_{i=1}^N \theta_i$. Hence, agent n's optimal behavior is determined by:

$$a_n = 1 \text{ if and only if } E\left[\sigma \sum_{i=1}^{n-1} a_i + \sum_{i=1}^{N} \theta_i \,|\, \mathcal{I}_n\right] \geq 0$$

Using the fact that the expected value of successors' signals is always zero, the decision rule can conveniently be expressed as a history-contingent cutoff-rule for θ_n:

$$a_n = 1 \text{ if and only if } \ \theta_n \geq -E\left[\sigma \sum_{i=1}^{n-1} a_i + \sum_{i=1}^{n-1} \theta_i \,|\, \mathcal{I}_n\right]. \qquad (2)$$

[1]This term is borrowed from Drehmann, Oechssler and Roider (2007).

Let $\widehat{\theta}_n$ denote the right-hand side of this inequality, i.e. the cutoff-level for agent n's signal. Note that $\widehat{\theta}_n$ is sufficient to describe the behavior of agent n. Therefore, I focus my analysis on the stochastic process $\{\widehat{\theta}_n\}$, which I will refer to as the *learning process.* In the remainder of this section, I study the properties of the learning process under different assumptions on the signal distribution F as well as different assumptions on the degree of observability of other agents' actions. In subsection 3.1, I assume that F is uniform, while subsection 3.2 shows that this assumption can be relaxed without loosing some of the key properties.

3.1 Uniform Signals

Let F be a uniform distribution on the interval $[-1, 1]$, i.e. $b = 1$. In the following, I study the corresponding learning processes in the perfect and the imperfect observability scenarios.

3.1.1 Perfect Observability and Uniform Signals

Let $\mathcal{I}_n = \mathcal{I}_n^{PO}$, i.e. assume that in addition to her private signal θ_n, each agent n observes the entire sequence $(a_i)_{i=1}^{n-1}$ of predecessors' actions before taking a decision. In this case, the learning process is a stochastic process which is characterized as follows.

Proposition 1. *The learning process satisfies* $\widehat{\theta}_1 = 0$ *and*

$$\widehat{\theta}_{n+1} = \begin{cases} 1 & \text{if } \widehat{\theta}'_{n+1} > 1 \\ -1 & \text{if } \widehat{\theta}'_{n+1} < -1 \\ \widehat{\theta}'_{n+1} & \text{otherwise} \end{cases}$$

where $\widehat{\theta}'_{n+1} = \frac{\widehat{\theta}_n}{2} - a_n\left(\frac{1}{2} + \sigma\right)$.

Proof. Using the fact that F is uniform and equation (2), it follows that

$$\begin{aligned} a_n = 1 \text{ iff } \theta_n \geq \widehat{\theta}_n &= -\mathrm{E}\left[\sum_{i=1}^{n-1} (\theta_i + \sigma a_i) \mid (a_i)_{i=1}^{n-1}\right] \\ &= \widehat{\theta}_{n-1} - \mathrm{E}\left[\theta_{n-1} + \sigma a_{n-1} \mid (a_i)_{i=1}^{n-1}\right] \\ &= \widehat{\theta}_{n-1} - \begin{cases} \frac{1}{2}(\widehat{\theta}_{n-1} + 1) + \sigma & \text{if } a_{n-1} = 1 \\ \frac{1}{2}(\widehat{\theta}_{n-1} - 1) - \sigma & \text{if } a_{n-1} = -1 \end{cases} \\ &= \frac{\widehat{\theta}_{n-1}}{2} - a_{n-1}\left(\frac{1}{2} + \sigma\right) \end{aligned}$$

Clearly, agent one's cutoff is $\widehat{\theta}_1 = 0$, because $a_1 = 1$ iff $\theta_1 \geq 0$. □

Figures 1 and 2 show the realizations of several Monte Carlo simulations of $\widehat{\theta}_n$ for different σ (and for sufficiently large N). In case $\sigma = 0.01$, the learning process eventually settles down in the point $\widehat{\theta} = 1$, which implies identical actions, regardless of private signals. This is equivalent to an informational cascade with herding. Moreover, the plots suggest that the autocorrelation of the learning process increases in σ. Intuitively, this is due to the fact that increasingly strong strategic substitutabilities increase agents' tendency to 'avoid' the actions of their predecessors. The following proposition describes the behavior of the learning process for different σ.

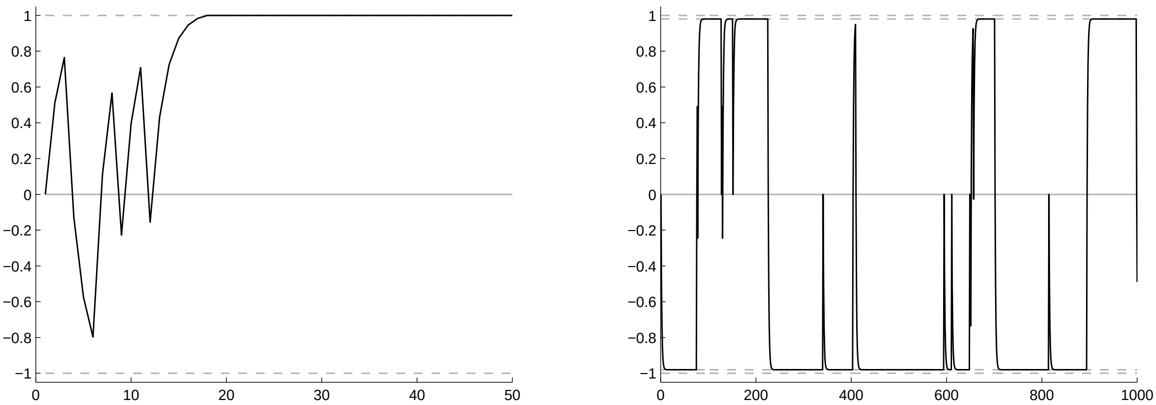


Figure 1: Simulation runs of $\widehat{\theta}_n$ for $\sigma = 0.01$ (left) and $\sigma = -0.01$ (right).

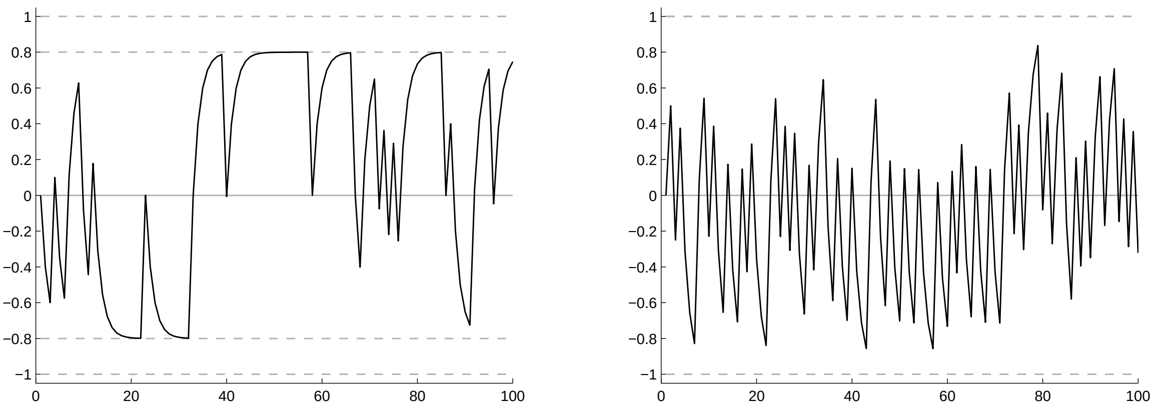


Figure 2: Simulation runs of $\widehat{\theta}_n$ for $\sigma = -0.1$ (left) and $\sigma = -1$ (right).

Proposition 2.

(a) If $\sigma > 0$, an informational cascade with herding almost surely occurs after finite time.

(b) If $\sigma = 0$, all agents are self-reliant with positive probability and there is no informational cascade but herding occurs after a finite time, see Çelen and Kariv (2004).

(c) If $\sigma \in (-2, 0)$, all agents are self-reliant with positive probability and neither an informational cascade nor herding occurs. In particular, if $\sigma = -0.5$, all agents are purely self-reliant.

(d) If $\sigma \leq -2$, an informational cascade with anti-herding occurs.

Proof. All claims follow directly from proposition 2 in section 3.2, with the exception of part (c). Let $\sigma \in (-2, 0)$. The boundaries of the learning process are given by its fixed points. The solution to the equation $\widehat{\theta} = \widehat{\theta}/2 - a_n(0.5+\sigma)$ is either $\widehat{\theta} = -(1 + 2\sigma)$ or $\widehat{\theta} = (1 + 2\sigma)$. Hence, $\widehat{\theta}_n \in (-(1 + 2\sigma), (1 + 2\sigma)) \subset [-1, 1]$ and therefore all agents are self-reliant with positive probability. □

Consider the case $\sigma \in (-0.5, 0)$. It is shown in the proof of proposition 2 that in this case $\widehat{\theta}_n \in (-(1+2\sigma), (1+2\sigma)) \subset [-1, 1]$ for all n. In particular, if σ is close to 0, the cutoff process is close to either 1 or -1 and switches occur only rarely (see figure 1, right). Therefore, agent's behavior is characterized by long spells of imitative behavior and only rarely (with probability close to σ) an agent receives a signal which is sufficiently strong to induce her to deviate from the previous run. By overturning a previous run, the deviator reveals an extreme contrary signal which induces a big jump in the cutoff process (see figure 1, right) and makes her successor close to indifferent between the two actions. It is easy to see that in the limit as N and n tend to infinity, the learning process never settles down because behavior overturns forever. This result contrasts with Çelen and Kariv (2004) who demonstrate that in the absence of strategic externalities, i.e. $\sigma = 0$, a herd must arise in finite time (intuitively, their result is based on the fact that the probability of overturning an ongoing run goes to zero quite rapidly). Hence, their finding is not robust in the sense that every arbitrarily small size of strategic substitutability causes behavior to overturn forever and therefore makes herding impossible.

An interesting phenomenon arises if $\sigma = -0.5$. In this case, strategic and informational externalities cancel each other out. Assume that some agent k observes the action a_n of some other agent $n \neq k$. This observation reveals that agent n's private signal has been in favor of alternative a_n, which has a positive effect on agent k's expected payoff of alternative a. In case $\sigma = -0.5$, the strategic substitutability imposed by this choice has an opposing effect of exactly the same magnitude and therefore the payoff of alternative a remains unaffected by the observation of agent n's choice. Therefore, despite being informative about the true state of the world, the observations of the behavior of others are worthless in this case and agents rely entirely on their private signals.

3.1.2 Imperfect Observability and Uniform Signals

In addition to her private signal θ_n, agent n observes only her immediate predecessor's action a_{n-1} before taking an action. The following proposition characterizes the corresponding learning process in terms of two cutoff levels, $\underline{\theta}_n$ and $\overline{\theta}_n$, depending on whether $a_{n-1} = -1$ or $a_{n-1} = 1$.

Proposition 3. *The learning process satisfies* $\widehat{\theta}_1 = 0$ *and*

$$\widehat{\theta}_{n+1} = \begin{cases} \overline{\theta}_{n+1} & \text{if } a_n = 1 \\ \underline{\theta}_{n+1} & \text{if } a_n = -1 \end{cases}$$

where $\underline{\theta}_{n+1} = -\overline{\theta}_{n+1}$,

$$\overline{\theta}_{n+1} = \begin{cases} 1 & \text{if } \overline{\theta}'_{n+1} > 1 \\ -1 & \text{if } \overline{\theta}'_{n+1} < -1 \\ \overline{\theta}'_{n+1} & \text{otherwise} \end{cases}$$

and $\overline{\theta}'_{n+1} = -\sigma - \frac{1}{2}\left(1 + \overline{\theta}^2_n\right)$.

Proof. See Appendix. □

Figure 3 illustrates the evolution of the cutoff levels $\underline{\theta}_n$ and $\overline{\theta}_n$ as specified in proposition 3. In case $\sigma > -0.5$ (left graph), $\underline{\theta}_n$ is larger than $\overline{\theta}_n$. Hence, whenever agent n receives a private signal within the interval $(\overline{\theta}_n, \underline{\theta}_n)$, she chooses the same action as her predecessor. This is the case because her private signal is too close to zero in order to be sufficiently informative to guide her decision. If, on the other hand, $\sigma < -0.5$ (right graph), $\underline{\theta}_n$ is smaller than $\overline{\theta}_n$ and whenever their private signal is within the interval $(\underline{\theta}_n, \overline{\theta}_n)$, agents choose the action opposite to their predecessor's action. This leads to the following definition.

Definition 4. Agent n is a *conformist* if $\overline{\theta}_n < \underline{\theta}_n$ and $\theta_n \in [\overline{\theta}_n, \underline{\theta}_n]$, and a *contrarian* if $\overline{\theta}_n > \underline{\theta}_n$ and $\theta_n \in [\underline{\theta}_n, \overline{\theta}_n]$. Agent n is a *pure* conformist (contrarian) if she is a conformist (contrarian) for all $\theta_n \in \Theta_S$.

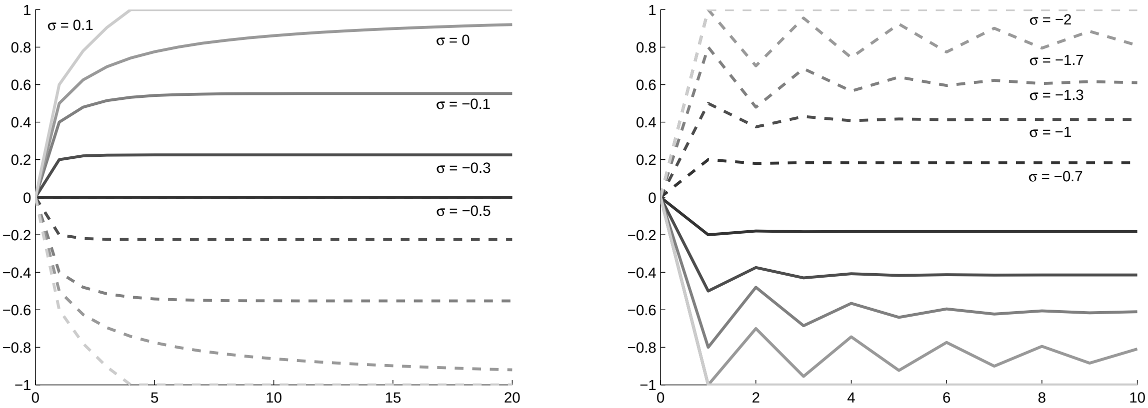


Figure 3: The cutoff-values $\overline{\theta}_n$ (dashed line) and $\underline{\theta}_n = -\overline{\theta}_n$ (straight line) as a function of n (for different σ).

Figure 3 suggests that in case $\sigma \in (-2, 0)$, all cutoff values $\underline{\theta}_n$ and $\overline{\theta}_n$ are strictly within the interval $(-1, 1)$ (with the exception of agent $n = 1$ in case σ is close to -2) and therefore all agents are self-reliant with positive probability, provided that their private signals are sufficiently close to either 1 or -1. Hence, an informational cascade does not arise in this case. As σ gets close to either -2 or 0, conformist, resp. contrarian, behavior becomes increasingly likely because the interval $[\overline{\theta}_n, \underline{\theta}_n]$, resp. $[\underline{\theta}_n, \overline{\theta}_n]$, approaches the entire signal space. Beyond these thresholds, agents always disregard their private information and select their optimal action entirely on the grounds of their predecessor's action, i.e. agents are *pure* conformists in the former case and *pure* contrarians in the latter. An informational cascade always arises in this case due to herding, resp. anti-herding. Note that, like under perfect observability, strategic and informational externalities cancel each other out if $\sigma = -0.5$. It is easy to see that behavior must be independent of the observational structure in this case. The following proposition summarizes.

Proposition 4.

(a) If $\sigma < -0.5$ ($\sigma > -0.5$), all agents are contrarians (conformists) with positive probability and there are no conformists (contrarians).

(b) If $\sigma > 0$, there exists some agent $n_0 \geq 2$ such that all agents $n \geq n_0$ are pure conformists and an informational cascade with herding occurs.

(c) If $\sigma \in (-2, 0]$, all agents are self-reliant with positive probability and neither an informational cascade nor herding occurs. In particular, if $\sigma = -0.5$, all agents are purely self-reliant.

(d) *If $\sigma \leq -2$, all agents are pure contrarians and an informational cascade with anti-herding occurs.*

Proof. Let me use the fact that $\overline{\theta}_n = -\underline{\theta}_n$ for all n (see proposition 3).

part (a): Let $x := -\sigma - 0.5 > 0$ and let $x_m := \min\{x, 1\}$. First, I show that $\overline{\theta}_n > 0 \, \forall n \geq 2$. Assume that $\overline{\theta}_n \in (0, x_m)$. This implies that $\overline{\theta}_{n+1} = \min\{x - \overline{\theta}_n^2/2, 1\} \in (0, x_m)$ and the claim follows by induction since $\overline{\theta}_1 = 0$, $\overline{\theta}_2 = x_m$ and $\overline{\theta}_3 = \min\{x - x^2/2, 1\} \in (0, x_m)$. The second part of the proof works in the same way.

part (b): See the general proof of proposition 9.

part (c): In case $\sigma \in (-2, 0)$, it can be shown (cf proof of proposition 5) that $\underline{\theta}_n \to -1 + \sqrt{-2\sigma}$ and $\overline{\theta}_n \to 1 - \sqrt{-2\sigma}$ as $N \to \infty$ and $n \to \infty$. Hence all agents are self-reliant with positive probability. In order to show that all agents are purely self-reliant if $\sigma = -0.5$, it is sufficient to show that $\overline{\theta}_n = 0$ $\forall\, n$. This follows by induction since $\overline{\theta}_n = 0$ implies that $\overline{\theta}_{n+1} = -\sigma - 1/2 = 0$ $\forall\, n$ and since $\overline{\theta}_1 = 0$. In the special case $\sigma = 0$, the interval $[\overline{\theta}_n, \underline{\theta}_n]$ converges to the interval $[-1, 1]$. Nevertheless, a herd does not exist as shown by Çelen and Kariv (2004).

part (d): It can be shown via direct calculation that in case $\sigma \leq -2$, $\overline{\theta}_n = 1$ for all $n \geq 2$ which implies an informational cascade with anti-herding. □

The question of whether herding occurs in the limiting case without strategic externalities is answered by Çelen and Kariv (2001). In this case, $\overline{\theta}_n$ monotonically increases in n and the interval $[\underline{\theta}_n, \overline{\theta}_n]$ converges to the entire signal space $[-1, 1]$ (in the limit as N and n tend to infinity). Çelen and Kariv (2001) show that the cutoff process $\{\widehat{\theta}_n\}$ as defined in proposition 3 does not converge in this case.[2] Since divergence of cutoffs implies divergence of actions, standard herd behavior is impossible, even though the expected length of runs is increasing and goes to infinity. In contrast to this result, it is shown here that if $\sigma > 0$, the interval $[\underline{\theta}_n, \overline{\theta}_n]$ reaches the entire signal space after some *finite* number of agents and therefore an informational cascade and herd behavior arise after finite time. Since this is the case for any arbitrarily small size of strategic complementarities, the above analysis shows that the no-herding result of Çelen and Kariv (2001) is not robust in this sense.

As shown by the proposition, if strategic externalities are within the interval

[2]For any agent k, it can be shown that, as $N \to \infty$, the probability that all of agent k's successors take the same action is equal to zero, i.e. $\Pi_{n=k}^{\infty} \frac{1-\overline{\theta}_n}{2} = 0$, see corollary 5 in Çelen and Kariv (2001).

$(-2,0)$, neither an informational cascade nor (anti-)herding occur because agents are self-reliant with strictly positive probability. As noted before, agents become increasingly likely to engage in conformist or contrarian behavior if σ approaches the boundaries of this interval and therefore one expects behavior to come close to cascade and (anti-)herding behavior at these boundaries. The following proposition supports this expectation with analytical expressions. In particular, the expected length of a run is calculated as a function of σ. This expression tends to 0 as σ approaches -2 (from above) due to an increasing tendency towards anti-herding. As σ approaches 0 (from below), the expected length of a run tends to $+\infty$ and hence approaches herding behavior in the limit (see figure 4). Similarly, the proposition shows that behavior approaches cascade behavior at the boundaries in the sense that the expected number of consecutive contrarians (conformists) approaches $+\infty$ as $\sigma \to -2$ (as $\sigma \to 0$). Moreover, as σ approaches -0.5, runs become increasingly short and entirely disappear in case $\sigma = -0.5$ since all agents are entirely self-reliant (see figure 4).

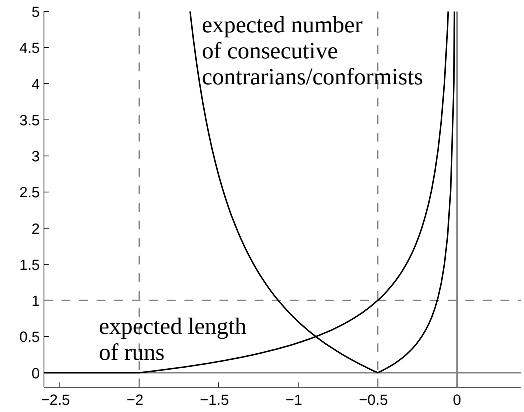


Figure 4: Behavior approaches cascade and (anti-)herding behavior as σ approaches the boundaries of the interval $[-2,0]$.

Proposition 5.

(a) If $\sigma \in (-2,0)$, $\underline{\theta}_n \to 1-\sqrt{-2\sigma}$ as $n \to \infty$ and the expected length of a run starting at agent n tends to $\frac{2}{\sqrt{-2\sigma}} - 1$.

(b) In case $\sigma \in (-2,-0.5]$, the expected number of consecutive contrarians following agent n tends to $\frac{\sqrt{-2\sigma}-1}{2-\sqrt{-2\sigma}}$ as $n \to \infty$. In case $\sigma \in (-0.5,0)$ the expected number of consecutive conformists following agent n tends to $\frac{1}{\sqrt{-2\sigma}} - 1$ as $n \to \infty$.

Proof. part (a): The limit point $1-\sqrt{-2\sigma}$ of the sequence $\{\underline{\theta}_n\}$ is given by the (feasible) root of the equation $\underline{\theta}_{n+1}-\underline{\theta}_n = 0$. Furthermore, the probability for agent n to choose action $a_n = a_{n-1}$ is equal to $\Pr(\theta_n \in (\overline{\theta}_n, 1)) = (1-\overline{\theta}_n)/2$ and approaches $1-\sqrt{-2\sigma}/2$ as $n \to \infty$. Hence, the distribution of the length of a run following agent n approaches a geometric distribution with parameter $p = \sqrt{-2\sigma}/2$ and expected value $(1-p)/p = \frac{2}{\sqrt{-2\sigma}} - 1$.

part(b): In case $\sigma \in (-2, -0.5]$, the probability that agent n is a contrarian is equal to $(\overline{\theta}_n - \underline{\theta}_n)/2$ and approaches $\sqrt{-2\sigma}-1$ as $n \to \infty$. Hence, the distribution of the number of successive contrarians approaches a geometric distribution with parameter $p = 2-\sqrt{-2\sigma}$ and expected value $(1-p)/p = \frac{\sqrt{-2\sigma}-1}{2-\sqrt{-2\sigma}}$. Similarly, in case $\sigma \in (-0.5, 0)$, the number of successive conformists approaches a geometric distribution with parameter $p = \sqrt{-2\sigma}$ and expected value $(1-p)/p = \frac{1}{\sqrt{-2\sigma}} - 1$. □

Figure 5 summarizes the findings of this section. Remember from definition 2 that herding and anti-herding are *complete* if *all* agents $n \neq 1$ engage in it. Anti-herding is always complete and it is easy to show that complete herding occurs beyond $\sigma = 0.5$.

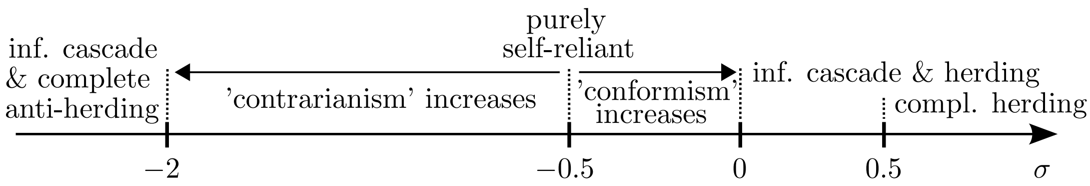


Figure 5: The relationship between informational cascades, (anti-)herds and the size of strategic externalities.

Propositions 2 and 4 suggest that the informational structure does not affect the long-run outcome of behavior whereas strategic externalities do. Only in the absence of strategic externalities, the long-run outcome is affected by the informational structure. This shows that the main finding of Çelen and Kariv (2004) strongly depends on the absence of strategic externalities. The following corollary summarizes.

Corollary 1. *For any $\sigma \neq 0$, an informational cascade (a herd, an anti-herd) occurs under perfect observability if and only if it occurs under imperfect observability. This property is violated in case $\sigma = 0$.*

3.2 General Signals

In this section, I drop the assumption of a uniform signal distribution and allow for a general symmetric distribution characterized by its continuous and differentiable c.d.f. F over the finite support $\Theta_S \equiv [-b, b]$. Let $\mathrm{E}^+(x)$ denote the expected value of the signal θ, conditional on being larger than some number x, i.e. $\mathrm{E}^+(x) = \int_x^b \theta dF(\theta)$.

3.2.1 Perfect Observability and General Signals

Assume perfect observability, i.e. $\mathcal{I}_n = \mathcal{I}_n^{PO}$. Proposition 6 specifies a law of motion of the learning process and proposition 7 describes its properties. The most important properties of the uniform case carry over to the general case, such as the existence of a σ under which both types of externalities cancel each other out, the occurrence of an informational cascade with herding in the presence of positive σ as well as the occurrence of an informational cascade with anti-herding in the presence of negative (and sufficiently small) σ. The main difference is that under a general signal distribution, some agents may be pure imitators or pure contrarians even if strategic externalities are moderate, i.e. $\sigma \in (-2b, 0)$. However, proposition 7 shows that an informational cascade or (anti-)herding is nevertheless impossible in this case.

Proposition 6. *The learning process satisfies $\widehat{\theta}_1 = 0$ and*

$$\widehat{\theta}_{n+1} = \begin{cases} b & \text{if } \widehat{\theta}'_{n+1} > b \\ -b & \text{if } \widehat{\theta}'_{n+1} < -b \\ \widehat{\theta}'_{n+1} & \text{otherwise} \end{cases}$$

where $\widehat{\theta}'_{n+1} = \widehat{\theta}_n - a_n \left[E^+ \left(a_n \cdot \widehat{\theta}_n \right) + \sigma \right]$.

Proof. Since all the information revealed by the history $(a_i)_{i=1}^{n-1}$ is already contained in $\widehat{\theta}_n$, $\widehat{\theta}_{n+1}$ is altered only by the new information revealed through a_n, i.e. $\widehat{\theta}_{n+1} = \widehat{\theta}_n - \mathrm{E}\left(\theta_n + \sigma a_n | \widehat{\theta}_n, a_n\right)$. The claim follows from the fact that $\mathrm{E}\left(\theta_n | \widehat{\theta}_n, 1\right) = \mathrm{E}^+(\widehat{\theta}_n)$ and $\mathrm{E}\left(\theta_n | \widehat{\theta}_n, -1\right) = -\mathrm{E}^+(-\widehat{\theta}_n)$. □

Proposition 7.

(a) If $\sigma > 0$, an informational cascade with herding almost surely occurs after finite time.

(b) *If* $\sigma = 0$*, all agents are self-reliant with positive probability and there is no informational cascade but herding occurs after a finite time, see Smith and Sorenson (2000).*

(c) *If* $\sigma \in (-2b, 0)$*, neither an informational cascade nor herding occurs. If* $\sigma = -E^{+}(0)$*, all agents are purely self-reliant.*

(d) *If* $\sigma \leq -2b$*, an informational cascade with anti-herding occurs.*

Proof. part (a): Consider any agent n_0 and let $\widehat{\theta}_{n_0} \geq 0$. For $n' = \lceil \frac{b}{\sigma} \rceil$, the probability that $\widehat{\theta}_{n_0+n'} = b$ is at least $p = \left(\frac{1}{2}\right)^{n'}$. Since the symmetric argument applies for $\widehat{\theta}_{n_0} < 0$, the claim follows from

$$\Pr(\min\{n : \widehat{\theta}_n \in \{-b, b\}\} = \infty) \leq \lim_{i \to \infty} (1-p)^i = 0.$$

part (c): It is sufficient to prove that for any agent n_0, there exists some agent $n \geq n_0$ such that $\widehat{\theta}_n \in (-b, b)$. Consider some $n \geq n_0$ and suppose that $\widehat{\theta}_n \notin (-b, b)$, i.e. either $\widehat{\theta}_n = -b$ or $\widehat{\theta}_n = b$. In this case either $\widehat{\theta}_{n+1} = -b-\sigma \in (-b, b)$ or $\widehat{\theta}_{n+1} = b+\sigma \in (-b, b)$, respectively. Moreover, $\sigma = -\mathrm{E}^{+}(0)$ implies $\widehat{\theta}_n = 0$ for all n.

part (d): It is easy to show that in case $\sigma \leq -2b$, $\widehat{\theta}_2 \in \{-b, b\}$ and $\widehat{\theta}_{n+1} = -\widehat{\theta}_n$ for all $n \geq 2$, which implies the claim. □

3.2.2 Imperfect Observability and General Signals

Assume imperfect observability, i.e. $\mathcal{I}_n = \mathcal{I}_n^{IO}$. Proposition 8 specifies a law of motion of the learning process and proposition 9 describes its properties. Like in the perfect observability case, the most important properties of the uniform case carry over to the general case.

Proposition 8. *The learning process satisfies* $\widehat{\theta}_1 = 0$ *and*

$$\widehat{\theta}_{n+1} = \begin{cases} \overline{\theta}_{n+1} & \text{if } a_n = 1 \\ \underline{\theta}_{n+1} & \text{if } a_n = -1 \end{cases}$$

where $\underline{\theta}_{n+1} = -\overline{\theta}_{n+1}$*,*

$$\overline{\theta}_{n+1} = \begin{cases} b & \text{if } \overline{\theta}'_{n+1} > b \\ -b & \text{if } \overline{\theta}'_{n+1} < -b \\ \overline{\theta}'_{n+1} & \text{otherwise} \end{cases}$$

and $\overline{\theta}'_{n+1} = \overline{\theta}_n - 2F(\overline{\theta}_n)\left[\overline{\theta}_n + E^{+}(-\overline{\theta}_n)\right] - \sigma$*.*

Proof. If can be shown that even in the case of a general signal distribution $P(a_n = 1) = 0.5$ and $\overline{\theta}_n + \underline{\theta}_n = 0$ is satisfied for all n (see Çelen and Kariv (2001) for a similar proof). Moreover,

$$\begin{aligned}\overline{\theta}'_{n+1} &= \left[1 - F(\overline{\theta}_n)\right]\left[\overline{\theta}_n - E^+(\overline{\theta}_n) - \sigma\right] + \left[1 - F(\underline{\theta}_n)\right]\left[\underline{\theta}_n - E^+(\underline{\theta}_n) - \sigma\right] \\ &= \overline{\theta}_n - 2\left[F(\overline{\theta}_n)\overline{\theta}_n + (1 - F(\overline{\theta}_n))\mathrm{E}^+(\overline{\theta}_n)\right] - \sigma. \qquad (3)\end{aligned}$$

The last transformation uses the fact that $1 - F(\underline{\theta}_n) = F(-\underline{\theta}_n) = F(\overline{\theta}_n)$. Finally, the symmetry of the signal distribution implies that

$$(1 - F(\theta))\mathrm{E}^+(\theta) = F(\theta)\mathrm{E}^+(-\theta).$$

Hence,

$$\overline{\theta}'_{n+1} = \overline{\theta}_n - 2F(\overline{\theta}_n)\left[\overline{\theta}_n + \mathrm{E}^+(-\overline{\theta}_n)\right] - \sigma. \qquad (4)$$

□

Proposition 9.

(a) If $\sigma > 0$, there exists some agent $n_0 \geq 2$ such that all agents $n \geq n_0$ are pure conformists and an informational cascade with herding occurs. Herding is complete if $\sigma \geq b - E^+(0)$.

(b) If $\sigma \in [-b, 0]$, all agents are self-reliant with positive probability. There is no informational cascade and no (anti-)herding. If $\sigma = -E^+(0)$, all agents are purely self-reliant.

(c) If $\sigma < -b$, all agents are either contrarians or self-reliant. If $|\sigma|$ is sufficiently large, all agents $n \geq 2$ are pure contrarians and an informational cascade with complete anti-herding occurs.

Proof. In the following, I only show that the claims hold for $\overline{\theta}_n$. The full claims easily follow from $\underline{\theta}_n = -\overline{\theta}_n$.

part (a): It needs to be shown that $\exists\, n_0$ s.t. $\overline{\theta}_n = -b$ and $\underline{\theta}_n = b$ $\forall\, n > n_0$. Plugging the inequality $\mathrm{E}^+(-\overline{\theta}_n) > -\overline{\theta}_n$ into equation (4) gives $\overline{\theta}'_{n+1} < \overline{\theta}_n - \sigma$. Hence $\overline{\theta}_n = -1$ after finitely many steps. Complete herding requires that $\theta_2 = \theta_3 = -1$ which holds if $\sigma \geq b - \mathrm{E}^+(0)$.

part (b): It needs to be shown that $\overline{\theta}_n, \underline{\theta}_n \in (-b, b)\ \forall\, n$. Using $\mathrm{E}^+(\overline{\theta}_n) > \overline{\theta}_n$ in both equations (3) and (4) gives $\overline{\theta}_{n+1} < \min\{\overline{\theta}_n, -\overline{\theta}_n\} - \sigma$. Since $\sigma \geq -b$ it follows that $\overline{\theta}_{n+1} < b$. On the other hand, equation (4) and the inequality $\mathrm{E}^+(-\overline{\theta}_n) < b$ imply $\overline{\theta}_{n+1} > -b - \sigma \geq -b$ (since $\sigma \leq 0$). Hence, $|\overline{\theta}_n| < b\ \forall n$ which implies the impossibility of an informational cascade. See Çelen and Kariv (2004) for the proof in the special case $\sigma = 0$. In case $\sigma = -\mathrm{E}^+(0)$, it is easy to check that $\overline{\theta}_n = 0$ implies $\overline{\theta}_{n+1} = -\mathrm{E}^{+}(0) - \sigma = 0$ (since $F(0) = 1/2$). Hence, the claim follows by induction since $\overline{\theta}_1 = 0$.

part (c): It needs to be shown that $\overline{\theta}_n > 0$ and $\underline{\theta}_n < 0$. This follows from part (b) since $\overline{\theta}_n > -b - \sigma > 0$. Moreover, it immediately follows from equation (3) that $\overline{\theta}'_{n+1} > b$ for all $n > 1$ provided that $|\sigma|$ is sufficiently large. □

4 Binary Signals

Let $N = \infty$. The space of states of the world as well as the space of agents' private signals are binary, i.e. $\Theta \equiv \Theta_S \equiv \{-1, 1\}$.[3] The prior distribution of θ_0 is symmetric and common knowledge. The precision of each agent n's signal θ_n satisfies $\Pr(\theta_n = 1 \,|\, \theta_0 = 1) = \Pr(\theta_n = -1 \,|\, \theta_0 = -1) = q > 0.5$. Assuming perfect observability, i.e. $\mathcal{I}_n = \mathcal{I}_n^{PO}$, the optimality condition (1) becomes:

$$a_n = 1 \text{ iff } \mathrm{E}\left(\theta_0 \,|\, \mathcal{I}_n^{PO}\right) \geq -\sigma \sum_{i=1}^{n-1} a_i.$$

Let p_n denote agent n's (private) belief about the likelihood of the event $\theta_0 = 1$, conditional on her information $\mathcal{I}_n^{PO}$. Hence, $\mathrm{E}(\theta_0 | \mathcal{I}_n^{PO}) = 2p_n - 1$ and the optimal decision rule of agent n can be rewritten in the following way:

$$a_n = 1 \text{ iff } p_n \geq \widehat{p}_n, \text{ where } \widehat{p}_n = 0.5(1 - \sigma \sum_{i=1}^{n-1} a_i).$$

$\{\widehat{p}_n\}_{n \geq 1}$ is a stochastic process which satisfies $\widehat{p}_1 = 0.5$ and its dynamics can be summarized by:

$$\widehat{p}_{n+1} = \widehat{p}_n - \frac{\sigma}{2} a_n. \tag{5}$$

This process specifies cutoff values of agent's private beliefs and is therefore referred to as the *cutoff process*. Furthermore, let b_n denote agent n's belief *before* she observes θ_n and let $b_n^{\theta_n}$ denote her belief after having observed θ_n, i.e. $p_n = b_n^{\theta_n}$. Bayesian updating implies $b_n^1 = \frac{q b_n}{q b_n + (1-q)(1-b_n)}$ and $b_n^{-1} = \frac{(1-q) b_n}{(1-q) b_n + q(1-b_n)}$. I refer to $\{b_n\}_{n \geq 1}$ as the process of *public beliefs* and I denote its state space by $B = \{B_i\}_{i=-\infty}^{+\infty}$. It is straightforward to show that

$$B_i = \left(1 + \left(\frac{1-q}{q}\right)^i\right)^{-1}. \tag{6}$$

[3] In terms of the basic model presented in section 2, this describes the special case where either $Q_1 = -Q_{-1} = 1/2$ or $Q_1 = -Q_{-1} = -1/2$ is satisfied in either of the two states of nature.

Let me first characterise the behavior of the first two agents before proceeding to a more general analysis. Agent one follows her private signal, i.e. $a_1 = \theta_1$, since $\widehat{p}_1 = 0.5$. For the sake of illustration, assume that $a_1 = \theta_1 = 1$ (in the other case, the inverse reasoning holds). Agent two observes a_1, correctly deduces θ_1 and updates her belief via Bayes' rule to $b_2 = q$. Upon observing her private signal and applying Bayes' rule for a second time, agent two's private belief either becomes $p_2 = b_2^1 = q^2/(q^2 + (1-q)^2)$ (if $\theta_2 = 1$) or $p_2 = b_2^{-1} = 0.5$ (if $\theta_2 = -1$). Her optimal behavior as a function of q and σ can accordingly be derived from her cutoff-value $\widehat{p}_2 = 0.5(1-\sigma)$. In case $\theta_2 = 1$, one obtains $a_2 = 1$ iff $\sigma \geq 1 - 2q^2/(q^2 + (1-q)^2)$ and in case $\theta_2 = -1$, one obtains $a_2 = 1$ iff $\sigma \geq 0$. Figure 6 summarizes agent two's behavior for different combinations of σ and q. Note that if $\theta_2 = -\theta_1$, signals cancel each other out and the optimal choice of agent two depends on whether strategic externalities are positive or negative. Hence, if $\sigma > 0$, agent two always imitates the action of agent one, irrespectively of her own signal (i.e. she is a *pure conformist* in the spirit of section 4). If, however, σ is negative and agent two observes a private signal which is equal to the action of agent one, i.e. $\theta_2 = a_1$, her optimal choice depends on the tradeoff between the positive informational externality of two signals in favor of action one against the strategic substitutability induced by agent one's action. Hence, this tradeoff depends on the relative magnitudes of q and $|\sigma|$. For any q, agent two chooses the action opposition to agent one's action, provided that σ is sufficiently small (i.e. she is a *pure contrarian* in the spirit of section 4). It is easy to check that the parameter combinations under which agent two reveals her private information, i.e. under which she is *purely self-reliant*, is given by $\{(\sigma, q) : 1 - 2q^2/(q^2 + (1-q)^2 < \sigma < 0\}$, see figure 6. In the following, I analyze the learning process more generally, by studying the behavior of agents beyond agent two.

First, consider the case $\sigma > 0$.[4] Agent one chooses the action that equals her signal, i.e. $a_1 = \theta_1$ and agent two chooses $a_2 = a_1$, since strategic complementarities provide a strict incentive to imitate (see above). Hence, agent two does not reveal her private information. Consequently, agent three finds herself in the same position as agent two and it follows by induction that no subsequent agent ever reveals her private signal. The following proposition summarizes this result.

Proposition 10. *If $\sigma > 0$, only agent one reveals her private information and all subsequent agents follow her action. In other words, there exists an informational cascade with herding.*

[4] I only study the two cases $\sigma > 0$ and $\sigma < 0$. For the case $\sigma = 0$ see standard models such as Banerjee (1992) and Bikhchandani et al. (1992).

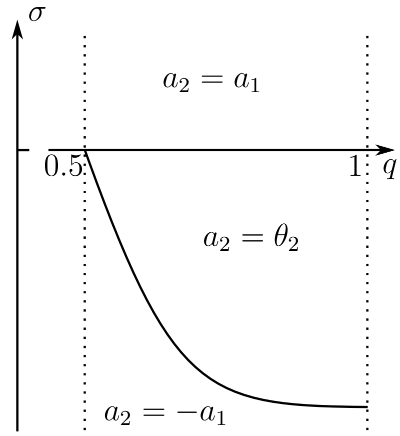


Figure 6: For different combinations of q and σ, agent two is either a pure conformist, purely self-reliant or a pure contrarian.

In the case of strategic complementarities, the learning process 'gets stuck' in a herd after the first agent. If $\sigma < 0$, a herd can not occur since negative externalities reduce the expected payoff of every new agent who joins a ongoing run and eventually the contrary action yields a higher expected payoff. Specifically, learning happens in two distinct ways. Either by some agent who breaks away from the uniform behavior of her predecessors and thereby reveals a contrary signal, or by an agent who deliberately imitates her predecessors despite the negative payoff effects of their uniform behavior and thereby reveals a signal that favors their actions. Hence, learning is necessarily more accurate in the case of strategic substitutabilities because it does not stop after agent one. However, I will show that in this case, an informational cascade occurs after some finite (and random) point in time and behavior settles down in an anti-herd. The following lemma characterizes the corresponding limit points of learning in terms of the public belief process $\{b_n\}_n$.

Lemma 1. *Let $\sigma < 0$. There exist two absorbing states $b^*_{-1}, b^*_1 \in (0,1)$ of $\{b_n\}_n$ that satisfy $b^*_{-1} + b^*_1 = 1$ and $b^*_{-1} < b^*_1$. $\{b_n\}_n$ does not reach any other absorbing states.*

Proof. Define $b^*_1 = \min\{b \in B \,:\, b^{-1} > 0.5 + j\frac{|\sigma|}{2}$ and $b^1 < 0.5 + (j+1)\frac{|\sigma|}{2}\}$ and $b^*_{-1} = 1 - b^*_1$. Note that b^*_{-1} and b^*_1 clearly exist since $|b^{-1} - b^1|$ goes to 0 as b approaches either 0 or 1 and since $|\sigma|$ is constant. Moreover, from equation (6) it is easy to check that $B_i - 0.5 = 0.5 - B_{-i}$ for all $i > 0$, which implies that $b^*_{-1} + b^*_1 = 1$ and $b^*_{-1} < b^*_1$. In order to show that b^*_{-1} and b^*_1 are absorbing states of $\{b_n\}_n$ let $b_n = b^*_{-1}$ or $b_n = b^*_1$. Since, by definition, there is no $j \in \mathbb{Z}$ that satisfies $0.5 + j\frac{|\sigma|}{2} \in (b^{-1}, b^1)$, it follows that $\widehat{p}_n \notin (b^{-1}, b^1)$ and therefore the private signals of agent n and any of her successors will not be revealed. Hence, $\{b_n\}_n$ remains constant forever and $\sigma < 0$ implies that actions satisfy $a_n = -a_{n+1}$. Finally, by definition of b_1^*, none of the states $b \in (b_{-1}^*, b_1^*)$ are absorbing. □

The two plots in figure 7 illustrate convergence to the absorbing state b_1^*. Once this state is reached, an informational cascade with anti-herding occurs (as shown by the proof of lemma 1). The cutoff-process $\{\widehat{p}_n\}_n$ fluctuates around the interval (b_n^{-1}, b_n^1) but b_n is constant since no agent reveals her private information.

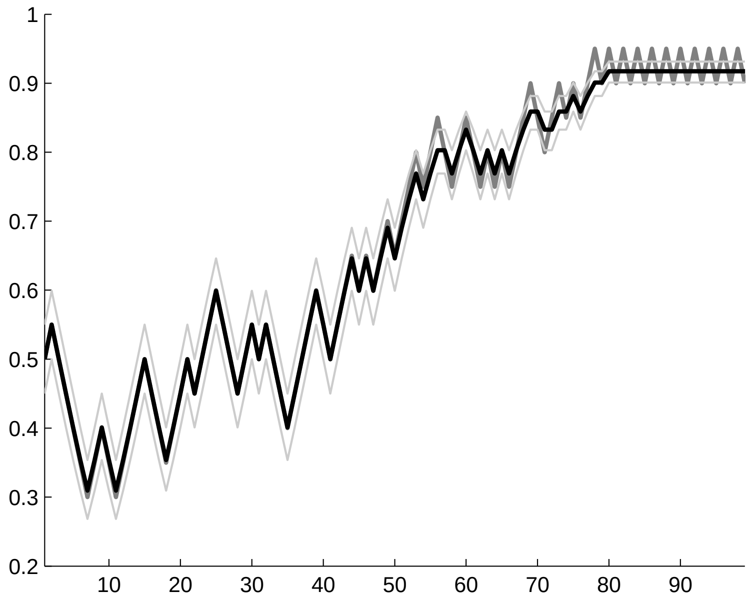
Lemma 1 states that there are two possible outcome of social learning, only one of which is close to the true state of the world θ_0. Even though θ_0 is never learned with certainty, I show below (lemma 3) that for certain parameter configurations, the public belief will eventually be very close to either one or zero. However, this does not exclude the possibility that social learning settles in near the 'wrong' state.

Definition 5 defines the notion of correct learning and lemma 2 provides two sufficient conditions for the probability of correct learning to be close to one.

Definition 5. Learning is *correct* if $\{b_n\}$ reaches the absorbing state $b_{\theta_0}^*$. $P_{\sigma,q}$ denotes the *ex ante* probability that learning is correct.

Lemma 2. *Let $\sigma < 0$ and consider any $\varepsilon > 0$. If $q > 0.5$ is fixed, $P_{\sigma,q} > 1-\varepsilon$ if $|\sigma|$ is sufficiently small. The same is true if σ is fixed and q is sufficiently close to one.*

Proof. To keep notation simple, I assume that $\theta_0 = 1$ (the proof in the other case is symmetric). Let me first prove that for all $b_n \in (b_{-1}^*, b_1^*)$ and $\widehat{p}_n$, there exists some finite integer $\delta \geq 1$ such that $\Pr(b_{n+\delta} = b_n^1) = q$, $\Pr(b_{n+\delta} = b_n^{-1}) = 1 - q$ and $b_{n+i} = b_n$ for all $i = 0, .., \delta - 1$. The claim follows immediately if $\widehat{p}_n \in [b_n^{-1}, b_n^1)$ (where $\delta = 1$) because in this case agent n chooses the action that equals her signal. Furthermore, let me show that if the condition $\widehat{p}_n \in [b_n^{-1}, b_n^1)$ is not satisfied, it will nevertheless hold after a finite and deterministic number of steps. Assume that $b_n > 0.5$ (the proof for the case $b_n < 0.5$ is symmetric) and let $\widehat{p}_n \notin [b_n^{-1}, b_n^1)$. Agent n does not reveal her action and therefore $b_{n+1} = b_n$. Moreover, $\widehat{p}_{n+1}$ approaches the interval $[b_n^{-1}, b_n^1)$ by a fixed step size of $|\sigma|/2$ (see equation 5). Since $b_n \in (b_{-1}^*, b_1^*)$ implies that there exists some $j > 0$ such that $0.5 + j\frac{\sigma}{2} \in [b_n^{-1}, b_n^1]$ and since the step size is constant, it follows by iteration that $\{\widehat{p}_n\}$ eventually (after δ steps) enters the interval $[b_n^{-1}, b_n^1)$. Depending on the signal of agent $n + \delta$, either $b_{n+\delta} = b_n^1$ (with probability q) or $b_{n+\delta} = b_n^{-1}$ (with probability $1 - q$).



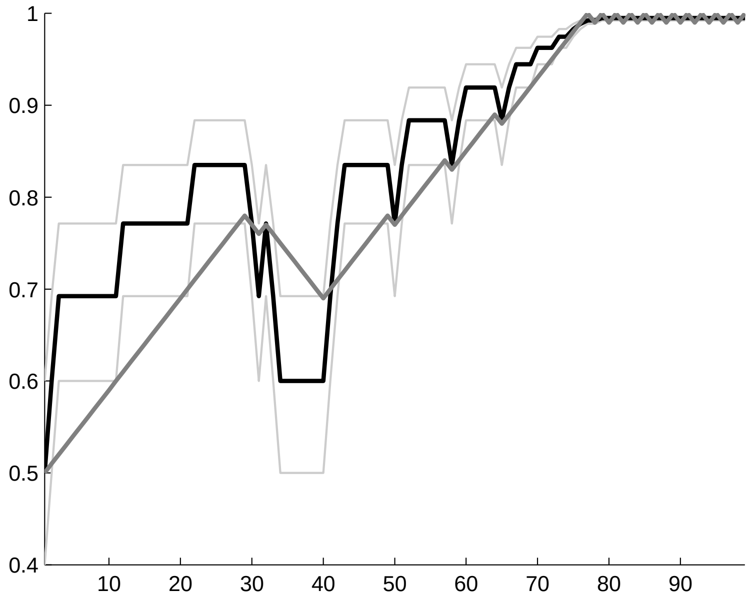


Figure 7: Two simulation runs of the processes $\{b_n\}$ (black), $\{\widehat{p}_n\}$ (dark grey), $\{b_n^{-1}\}$ (light grey) and $\{b_n^1\}$ (light grey) for $\sigma = -0.1$ (first plot) and $\sigma = -0.02$ (second plot).

Let k^* denote the positive integer that satisfies $b_1^* = B_{k^*}$. It is easy to check that $k^* \to \infty$ if either $q \to 1$ (for fixed σ) or $|\sigma| \to 0$ (for fixed q), since $B_i - B_{i-1}$ monotonically approaches zero (in i) and since $B_i \to 1$ as $q \to 1$ $\forall i > 0$, resp. $B_i \to 0$ $\forall i < 0$. Let the random sequence $\mathcal{T} = \{n_1, n_2, ..., n^*\}$ denote the positions at which the public belief changes (i.e. $n \in \mathcal{T} \Rightarrow \widehat{p}_{n-1} \in (b_n^{-1}, b_n^1)$). The first part of the proof implies that all $n_{i+1} - n_i$ are finite. Let me define a stochastic process $\{\mathcal{B}_t\}_{t=1}^{\infty}$ by

$$\mathcal{B}_t = \begin{cases} b_{n_t} & \text{if } t \leq n^* \\ \mathcal{B}_{t-1} & \text{if } t > n^* \end{cases}$$

$\{\mathcal{B}_t\}$ is a simple Markov Process with finite state space $B \cap [b_n^{-1}, b_n^1]$ and transition probabilities $\Pr(B_i \to B_{i+1}) = 1 - \Pr(B_i \to B_{i-1}) = q$ and $\Pr(b_1^* \to b_1^*) = \Pr(b_{-1}^* \to b_{-1}^*) = 1$. Hence, the problem is equivalent to the 'Gambler's Ruin Problem' which states that $P_{\sigma,q} = \Pr(\exists\, t > 0 \,:\, \mathcal{B}_t = b_1^*) = (1 + m^{k^*})^{-1}$, where $m = (1-q)/q$. Hence, $P_{\sigma,q} \to 1$ is implied by either $q \to 1$ or $|\sigma| \to 0$. □

The intuition of the proof of lemma 2 is the following. The social learning process can be rewritten as a Markov process which 'jumps' towards the correct state with probability q and away from it with probability $1-q$. If q is large, the outcome of the process is therefore more likely to be correct. On the other hand, the smaller $|\sigma|$, the more agents find it optimal to herd on the same action. In other words, it takes more steps for the learning process to reach one of the absorbing states. This implies that learning is correct with higher probability (the solution to the problem of convergence is equivalent to the solution of a special version of the *Gambler's Ruin Problem*[5]).

As noted before, an informational cascade always occurs after finite time. However, the following proposition implies that the learning process may nevertheless end up very close to either zero or one. This is the case because the expected start of a cascade goes to infinity as $|\sigma|$ tends to zero. The lemma establishes this result via a simple relation between the precision of learning and the probability of correct learning.

Lemma 3. *$b_{\theta_0}^* = P_{\sigma,q}$ and $b_{-\theta_0}^* = 1 - P_{\sigma,q}$.*

Proof. Let k^* denote the positive integer that satisfies $b_1^* = B_{k^*}$ and let $m = (1-q)/q$. It is a well known fact that $P_{\sigma,q} = (1 - m^{k^*})/(1 - m^{2k^*})$,

[5] A special version of the Gambler's Ruin Problem can be describes as follows. A gamber is offered an (infinite) sequence of bets. In each bet she gains one unit of money with probability q and looses one unit with probability $1-q$. Starting out with a fortune of size M, one can determine the probability that her fortune reaches $2M$ before she goes broke.

provided that $q \neq 0.5$ (this is the solution to a version of the Gambler's Ruin Problem in which the gambler seeks to double her fortune, cf. footnote 5). From the fact that the denominator can be rewritten as $(1-m^{k^*})\cdot(1+m^{k^*})$, it easily follows that $P_{\sigma,q} = B_{k^*} = 1/(1+m^{k^*})$. □

The following proposition summarizes the results of this section.

Proposition 11. *If $\sigma < 0$, the outcome of social learning is arbitrarily close to the true state of the world with arbitrarily high probability, provided that either $|\sigma|$ is sufficiently small (for given q) or q is sufficiently close to one (for given σ).*

Proof. The claim is implied by lemma 2 and lemma 3. □

Proposition 11 states that in the limit as $|\sigma|$ tends to 0, the true state of the world is learned with probability one. This contrasts with the standard herding models without strategic externalities such as the ones in Banerjee (1992) and Bikhchandani et al. (1992). Table 8 shows explicit values of $b_1^* = P_{\sigma,q}$ for different parameter combinations, provided that the true state of the world satisfies $\theta_0 = 1$.[6]

Note, however, that if $|\sigma|$ and q are both small, learning may take a long time to reach one of the absorbing states. For example, in case $q = 0.51$ and $\sigma = -0.01$, one obtains an average time of convergence of around 3555 periods.[7] Other parameter combinations produce considerably faster convergence (e.g. around 10.5 periods for $q = 0.7$ and $\sigma = -0.2$, despite the fact that the probability that learning is correct is as high as 96.7%), see figure 9.

$q \mid \sigma$	-0.2	-0.1	-0.05	-0.02	-0.01
0.51	0.520	0.520	0.540	0.885	0.943
0.52	0.540	0.579	0.863	0.954	0.977
0.55	0.833	0.917	0.961	0.985	0.992
0.60	0.945	0.975	0.989	0.995	0.998
0.70	0.967	0.986	0.994	0.997	0.999

Figure 8: Probabilities $P_{\sigma,q}$ of convergence to the correct belief b_1^*, where $P_{\sigma,q} = b_1^*$.

[6] These values have been determined analytically according to the procedures used in the proofs of lemma 1 and lemma 2.

[7] The average has been calculated over one million simulation runs.

$q \mid \sigma$	-0.2	-0.05	-0.01
0.51	7.0	26.0	3554.5
0.55	39.1	173.2	577.4
0.70	10.5	33.4	147.1

Figure 9: Average time of reaching the correct absorbing state b_1^*, conditional on reaching it (1 million simulation runs each).

5 Conclusion

In this paper, I studied the occurrence of herds and informational cascades in the presence of strategic externalities. I presented two specifications of a basic model, one with continuous signals and one with binary signals.

In the continuous case, I showed that if strategic externalities are negative and sufficiently weak, (anti-)herding and cascade behavior never occur. On the other hand, I showed that an informational cascade always occurs if strategic externalities are either positive or negative and sufficiently strong. I demonstrated that these results can be obtained by focusing on individuals' incentives to engage in either imitative or contrarian behavior. Moreover, I argued that my results underline that the *herding* as well as *non-herding* results of Çelen and Kariv (2004) strongly rely on their assumption of *pure* informational externalities. Finally, I showed that the obtained results are not altered by different assumptions on the observability of predecessors' actions. This contrasts with the case of pure informational externalities.

In the discrete version of the model, herding always occurs provided that strategic externalities are non-negative. In case they are negative, I showed that the learning process eventually reaches one of two possible absorbing states, one in which the corresponding public belief is close to the 'correct' state of the world and one in which it is close to the 'wrong' state of the world. I showed that if either strategic externalities are sufficiently small or signals are sufficiently accurate, the public belief ends up arbitrarily close to the 'correct' state of the world - with arbitrarily high probability.

Appendix A: Continuous State Space, Binary Signals, Complete Learning and the Rule of Succession

In this appendix, I study a simple model which is different from the basic model of section 2. It also incorporates the idea of learning in the presence of strategic externalities (negative externalities in this case) but unlike the previous model, agent's payoffs do not only depend on predecessors' actions, but on the actions off *all* agents.

Let there be N firms, indexed by $n = 1, 2, ..., N$. Each firm has to make a once-in-a-lifetime entry decision a_n between two markets A and B. Total profits in both markets are known to be of size one, but the exact distribution is unknown. Let θ_0 denote the fraction of total profits that are earned in market A. It is common knowledge that θ_0 is drawn from a uniform distribution on $[0, 1]$. Each firm n employs an expert who makes a recommendation $\theta_n \in \{A, B\}$, and each firm only observes their own expert's recommendation. Moreover, recommendations are independent of each other and accurate in the sense that $\Pr(\theta_n = A) = 1 - \Pr(\theta_n = B) = \theta_0$. Profits in each market are divided equally among all companies who operate within this market. Let $\mathcal{N}_A$ denote the total number of firms who choose market A. Firms are risk-neutral with utility function

$$u(a) = \begin{cases} \frac{\theta_0}{\mathcal{N}_A} & \text{if } a = A \\ \frac{1-\theta_0}{N-\mathcal{N}_A} & \text{if } a = B \end{cases}$$

Decisions are taken sequentially, where each firm n observes the actions of a subset of size $\lambda_n \in \{0, 1, ..., n-1\}$ of its predecessors.[8] Moreover, firms update their beliefs about θ_0 via Bayes' rule. The following proposition shows that there always exists an efficient equilibrium in which all firms are purely self-reliant (i.e. they entirely rely on their own information and neglect the actions of others). This result is independent of the information structure. The reason why the information about the behavior of others is worthless is that for each agent's action choice, the strategic substitutability induced by the action exactly outweighs the corresponding positive informational externality.

Proposition 12. *There exists an equilibrium in which all firms reveal their private information, i.e. $a_n = \theta_n$ for all n, regardless of their sample size λ_n.*

[8] It is irrelevant to the solution whether the information structure of the firms are known by the other firms.

Proof. Consider any firm n and assume that all other firms $n' \neq n$ play according to the strategy $a_{n'} = \theta_{n'}$. I show that agent n's best response to this strategy profile is $a_n = \theta_n$.

Let $k \leq \lambda_n$ denote the number of predecessors in firm n's sample who chose actions A and let $\widetilde{\mathcal{N}}_A$ denote the total number of agents other than agent n who choose action A. Hence,

$$\mathrm{E}\left[u(A)|\theta_n\right] = \sum_{i=0}^{N-1-\lambda_n} \Pr\left[\widetilde{\mathcal{N}}_A = k+i|\theta_n\right] \frac{\mathrm{E}\left[\theta_0|\widetilde{\mathcal{N}}_A = k+i, \theta_n\right]}{1+k+i} \tag{7}$$

It follows from the uniform prior and Bayesian updating that, conditional on $\widetilde{\mathcal{N}}_A$ and θ_n, θ_0 follows a Beta distribution with parameters $\alpha = 1 + \widetilde{\mathcal{N}}_A + \mathbf{1}_{\{\theta_n = A\}}$ and $\beta = 1 + N - \widetilde{\mathcal{N}}_A - \mathbf{1}_{\{\theta_n = A\}}$ (cf. the *rule of succession*). Since the expected value of a Beta distributed random variable with parameters α and β is $\alpha/(\alpha+\beta)$, it follows that:

$$\mathrm{E}\left[u(A)|\theta_n\right] = \frac{1}{N+2} \sum_{i=0}^{N-1-\lambda_n} \Pr\left[\widetilde{\mathcal{N}}_A = k+i|\theta_n\right] \frac{1+k+i+\mathbf{1}_{\{\theta_n=A\}}}{1+k+i}.$$

Similarly,

$$\mathrm{E}\left[u(B)|\theta_n\right] = \frac{1}{N+2} \sum_{i=0}^{N-1-\lambda_n} \Pr\left[\widetilde{\mathcal{N}}_A = k+i|\theta_n\right] \frac{N-k-i-\mathbf{1}_{\{\theta_n=A\}}}{N-k-i}.$$

The claim follows from pairwise evaluation of the summands of these two expressions. If $\theta_n = A$, one gets

$$\frac{2+k+i}{1+k+i} > \frac{N-k-i}{N-k-i} = 1,$$

and if $\theta_n = B$, one gets

$$\frac{1+k+i}{1+k+i} = 1 < \frac{N-k-i+1}{N-k-i}.$$

Hence, $\mathrm{E}\left[u(A)|A\right] > \mathrm{E}\left[u(B)|A\right]$ and $\mathrm{E}\left[u(A)|B\right] < \mathrm{E}\left[u(B)|B\right]$ are satisfied regardless of the actions in agent n's sample. This implies that firm n's best-response to the strategy profile is given by $a_n = \theta_n$. □

Appendix B

Proof of proposition 3

Assume that $\overline{\theta}'_n$ is always within the interval $[-1, 1]$, i.e. $\overline{\theta}_n = \overline{\theta}'_n$. The case in which $\overline{\theta}'_n$ 'jumps' out of the interval is mathematically straightforward and challenging only in terms of the notation.

First, let me show by induction that $\Pr(a_n = 1) = \frac{1}{2}$ and $\overline{\theta}_n + \underline{\theta}_n = 0$ for all n. Assume that $\Pr(a_m = 1) = \Pr(a_m = -1) = \frac{1}{2}$ and $\overline{\theta}_{m+1} + \underline{\theta}_{m+1} = 0$ is satisfied for some m and let me show that this implies that it must be satisfied for $m+1$ as well. Symmetry implies directly that $\Pr(a_{m+1} = 1) = \Pr(a_{m+1} = -1) = \frac{1}{2}$. Moreover,

$$\begin{aligned}
& \underline{\theta}_{m+2} + \overline{\theta}_{m+2} \\
= \; & -\mathrm{E}(\theta_{m+1} \,|\, a_{m+1} = 1) - \mathrm{E}(\theta_{m+1} \,|\, a_{m+1} = -1) \\
& - \sum_{j \in \{-1,1\}} \sum_{k \in \{-1,1\}} \Pr(a_m = j | a_{m+1} = k) \, \mathrm{E}\left[\sum_{i=1}^{m-2} (\theta_i + \sigma a_i) \,|\, a_m = j, a_{m+1} = k \right] \\
= \; & \overline{\theta}_{m-1} (\Pr(a_m = 1 | a_{m+1} = 1) - \Pr(a_m = -1 | a_{m+1} = 1) \\
& + \Pr(a_m = 1 | a_{m+1} = -1) - \Pr(a_m = -1 | a_{m+1} = -1)) = 0.
\end{aligned}$$

The last equality is implied by the fact that the second and the third (as well as the first and the fourth) summand within the brackets add up to zero (this is easy to proof). It remains to show that the statement holds for $m = 1$. Clearly, $\Pr(a_1 = 1) = \Pr(\theta_1 \geq 0) = 0.5$ and straightforward evaluation of equation (2) in section 3 yields $\overline{\theta}_2 = -(\sigma + 1/2)$ and $\underline{\theta}_2 = \sigma + 1/2$.

Moreover, equation (2) in section 3 implies that

$$\overline{\theta}_n \;=\; -\mathrm{E}\left[\sum_{i=1}^{n-2} (\theta_i + \sigma a_i) \,|\, a_{n-1} = 1 \right] - \mathrm{E}\left[\theta_{n-1} + \sigma a_{n-1} \,|\, a_{n-1} = 1 \right]$$

Let me calculate the two parts of the right-hand expression separately.

$$\begin{aligned}
& \mathrm{E}\left[\sum_{i=1}^{n-2}(\theta_i + \sigma a_i) \,|\, a_{n-1} = 1\right] \\
= & \sum_{k \in \{-1,1\}} \Pr(a_{n-2} = k | a_{n-1} = 1)\,\mathrm{E}\left[\sum_{i=1}^{n-2}(\theta_i + \sigma a_i) \,|\, a_{n-1} = 1, a_{n-2} = k\right] \\
= & \frac{\Pr(a_{n-2} = 1)\ \Pr(a_{n-1} = 1 \,|\, a_{n-2} = 1)}{\Pr(a_{n-1} = 1)}(-\overline{\theta}_{n-1}) \\
& + \frac{\Pr(a_{n-2} = -1)\ \Pr(a_{n-1} = 1 \,|\, a_{n-2} = -1)}{\Pr(a_{n-1} = 1)}(-\underline{\theta}_{n-1}) \\
= & \frac{\frac{1}{4}(1 - \overline{\theta}_{n-1})}{\frac{1}{4}(1 - \overline{\theta}_{n-1}) + \frac{1}{4}(1 + \overline{\theta}_{n-1})}(-\overline{\theta}_{n-1}) + \frac{\frac{1}{4}(1 + \overline{\theta}_{n-1})}{\frac{1}{4}(1 - \overline{\theta}_{n-1}) + \frac{1}{4}(1 + \overline{\theta}_{n-1})}(+\overline{\theta}_{n-1}) \\
= & \frac{1 - \overline{\theta}_{n-1}}{2}(-\overline{\theta}_{n-1}) + \frac{1 + \overline{\theta}_{n-1}}{2}(+\overline{\theta}_{n-1}) = \overline{\theta}^2_{n-1}
\end{aligned}$$

Furthermore,

$$\begin{aligned}
\mathrm{E}\left[\theta_{n-1} \,|\, a_{n-1} = 1\right] &= \sum_{k \in \{-1,1\}} \Pr(a_{n-2} = k | a_{n-1} = 1)\,\mathrm{E}\left[\theta_{n-1} \,|\, a_{n-2} = k\right] \\
&= (1 - \overline{\theta}_{n-1})\frac{1 + \overline{\theta}_{n-1}}{2} + (1 + \overline{\theta}_{n-1})\frac{1 - \overline{\theta}_{n-1}}{2} = \frac{1}{2}\left(1 - \overline{\theta}^2_{n-1}\right)
\end{aligned}$$

Hence,

$$\overline{\theta}_n = -\overline{\theta}^2_{n-1} - \frac{1}{2}\left(1 - \overline{\theta}^2_{n-1}\right) - \sigma = -\sigma - \frac{1}{2}\left(1 + \overline{\theta}^2_{n-1}\right)$$

and $\underline{\theta}_n = -\overline{\theta}_n = \sigma + \frac{1}{2}\left(1 + \overline{\theta}^2_{n-1}\right)$.
Q.E.D.

References

[1] Banerjee, A., 1992. A simple model of herd behavior. *Quarterly Journal of Economics* 107, 797-817.

[2] Bikhchandani, S., Hirshleifer, D. and Welch, I., 1992. A theory of fads, fashion, custom, and cultural change as informational cascade. *Journal of Political Economy* 100, 992-1026.

[3] Choi, J., 1997. Herd Behavior, the "Penguin Effect," and the Suppression of Informational Diffusion: an Analysis of Informational Externalities and Payoff Interdependency. *Rand Journal of Economics* 28 (3), 407-425.

[4] Çelen, B. and Kariv, S., 2001. Observational learning under imperfect observability. CESS working paper 02-03. New York University.

[5] Çelen, B. and Kariv, S., 2004. Observational learning under imperfect observability. *Games and Economic Behavior* 47, 72-86.

[6] Debo, L.G., Parlour, C.A. and Rajan, U., 2005. Herding and Congestion. Mimeo. http://ozymandias.tepper.cmu.edu

[7] Debo, L.G. and Veeraraghavan, S.K., 2008. Models of Herding Behavior in Operations Management. Mimeo. University of Pennsylvania.

[8] Dasgupta, A., 2000. Social Learning with Payoff Complementarities. Mimeo. Yale University.

[9] Drehmann, M., Oechssler, J. and Roider, A., 2007. Herding with and without payoff externalities an internet experiment. *International Journal of Industrial Organization* 25, 391-415.

[10] Frisell, L., 2003. On the Interplay of Informational Spillovers and Payoff Externalities. *Rand Journal of Economics* 34 (3), 582-592.

[11] Kahan, M., 1997. Social Influence, Social Meaning, and Deterrence. *Virginia Law Review* 83, 276-304.

[12] Smith, L. and Sørensen, P., 1996. Rational social learning by random sampling. Mimeo. MIT.

[13] Smith, L. and Sørensen, P., 2000. Pathological outcomes of observational learning. *Econometrica* 68, 371-398.

[14] Welch, I., 1992. Sequential sales, learning, and cascades. *Journal of Finance* 47 (2), 695-732.