A game-theoretic Analysis of a Market for long-term Relationships

Inaugural-Dissertation
zur Erlangung des akademsichen Grades eines
Doktors der Wirtschaftswissenschaften
der Universität Mannheim

vorgelegt von

Heiner Schumacher

Mai 2008

Abteilungssprecher: Prof. Dr. Hans Peter Grüner
Referent: Prof. Dr. Ernst-Ludwig von Thadden
Korreferent: Prof. Jörg Oechssler, Ph.D.
Tag der Verteidigung: Montag, 21.07.2008

Acknowledgements

First and foremost I would like to thank my supervisor Ernst-Ludwig von Thadden for his excellent guidance, continuing encouragement and the great research environment at his chair. Jörg Oechssler and Francesco Squintani complete my thesis committee. I am much indebted for their support and helpful comments. Francesco supervised the first part of this thesis and contributed a lot to the final version. Jörg Oechssler offered a great introduction into the theory of learning in games and helped a lot with writing the second and third part. The joint work with Alexander Gaigl, co-author of chapter four, was and continues to be a great pleasure and source of motivation. Financial support from the Deutsche Forschungsgemeinschaft is gratefully acknowledged.

I am very grateful to Mike Felgenhauer, Tri Vi Dang, Michal Kowalik, Sabine Münch for the excellent atmosphere at the Chair of Economic Theory and to my fellows from the Center for Doctoral Studies in Economics (CDSE) for creating a stimulating "environment", especially Gonzague Vanoorenberghe, Malte Hübner, Daniel Harenberg, Christoph Nagel and Björn Sass. I also want to thank the faculty of CDSE for the excellent courses in the first year of the program, namely Andreas Irmen, Felix Kübler, Ulrich Schlieper, Markus Nöth, Hans-Peter Grüner, Enno Mammen, Stephan Hoderlein, Michael Gerfin, Günter Franke and Konrad Stahl.

Moreover, I thank the University College London for their hospitality and V. Bhaskar, Daniele Condorelli and Andreas Uthemann for helpful discussions. This work profited from comments and suggestions by several people. I am especially grateful for discussions with and support by Eckhard Janeba, Martin Peitz, Jürgen Eichberger, Jeanine Miklós-Thal, Ludwig Reßner, Melanie Schienle, Markus Kinateder, Kerstin Gerling, Thomas Gall, Iris Kesternich, Andrey Ivanov, Nicola Persico, Angelino Viceisza, the iCERN-Group, Frank Rosar and Sonya Mihaleva.

Last but certainly not least, I thank my parents for their permanent backing and encouragement.

Bei Mannheim weint man zwei mal:
Wenn man kommt und wenn man geht.

Contents

Chapter 1

Introduction

One of the most common institutions in a society is the long-term relationship. Activities which can be done either with varying or with permanent partners include social or professional relations, participation in the labor market or coauthoring. We observe that in most of these activities, we are linked with a fixed set of agents: The largest share of current employment contracts is of long-term nature. The marriage usually is celebrated with the intention to maintain the relationship until one partner dies.

Strangely enough, the option to maintain or to quit relationships is largely ignored in the theory of repeated games and its experimental counterparts: mostly, an agent's opponents are given exogenously, i.e. they are fixed or there is random matching in every period. With this option, the strategic opportunities in a repeated games setting may change dramatically. Consider, for example, the infinitely repeated prisoner's dilemma. In a setting with fixed opponents, it is simple to punish deviant behavior of opponents by playing a strategy like tit-for-tat. However, if a player can quit the relationship and switch to other opponents immediately, such a punishment will no longer be a threat. We have to find other mechanisms which implement cooperation.

In this thesis, we therefore study social interaction of the following form: each agent of a large population plays a game of conflicting interests—as, for example, the prisoner's dilemma—with some opponent in each period. After observing the partner's action choice, each player has the option to maintain or to quit the current relationship. If the first action is chosen by both agents, they will play the game together in the next period, otherwise they return to a "market" for long-term relationships and will be matched randomly with another opponent. The matching process in the market is global and non-assortative: everybody can be matched together with anybody and own behavior does not affect the probability of being paired up with an agent who plays a certain strategy. Furthermore, there are no information flows between pairs. With (small) positive probability, a relationship is broken up, regardless of the agent's action choice. Thus, it is unlikely that there is no agent in the market whenever the population is large.

Note that in each period of the game, the population is divided into two parts: Fixed relationships in which agents choose to stay with their current opponent and the market for long term relationships, in which agents are matched randomly.

In the first part of the thesis, we analyze the outcome of this game by using the standard game-theoretic approach. We extend the Folk Theorems of Friedman (1971) and Fudenberg and Maskin (1986) to games of this form, and establish a structural difference between models with finitely and infinitely many agents: In an infinite population, the probability of meeting the same opponent another time after the relationship with her has been broken up, is zero. There cannot be punishment by the same opponent after a deviation. Whenever there are only finitely many individuals in the population, one may well meet the same opponent again though the relationship was broken up voluntarily in the last period. We therefore can show the following: while with finitely many players any individually rational average payoff can be reached, this is not possible with infinitely many agents in games of conflicting interests. However, it is possible to establish a Folk Theorem for the latter case, using a strategy which prescribes to "start small": at the beginning of a new relationship, both agents play a profile with low payoffs—the Nash outcome of the stage game, for example—and start to cooperate in later periods. Whenever a player deviates from this path of play, the opponent chooses to quit the relationship. Thus, any gain from deviation is wiped out by the subsequent play of a profile with low payoffs in the new relationship.

However, this solution has serious shortcomings:

- Without further assumptions, a strategy which prescribes to "start small", is not robust against communication: given that all other players in the population stick to the described pattern, it is optimal for two agents who meet in the market for the first time, to start the relationship with the cooperative profile immediately. This would not violate any incentive constraint. However, if all pairs act in this way, we are no longer in equilibrium. Gosh and Ray (1996) solve this problem by introducing heterogeneous time preferences: a fixed share of players is myopic while the rest of the population is patient to some extent. Thus, a period of less cooperation at the beginning of a new relationship serves to "test" the opponent's patience. However, as long as we do not make this assumption, we do not have a convincing solution for the considered game.

- The second disadvantage of the existing solution is more fundamental to game theory: the equilibrium concept requires knowledge of the opponents' strategies. Especially for large populations, this assumption is unrealistic. Conventional solutions leave open the question how agents manage to reach an equilibrium in the absence of abundant sophistication. Also note that in the considered game, there are many equilibria, including a situation in which agents never cooperate.

We therefore pursue a different approach in the remaining two parts of this thesis. Agents are no longer assumed to have knowledge about the aggregate play in the market. Our main goal is to show that in the absence of knowledge of aggregate play, a large share of the population will play a very simple and intuitive strategy in a stable, equilibrium-like situation: this strategy is "cooperate in every period and maintain the relationship if and only if your opponent does so as well". In particular, we are concerned with the dynamics in the population, i.e. how aggregate behavior evolves over time and under what circumstances a cooperative outcome can be reached.

In the second part, agents are boundedly-rational: they are myopic and do not have the ability to follow history-dependent strategies. Strategies are very simple heuristics, in particular "never cooperate and quit every relationship" and "cooperate and maintain the relationship if and only if your opponent cooperates as well". There is nothing like a punishment phase at the beginning of a relationship. Agents have sometimes access to information about what behavior is on average the most successful one in the population. They imitate this strategy whenever the period payoff in the current relationship is smaller than the average payoff of the most successful strategy. With small probability, they switch to another strategy randomly and quit the current relationship, i.e. they choose to "experiment".

We can show that in this framework, cooperative behavior emerges in the population if both the imitation and experimentation rate are sufficiently small. The share of cooperative long-term relationships only increases if the share of cooperators in the market is not too small over a larger number of periods, i.e. if cooperators in the market do not switch to the non-cooperative strategy too fast. Additionally, if too many agents imitate in each period, the share of cooperative players in the market—and therefore the average utility of non-cooperative players—is sufficiently high, such that cooperative long-term relationships are broken up, as many individuals switch to the non-cooperative strategy.

In order to relax the assumption on the imitation rate, we then introduce heterogeneity into the population: an agent does not want to interact with any agent of the population, but prefers to meet individuals with certain characteristics or manners. Long-term relationships only can exist between agents who “get along with each other”. The period payoff a player receives increases by a fixed amount if she gets along with her current opponent, regardless of the strategies played by agents. Thus, the strategy “never cooperate and maintain the relationship if and only if you get along with your opponent”, becomes attractive if the probability of meeting such an opponent in the market is sufficiently small. Under the assumption of heterogeneity of players, we maintain the same result as in the homogeneous setting, but without a boundary on the imitation rate: even though almost all agents switch to the most successful strategy in each period, learning agents adapt to the cooperative strategy in almost all periods if the experimentation rate is sufficiently small.

In this chapter, we emphasize the importance of imitation as a form of naive learning which leads to a cooperative outcome. We will see that other forms of naive learning—as for example fictitious play or regret matching—do not have this attribute in the considered game.

We do not have to assume that players are boundedly-rational and myopic in order to maintain cooperative outcomes in simple strategies. In the last part of this thesis, agents once again understand the trade-off in the game above, but still have no knowledge about the aggregate play of individuals in the market. They know that one can play according to a non-cooperative strategy, which prescribes to defect in each period of a relationship, or according to a cooperative strategy, which leads to a cooperative long-term relationship if both agents follow this strategy. We allow for “starting small” strategies, but in particular we are interested in the outcome if the cooperative strategy prescribes to start with cooperation immediately.

With a restricted strategy space, agents learn their current opponent’s strategy in finitely many periods. Their subjective believe about the aggregate play in the market is based on past experiences. If almost none or almost everybody of her previous opponents played according to the non-cooperative strategy, it is optimal for the agent to play according to this strategy as well. If she made mixed experiences—opponents sometimes played cooperatively, sometimes not—then it is optimal for her, to play according to the cooperative strategy.

As updating rule we take various forms of "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 his opponent's play.

In a first step, we assume that there are infinitely many agents in the population. Analytically, and by simulating the model, we derive conditions under which a significant degree of cooperation can be expected. For a large measure of initial distributions of beliefs, aggregate play converges to a cooperative outcome. This remains true if strategies are very simple and punishment within a relationship is not possible. The result is a population in which different agents make different experiences in the market and therefore act differently even if aggregate play remains constant.

In a second stage, we extend the model to finite populations and show that the dynamics are similar to the infinite case if there are sufficiently many agents.

The thesis is organized in such a way that the chapters can be read independently of each other. As the chapters are closely connected, this involves a certain degree of repetition for readers who read through the entire manuscript at the benefit of readers who read the chapters selectively. All references are collected in the bibliography.

Chapter 2

Cooperative Equilibria in Repeated Games with Endogenous Matching Decision

2.1 Introduction

The literature on Folk Theorems in infinitely repeated games in most cases considers one of two common settings: Either an agent is forced to play some stage game against a fixed set of opponents all the time—see Friedman (1971), Maskin and Fudenberg (1986)—or she plays against different opponents in each period—see Ellison (1994) and Kandori (1992). In neither case agents have the opportunity to decide on the maintenance of the current relationship.

In this paper we establish versions of the Folk Theorem for games in which this option exists: A n-player stage game is played repeatedly and simultaneously by many groups. These groups are matched randomly, but the members of each group can decide whether they want to maintain or quit the current relationship. Any player can leave her group in order to be matched to other opponents. An agent only observes the history of her own matches, i.e. the identity of current and former opponents and the actions they have chosen in the periods in which she played the stage game with them. There are no information flows between the groups. We assume that for every player there is a positive (but small) chance of being deleted from the set of players and replaced by a new agent. Therefore, the set of agents to be matched randomly in the next period is always non-empty with positive probability.

Compared to the standard repeated games setting we only introduce the multiplicity of n-player groups and the exogenous rate of detachment from the set of players.

Our main interest is the equilibrium set of action profiles and expected average payoffs which can be supported by some strategy. We show that for finite population size the Folk Theorems of Friedman (1971) and of Maskin and Fudenberg (1986) hold, i.e. every payoff vector which strictly dominates a Nash equilibrium of the stage game (or every individually rational payoff vector respectively) can be supported as an expected average payoff whenever the set of players is stable enough and the discount factor is sufficiently high. This result relies on the fact that, even if there is the opportunity to quit a relationship, a player may not find new opponents immediately when all other players choose to stay in their groups.

When the set of players is of infinite size, every agent is able to switch to other opponents immediately. However, it is possible to support the repeated play of profiles which are not Nash equilibria of the stage game: All agents start a long-term relationship by choosing a low payoff profile in the first periods of a new group. After this time they play the desired profile as long as the match exists. Whenever some player deviates, each agent of the group quits the relationship. If the discount factor is sufficiently high and agents remain in the set of players for sufficiently long time, the gains from any deviation of any player will be wiped out by the subsequent period of low payoffs in the next match. With this strategy it is possible to approximate many efficient expected average payoff profiles arbitrarily closely.

Furthermore, we investigate how to construct efficient strategies for fixed values of the discount factor and the probability of detachment in both the infinite and the finite setting. For certain stage games of interest—like the prisoner's dilemma—it is possible to show that no other strategy combination dominates the described patterns of behavior. In these cases we get a clear justification for the existence of long-term relationships in repeated games in which players have the opportunity to maintain or to quit relationships.

The rest of the paper is organized as follows: The next section introduces the basic model. The main results for the case of infinitely many agents are derived in section 3.1, in section 3.2 we show that for some stage games it is possible to show the optimality of the supporting strategies. In section 3.3, we discuss the concept of "bilateral rationality" which is related to the strategies of this section. Chapter 4.1 presents the main result for the model with finitely many agents. In chapter 4.2, we apply it to a stylized version of a labor market model. Section 4.3 shows how to combine the strategies of the third and fourth chapter in order to establish cooperation. In chapter 5, we provide an overview of the related literature. The last chapter concludes.

2.2 Outline of the Model

Time is discrete in our model and periods are marked by $t = (0, 1, 2, ...)$. Denote by $C_i(t)$, $i \in (1, ..., n)$, sets of countably many agents where

$$| C_i(t) |=| C_j(t) |= m \tag{2.1}$$

for all $i, j \in (1, ..., n)$, all periods t, and $m \in \mathbb{N} \cup (\infty)$ is a constant. Hence, $C(t) = \bigcup_{i=1}^{n} C_i(t)$ is the set of all players in a given period t and m is the number of matches who play the stage game simultaneously in each period. An agent from the set $C_i(t)$ will be called an agent from class i. As we will focus on equilibria in which all agents of class i follow the same strategy, we do not introduce notation for individual players.

Let $\sigma \in (0, 1]$ be the probability that nature replaces a single agent who was a member of $C_i(t)$ by a new agent in the set $C_i(t + 1)$ at the beginning of period $t+1$. Let a new agent in period t be a player who is not in the sets $C_i(\tau)$, $\tau < t$, but is in the set $C_i(t)$. We say that an agent of class i is deleted from the set of players in period t if she was in the set $C_i(t)$ but not in the sets $C_i(\tau)$, $\tau \in (t+1, t+2, ...)$. The probability $\sigma > 0$ is equal for all classes i.

Call a group of n agents of different classes a match r. Every match plays in each period a finite stage game g. The stage game g can be any normal or extensive form game of complete information. Agents from the sets C_i always take on position i in g and have the pure finite action set A_i. An element of this set played by an agent of class i in period t is denoted by a_i^t. The set of all available pure and mixed actions of an agent of class i is Σ_i. The space of all pure action profiles then is $A = \times A_i$ while the space of all action profiles is $\Sigma = \times \Sigma_i$.

After the game has been played and payoffs are realized, each player has the opportunity to quit or to maintain the relationship with her current opponents. The extended action set for an agent of class i therefore is $\Sigma_i \times \{Q, M\}$ where the decision to play Q (quit) or M (maintain) may be conditional on the outcomes of g in the recent and preceding periods. If all players of a group r choose option M in period t, then they play the stage game against each other in period $t+1$ unless one or more players of r are deleted from the set of agents.

In a given group of agents r, the position of player i is said to be "vacant" at the beginning of period $t+1$ if either the player of class i was deleted from the set of players at the end of period t or the respective player chose to quit the relationship in t. By the assumption in (2.1), the number of vacant positions for players of class i coincides in each period with the number of those agents of the same class who either are new in the set of players or chose Q in the previous period. We further assume that all possible ways of pairing up those vacant positions and agents occur with the same probability. Hence, if m is finite, there is a positive probability for each agent to be matched together with the same opponents even when she chose Q in the previous period.

Example 1. To illustrate the matching process we label the individual players in this example. Assume that $m = 2$ and there are 4 classes of players, C_1, C_2, C_3 and C_4 where each class i contains the agents a_i and b_i in period t. The groups in period t are given by

a_1	a_2	a_3	a_4
b_1	b_2	b_3	b_4

If all players except b_3 choose M in period t then the second match (the second line) will be the same in period $t+1$ if and only if

- no agent from the set $\{b_1, b_2, b_3, b_4\}$ is replaced by a new player and
- a_3 is not replaced by a new player *or* a_3 is replaced by a new player but this player is matched to the group $\{a_1, a_2, a_4\}$ and b_3 again is grouped together with his previous opponents (where the probability of this event would be $\frac{1}{2}$).

If all agents except a_3 and b_3 choose M in period t then the groups will be the same in period $t+1$ if and only if

- no agent is replaced by a new player and
- a_3 is matched to the group $\{a_1, a_2, a_4\}$ and b_3 to the group $\{b_1, b_2, b_4\}$ (where again the probability of this event would be $\frac{1}{2}$).

Finally, if all agents from the set $\{a_1, a_2, a_3, a_4\}$ choose M in period t while all other agents choose Q, then the groups will be the same in period $t+1$ if and only if no agent is replaced by a new player. [**End of example**]

For convenience, we introduce the following two definitions: A match r is "formed" in period t if its members were not in one group in period $t-1$. A match r is "dissolved" in period t if its members are not matched together in period $t+1$.

In order to convexify the equilibrium set, we assume throughout the paper that a public signal θ is realized in every period. It takes on all real values in the interval $[0,1]$ with equal probability.

The sequence of events in every period is as follows:

(i) New agents enter the set of players. Together with those agents who chose Q in the previous period they are paired up randomly to matches with the respective vacant positions.

(ii) A public signal is realized.

(iii) All agents choose their actions, payoffs are realized.

(iv) After observing the opponents' action choice, each agent chooses between the options M and Q.

(v) Nature decides which agents leave the set of players.

Each player knows only the history of her own matches, i.e. she has perfect recall with respect to her former opponents' identities and what actions they chose in the periods she was matched with them. Denote the respective history of an agent in period t by h_t. Thus, each entry of h_t consists of an action profile $a \in A$ and the opponents' identities. The space of all histories at t is given by H_t. Let

$$\bar{H} = \bigcup_{t \in \mathbb{N}} H_t, \tag{2.2}$$

be the set of all finite histories. Agents are not informed about the action choice simultaneously made in other matches. Beliefs are given by the equilibrium strategies. Our solution concept will be subgame perfect equilibrium.

A strategy f^i of class i of players has two parts: the action-choice, f_1^i, and the matching-decision, f_2^i. The former is a function

$$f_1^i : \bar{H} \to \Sigma_i, \tag{2.3}$$

while the latter one is given by a function

$$f_2^i : \bar{H} \times \Sigma \to \{Q, M\}. \tag{2.4}$$

Denote by $\tilde{f}^i$ the belief of an agent of class i about all other agents' strategies. The utility of a player of class i is given by a continuous function $\Pi_i : \Sigma \to \mathbb{R}$. An action profile $x \in \Sigma$ yields the utility $\pi_i(x)$ for class i. Agents discount utility with $\delta \in [0, 1]$ and maximize over the sum of discounted expected utility in the current period t

$$E_i(f^i, \tilde{f}^i, t) = \sum_{\tau=t}^{\infty} \delta^{\tau-t} \mathbf{E}_{\tilde{f}^i,m,\sigma}[\pi_i(a^\tau) \mid f^i, t], \tag{2.5}$$

where $\mathbf{E}$ is the expectations operator induced by $\tilde{f}^i$, m and σ.

The assumption of a constant σ is crucial: It implies that agents cannot gain by conditioning behavior upon the number of periods in which their current group was matched together or the duration of their affiliation to the set of players. Hence, the repeated game maintains a recursive structure. We denote the described game with $\Gamma(g, \delta, \sigma, m)$.

Some more notation is needed: Let Σ^{NE} be the set of all Nash equilibria of g (which of course is non-empty). Assume that $y^k \in \Sigma$ (where k is some index). A best response of players of class i to the profile y^k_{-i} is denoted by $y_i^{b,k}$. The minimax profile for class i is given by $\underline{x}^i$ and the respective payoff is

$$\pi_i(\underline{x}^i) = \min_{x_{-i} \in \Sigma_{-i}} \max_{x_i \in \Sigma_i} \pi_i(x_i, x_{-i}). \tag{2.6}$$

All payoff profiles π with $\pi_i > \pi_i(\underline{x}^i)$ for all i are called individually rational. Let $V = co(\pi(x) \mid x \in \Sigma)$ be the set of feasible payoffs. Then the set of individually rational payoffs is

$$V^* = \{\pi \in V \mid \pi_i > \pi_i(\underline{x}^i), i \in (1, ..., n)\}. \tag{2.7}$$

Of importance will be the set of payoff profiles which dominate some convex combination of Nash outcomes. Define $U^{NE} = co(\pi(x) \mid x \in \Sigma^{NE})$, i.e. the convex hull of all Nash payoff profiles. Then the required set is given by

$$V^{NE} = \{\pi \in V \mid \exists \bar{\pi} \in U^{NE} : \pi_i > \bar{\pi}_i, i \in (1, ..., n)\}. \tag{2.8}$$

We will speak of a mixed profile x when a collection of pure profiles $X = \{x^1, ..., x^k\} \subseteq A$, i.e. the support of x, is played according to the realization of θ. Thus, the probability that a certain profile x^l is played is given by a function $p_\theta : A \to [0, 1]$. The expected payoff from the mixed profile x for a player of class i is then given by

$$\pi_i(x) = \sum_{l=1}^{k} p_\theta(x^l) \pi_i(x^l). \tag{2.9}$$

Sometimes we will refer to x as a combination of Nash profiles with support $X \subseteq \Sigma^{NE}$ and to y as a combination of some action profiles with support $Y \subseteq \Sigma$. See Fudenberg and Maskin (1991) and Sorin (1986) for discussion and possibilities to avoid the assumption of public randomization under certain conditions.

2.3 Cooperative equilibria in games Γ with infinitely many players

2.3.1 On the support of cooperative outcomes

In this section we analyze the support of profiles which are not elements of Σ^{NE} in games Γ with $m = \infty$. With infinitely many groups, the probability that two players meet again when their previous match has been dissolved, is zero. The most obvious difference to the traditional results for infinitely repeated games with observable outcomes is captured in our first observation:

Observation *No action profile $y \in \Sigma$ which is not an element of Σ^{NE} can be supported to be played by all groups in all periods in a Nash equilibrium of the game $\Gamma(g, \delta, \sigma, \infty)$.*

Proof: Assume that all matches play y in every period. As $y \notin \Sigma^{NE}$ for at least one class i of players there is a best response y_i^b with $\pi_i(y_i^b, y_{-i}) > \pi_i(y)$. As $\sigma > 0$, the probability that she is matched to new opponents is 1 whenever she chooses to quit the current relationship. Then a player of class i can deviate profitably by playing y_i^b and Q in every period.
Q.E.D.

The statement implies that in games with conflicting interests like the prisoner's dilemma (PD), V^* or V^{NE} cannot be the equilibrium sets of expected average payoffs even when δ is very close to 1. As long as there are infinitely many agents to be matched at the beginning of each period, every player has the outside option to switch to new opponents immediately. If the sum of discounted expected utility in the next match is not lower than in the current one, there is no reason to avoid the dissolution of r or a change in the current opponents' behavior.

However, there is at least one possibility to implement the repeated play of profiles which are not Nash equilibria of the stage game: At the beginning of a new relationship r a Nash profile (or a combination of Nash equilibria)

x is played by r for T periods. After this time, the players of r switch to a profile y which yields higher payoffs for every member of the group. Whenever an agent fails to choose the prescribed action, all of his opponents quit the relationship. T is chosen such that any gain that could be attained by a deviation of some class of players is wiped out by the disadvantage a player has to incur when matched to a new group which again starts to play the Nash profile.

Although we assumed time to be discrete, we can treat T like a continuous parameter: If the current match was formed in period t, then in period $t + [T] + 1$ (where the brackets here denote the highest integer smaller than T) the players randomize via the public signal whether they play y or x in the current period. The probability for the latter event is then given by the number $T - [T]$.

In the following we will call this behavior a "*two-stage strategy*", which is associated with the first stage profile x, the second stage profile $y \in \Sigma$ and the number T of periods in which the inferior profile x is played after a new match is formed. Let $E_i((x, y), T)$ be the sum of discounted expected payoffs of a new agent of class i if all groups in the population play the corresponding *two-stage strategy*. The same term is also the continuation value for a player of class i, whose group dissolved in the last period. Note, that the sum of expected discounted average payoffs for each player of class i is equal or higher than $E_i((x, y), T)$.

Our first result establishes that *two-stage strategies* constitute a subgame perfect equilibrium whenever T is chosen large enough, the discount factor is sufficiently high and agents remain in the set of players for a sufficiently long time. We will call a profile y to be "implementable" if it is played repeatedly in a match after finitely many periods subsequent to the formation of this match:

Lemma 1a *For any action profile $y \in \Sigma$ with $\pi(y) \in V^{NE}$, there are values $\bar{\delta} < 1$ and $\bar{\sigma} > 0$, such that y is implementable in a subgame perfect equilibrium of $\Gamma(g, \delta, \sigma, \infty)$ whenever $\delta \geq \bar{\delta}$ and $\sigma \leq \bar{\sigma}$.*

Proof: see Appendix.

The first stage profile x must not necessarily be a Nash equilibrium of g and hence, in some stage games the second stage profile y does not have to dominate a combination of Nash outcomes. We can modify lemma 1a such that any y that dominates a profile x, can be implemented whenever the payoff structure of g exhibits certain properties:

Lemma 1b *For any action profile $y \in \Sigma$ for which there exists a pure action profile x such that*

$$\pi_i(y) > \pi_i(x_i^b, x_{-i}) \geq \pi_i(x) \tag{2.10}$$

for all i, there are values $\bar{\delta} < 1$ and $\bar{\sigma} > 0$ such that y is implementable in a subgame perfect equilibrium of $\Gamma(g, \delta, \sigma, \infty)$ whenever $\delta \geq \bar{\delta}$ and $\sigma \leq \bar{\sigma}$. For mixed profiles x with support $X \subseteq \Sigma$, the same is true if

$$\max_{x^l \in X}\{\pi_i(x_i^{b,l}, x_{-i}^l)\} < \pi_i(y) \tag{2.11}$$

and

$$\pi_i(y) - \pi_i(x) > \max_{x^l \in X}\{\pi_i(x_i^{b,l}, x_{-i}^l) - \pi_i(x^l)\} \tag{2.12}$$

hold for all i.

Proof: see Appendix.

Not every profile with individually rational payoffs is implementable by a *two-stage strategy*. To see this, consider the following stage game where player 1 chooses rows, player 2 chooses columns and player 3 chooses between the matrices X and Y:

X	L	M	R
T	10,10,10	10,1,1	0,0,0
M	1,1,1	0,10,0	0,10,0
B	0,0,0	0,0,0	0,0,0

Y	L	M	R
T	10,10,1	10,1,1	10,10,10
M	1,1,10	1,10,0	10,10,10
B	1,1,10	1,1,10	1,1,10

The minimax payoff for all players is 0 and thus, the profile $\{M, L, X\}$ is individually rational. Every player has the opportunity to deviate profitably from this profile so we need to find a (possibly mixed) first stage profile with smaller payoffs on all coordinates. Whenever the first [the second] player chooses B [R], player 3 can get a payoff of 10 by choosing Y, so the probability that a profile $\{B, ., X\}$ or $\{., R, X\}$ is played in a period of the first phase cannot exceed 0.1. Then only $\{M, M, X\}$ and $\{M, M, Y\}$ remain to be an option (they get a payoff of 0 for at least one player). As these profiles yield a payoff of 10 for player 2 they again must be chosen with a probability less than 0.1. Hence, $\{M, L, X\}$ cannot be implemented by a *two-stage strategy.*

However, if we consider only two-player games, then under certain restrictions every individually rational payoff profile is implementable:

Lemma 1c *If $n = 2$ and $\underline{x}^i$, $i \in \{1, 2\}$, are pure action profiles for both players, then for any action profile $y \in \Sigma$ with $\pi(y) \in V^*$, there are values $\bar{\delta} < 1$ and $\bar{\sigma} > 0$, such that y is implementable in a subgame perfect equilibrium of $\Gamma(g, \delta, \sigma, \infty)$ whenever $\delta \geq \bar{\delta}$ and $\sigma \leq \bar{\sigma}$.*

Proof: By definition, $\pi_i(\underline{x}^i) \geq \pi_i(a_i, \underline{x}^i_{-i})$ for all $a_i \in A_i$ and $i \in \{1, 2\}$. As y is individually rational it holds that

$$\pi_i(y) > \pi_i(\underline{x}^i) \geq \pi_i(\underline{x}_i^{-i}, \underline{x}^i_{-i}) \tag{2.13}$$

for both players. Use $x = (\underline{x}_i^{-i}, \underline{x}^i_{-i})$ as the first stage profile. Then condition (2.10) is satisfied and the result follows from lemma 1b.
Q.E.D.

The requirement that the minimax strategies must be pure action profiles can be dropped if we furthermore assume that the randomization of individual players is observable.

The outcome of cooperative equilibria supported by *two-stage strategies* cannot be pareto-efficient. The "punishment" has to be carried out at the beginning of every group, although no agent of r ever deviated. Nevertheless, in the limit we can approximate all pareto efficient payoff profiles π with $\pi \in V^{NE}$ as an expected average payoff in an equilibrium arbitrary closely:

Theorem 1

For any payoff vector $\bar{\pi}$ in the interior of V^{NE}, there are values of $\bar{\delta} < 1$ and $\bar{\sigma} > 0$, such that a π with $\pi_i \geq \bar{\pi}_i$ for $i \in \{1, ..., n\}$ can be supported as an expected average payoff of new players of class i in a subgame perfect equilibrium of the game $\Gamma(g, \delta, \sigma, \infty)$ whenever $\delta > \bar{\delta}$ and $\sigma < \bar{\sigma}$.

Proof: see Appendix.

It is important to see that π must be in the interior of V^{NE} even in the no-discounting case as long as σ is positive: A group is matched together for finitely many periods with probability 1, and at the beginning of every match there is always a period of reduced payoffs.

Payoff profiles $\pi(y)$ with $\pi(y) \notin V^{NE}$ can also be approximated as an expected average payoff when y is implementable by a first stage profile with the properties required in lemmas 1b or 1c. However, a convex combination of such payoff profiles may not be implementable as lemma 1b and 1c place additional requirements on the first stage profile x. Thus, in general not all pareto efficient outcomes in the interior of V can be approximated as an expected average payoff.

2.3.2 Extension: Efficient equilibria in *two-stage strategies*

For fixed values of $\delta < 1$ and $\sigma > 0$ we cannot expect every payoff profile $\pi \in V^{NE}$ to be implementable. It also may not be the utility maximizing strategy to choose a pareto optimal second stage profile y if it is implementable. The reason for this is that there may be—depending on the stage game g—a tradeoff between the size of payoffs and the required number of periods in which the inferior profile x is played. In this section we will try to characterize the optimal behavior of agents in games $\Gamma(g, \delta, \sigma, \infty)$.

We first state the maximal expected average payoff $\bar{\pi}_i$ for a new player of class i when a *two-stage strategy* is played which uses x as first and y as second stage profile:

Lemma 2 *For given values of δ and σ, let a profile $y \in A$ be implementable with $x \in \Sigma$ as first stage profile. Then, in a subgame perfect equilibrium of $\Gamma(g, \delta, \sigma, \infty)$, the maximal expected average payoff for a new player of class i attainable by a two-stage strategy using those two profiles is given by*

$$\bar{\pi}_i = \pi_i(x) + \delta^T (1 - \sigma)^{Tn} [\pi_i(y) - \pi_i(x)], \qquad (2.14)$$

where $T \in \mathbb{R}$ is the minimal value such that

$$(1 - \sigma)^n \frac{\delta - \delta^{T+1}(1 - \sigma)^{Tn}}{1 - \delta(1 - \sigma)^n} \geq \frac{\pi_i(y_i^b, y_{-i}) - \pi_i(y)}{\pi_i(y) - \pi_i(x)} \qquad (2.15)$$

holds for all classes i.

The respective calculations can be taken from the proofs of theorem 1 and lemma 1a.

Assume that in a certain stage game g there are two profiles y and $\bar{y}$ implementable with $\pi_l(y) = \pi_l(\bar{y})$ for all classes $l \neq i$ and $\pi_i(y) < \pi_i(\bar{y})$ for one class i. Additionally, for at least some class j, a deviation would be more profitable, i.e.

$$\pi_j(\bar{y}_j^b, \bar{y}_{-j}) - \pi_j(\bar{y}_j, \bar{y}_{-j}) > \pi_j(y_j^b, y_{-j}) - \pi_j(y_j, y_{-j}).$$

As equation (2.15) has to hold, we would have to increase T. From (2.14) we immediately see that the expected average payoff for each class of players decreases whenever T increases. Thus, there is a tradeoff between the payoff of the second stage profile and the required duration of the first stage: The optimal *two-stage strategy* in a game $\Gamma(g, \delta, \sigma, \infty)$ depends on the payoff structure of the stage game g and the parameters δ and σ.

Example 2. Consider the following game where player 1 chooses rows and player 2 chooses columns:

	D	C1	C2
D	0,0	5,-1	6,-5
C1	-1,5	4,4	0,0
C2	-4,12	0,0	5,4

There are two cooperative profiles, $\{C1, C1\}$ and $\{C2, C2\}$, where the second one is more favourable to player 1. Assume that $\delta = 0,75$ and $\sigma = 0,01$. With (2.15) one can verify that with these values both $\{C1, C1\}$ and $\{C2, C2\}$ can be implemented with the Nash outcome $\{D, D\}$ as first stage profile. In order to implement the second profile, the duration of the first stage would be 4.15 periods and the corresponding expected average payoff for the first player—given by (2.14)—would be 1.39. This is considerably smaller than if we would implement $\{C1, C1\}$, which only requires 0.31 periods of Nash play and therefore yields an expected average payoff of 3.64. [**End of Example**]

Up to now we have only considered a certain type of strategy. In a game $\Gamma(g, \delta, \sigma, \infty)$ there may be other strategy combinations which implement a profile $y \notin \Sigma$ as well and yield higher expected utility for at least one class of players than a *two-stage strategy*. In this case, we say that such a strategy combination f dominates the *two-stage strategy*. One can rule out this possibility for some stage games:

Lemma 3a *Assume that in a stage game g a profile $y \in \Sigma$ is implementable by a two-stage strategy with $x \in \Sigma$ as first stage profile. If the number*

$$\frac{\pi_i(y_i^b, y_{-i}) - \pi_i(y)}{\pi_i(y) - \pi_i(x)} \tag{2.16}$$

is equal for all i, then there is no strategy combination f which implements y and dominates all two-stage strategies using x as first and y as second stage profile.

We show this result informally: The assumption on the payoffs ensures that we can choose T in a *two-stage strategy* such that (2.15) holds with equality for all i, i.e. all players are indifferent between deviating or not whenever y is played. If for another strategy f, the sum of discounted expected utility of a new player is higher, this implies that the expected discounted utility from the periods in which y is not played with probability 1 is also larger for at least one class i. But then the agents of this class can gain by playing a best response and dissolving the match whenever y is supposed to be played.

This result is sensible for symmetric games, but unfortunately does not hold generically. However, with additional requirements we can generalize it for two player games. By using mixed profiles we can adjust the expected period payoff of the first stage profile x such that both players are indifferent between conforming and deviating when the second stage profile y is played.

Lemma 3b *Let g be a two-player game and $y \in \Sigma$ with $\pi(y) \in V^{NE}$ a pure action profile. Assume that there are two profiles $x^i \in A$, $i \in \{1,2\}$, where x^i_i is a best response of player i to profile x^i_{-i}, with the following properties:*

$$\pi_i(x^{-i}) < \pi_i(y) < \pi_i(x^i) \tag{2.17}$$

$$\pi_i(x_i^{b,-i}, x_{-i}^{-i}) - \pi_i(x^{-i}) < \pi_i(y) - \max\{\pi_i(x^{NE}), \pi_i(x^{-i})\} \tag{2.18}$$

for some profile $x^{NE} \in \Sigma^{NE}$ with $\pi_i(x^{NE}) < \pi_i(y)$ for both i. Then, there are values $\bar{\delta} < 1$ and $\bar{\sigma} > 0$, such that there is no other strategy combination f which implements y and dominates every two-stage strategy whenever $\delta > \bar{\delta}$ and $\sigma < \bar{\sigma}$.

Proof: see Appendix.

Requirement (2.17) says that for each player, there is a profile which favors her relative to the second stage profile y and punishes the opponent at the same time. Condition (2.18) is needed to make deviations from the first stage profile unattractive by delaying the second stage profile by one period.

Lemmas 3a and 3b establish that for certain stage games—including versions of the PD which satisfy the respective conditions—and sufficiently high [low] values of δ [σ], there cannot be a better strategy than to establish a long-term relationship in the game $\Gamma(g, \delta, \sigma, \infty)$ by using an appropriate *two-stage strategy*. The corresponding empirical observation would be that players who have a larger incentive to deviate in the second stage have a relatively smaller utility at the beginning of a relationship without being rewarded for that in later periods.

In order to formulate the result as general as possible, we required the conditions (2.17) and (2.18) to be fulfilled by the stage game g. Neither is necessary as our next example shows:

Example 3. In the stage game above, the first cooperative profile exhibits the payoff structure required in lemma 3a: Whenever this profile is implementable by a *two-stage strategy* using $\{D, D\}$ as first stage profile, T can be chosen such that both players are indifferent between deviating or not when $\{C1, C1\}$ is played.

The same is not true for $\{C2, C2\}$. However, by playing the profile $\{D, C2\}$ with a probability of (approx.) 0.68 and $\{D, D\}$ with a probability of (approx.) 0.32 in every period of the first stage, agents can adjust their period payoffs such that both again are indifferent in the second stage. The expected payoffs for this first stage profile are then $\pi_1(x) \approx 4.08$ and $\pi_2(x) \approx -3.4$. One can check that with $\delta = 0.95$, no player ever would deviate from this *two-stage strategy*. However, with $\delta = 0.75$ the duration of the first stage would be 1.6 periods and the second player would deviate whenever the profile $\{D, C2\}$ is supposed to be played. Therefore, this strategy would not be available in example 2. [**End of example**]

2.3.3 Extension: Bilateral rationality of *two-stage strategies*

If we consider the first phase of a *two-stage strategy*, it may strike that players have an incentive to start with the cooperative profile y immediately after the match is formed given that all other groups stick to the *two-stage strategy*. Assume for the moment that y is a pure action profile. Gosh and Ray (1996) establish the criterion of "bilateral rationality" of strategies in games which are very similar to $\Gamma(g, \delta, \sigma, \infty)$:

A strategy is said to satisfy bilateral rationality if, given all other groups follow this strategy, no group of agents who have not deviated before in their current relationship, can propose a change in the strategies that both increases the sum of discounted expected payoffs and satisfies the individual incentive constraints.

However, we see from (2.15) that by the public signal, T can be chosen such that at least one class of players is indifferent between deviating or not when y is played. For these players, it does not violate any incentive constraint if their strategy prescribes to choose a best response to y_{-i} whenever profile y is supposed to be played before the end of the first stage as a result of communication. As the group dissolves upon the event of a deviation, no player has the incentive to skip the first stage given that all other groups stick to the *two-stage strategy*. Thus, the play of a *two-stage strategy* is robust against communication.

2.4 Finitely many players

2.4.1 Extension of the standard Folk Theorems

In the last section we considered a limit case in which there are infinitely many players and hence, the probability of meeting the same person a second time after the dissolution of a relationship with her is zero. This section is devoted to games $\Gamma(g, \delta, \sigma, m)$ with $m < \infty$.

With finitely many groups, there is a positive probability that agents of a match r are paired up again with each other in $t+1$ although they chose Q in t. Assume that the player of class i in the match r quits the relationship in period t, while all other members of r choose to maintain it. If the agents of all other $m-1$ groups also choose to maintain their relationships in t, the probability for the single player (given that she is not deleted from the set of players in t) of being matched to the same opponents in $t+1$ is given by

$$(1-\sigma)^{n-1} \sum_{k=0}^{m-1} \left(\frac{1}{k+1} \right) \binom{m-1}{k} \sigma^k (1-\sigma)^{m-1-k} \tag{2.19}$$

which can be transformed to

$$(1-\sigma)^{n-1} \frac{1-(1-\sigma)^m}{\sigma m}. \tag{2.20}$$

If all players of the match r choose to quit the relationship in t, then the probability that she is matched to the agents of r again in $t+1$ (given that she is not deleted from the set of players in t) is

$$A(\sigma) = (1-\sigma)^{n-1} \left[\frac{1-(1-\sigma)^m}{\sigma m} \right]^n. \tag{2.21}$$

Accordingly, $1-A(\sigma)$ is the probability that at least one player of r is matched to another group in $t+1$.

Recall that it is possible for each agent to identify former opponents. In order to characterize equilibrium strategies, we then have to specify the players' actions when meeting again. It is not certain that a deviating player finds a new group of opponents immediately and hence, it is possible to punish deviant behavior (for example by playing a minimax profile). This punishment can be carried out as long as the group is matched together. With this mechanism, it is possible to establish the Folk Theorem:

Theorem 2

(a) For finite m and any $\pi \in V^{NE}$, there are values $\bar{\delta} < 1$ and $\bar{\sigma} > 0$, such that there is a subgame perfect equilibrium of the game $\Gamma(g, \delta, \sigma, m)$ with π_i as the average payoff of class i whenever $\delta \geq \bar{\delta}$ and $\sigma \leq \bar{\sigma}$.

(b) Assume further that the dimension of V equals n. Then the same holds for any $\pi \in V^*$.

Proof: see Appendix.

Part (a) of the theorem is an extension to Friedman's (1971) Folk Theorem. A profile y with $\pi(y) \in V^{NE}$ is played in all periods by all groups. Whenever a player of class i of a given match r deviates from y_i, her opponents choose a (possibly mixed) Nash profile x with

$$\pi_i(x) < \pi_i(y) \tag{2.22}$$

for all i until r is dissolved[1]. Additionally, every player of r chooses Q after a deviation: As (2.22) holds for all classes of players, there is no incentive to choose M. If σ is sufficiently small the expected number of periods until the players of r are not matched together again is long enough to deter every agent from deviation.

Part (b) is a modified version of the Folk Theorem of Fudenberg and Maskin

[1]We do not require players to punish a deviant player whenever matched with her in *all* future periods. The reason for this is robustness: If agents make mistakes with arbitrary small probability $\epsilon > 0$ and we take the limit of $\sigma \to 0$ we would face a situation in which all matches play the Nash profile and do not return to the cooperative outcome for a period approaching infinity. Hence, for all $\epsilon > 0$ the theorem no longer would hold.

(1986): If the space of feasible payoffs is of full dimension, then it is possible to compute strategies, in which punishing a deviator can be rewarded by future gains without rewarding the deviator herself. Compared to the strategies of the original version of this theorem, we add that as long as a kind of punishment is carried out, all agents of the match choose to quit the relationship. This preserves the subgame perfectness of the equilibrium. Again, σ has to be sufficiently small such that the expected duration of every phase of punishment is sufficiently long.

As these strategies do not require players to start with an inferior profile at the beginning of a relationship, we get a complete Folk Theorem, i.e. we can also find support for payoff vectors at the boundary of V^{NE} (or V^* respectively) as expected average payoffs.

The role of the option to maintain or to dissolve a relationship differs between the infinite and finite setting: The expected payoff from a deviation is reduced in equilibrium by the fact that all agents outside the match may choose to maintain their relationship and thus it is harder to find new opponents. With infinitely many agents, the dissolution was the only available punishment, while with finitely many players, it offers an opportunity to evade the punishment of the opponents.

2.4.2 Extension: An application of the model to a labor market

We now apply the above reasoning informally to a stylized version of the Stiglitz-Shapiro (1984) model of a labor market. Assume that there are three types of agents: Employers, type-A workers and type-B workers. The stage game is as follows: When paired up with two workers, the employer can decide about the employment and wage payment w_A, w_B to each of them. The wage can take on three values, $w_A, w_B \in \{0, w_p, w_{eff}\}$, where

$$1 < w_p < w_{eff} \tag{2.23}$$

and a payment of 0 means that the worker is not hired. We call w_{eff} the "efficiency wage".

After wages are paid out those workers who received a positive wage decide whether they exert effort e or not, i.e. $e_A, e_B \in \{0, 1\}$. Exerting effort causes personal costs equal to e for the worker. If both exert effort, then the revenue to the employer is G. If only one worker exerts effort, it is H and if both choose $e = 0$, then it is equal to 0. We furthermore assume that

$$\begin{aligned} G - 2w_p &> H - w_p, &(2.24)\\ w_{eff} - 1 &> w_p, &(2.25)\\ G - 2w_{eff} &< 0, &(2.26)\\ H - w_{eff} &> 0. &(2.27) \end{aligned}$$

Equation (2.24) specifies that it is more profitable to hire two workers if both exert effort, (2.25) implies that working at w_{eff} is better than shirking at w_p, (2.26) rules out that the employer pays the efficiency wage to both workers and finally, (2.27) assures that hiring one worker at w_{eff} generates positive profits as long as she exerts effort. For simplicity, we furthermore assume that all employers can credibly commit to pay a uniform wage to employed workers. The only Nash equilibrium of the stage game is of course $\{(w_A = 0, w_B = 0), e_A = 0, e_B = 0\}$.

After revenues have been realized, each party can decide whether to maintain the relationship or not. If an agent quits the relationship, then she will be matched randomly to a match with the respective vacant position. We assume that each agent is replaced by a new player with positive probability.

Assume first that there are infinitely many agents. As stated in the first observation, with $\delta < 1$ and $\sigma > 0$ it is not possible that the profile $\{(w_p, w_p), 1, 1\}$ is played by all matches in all periods as each worker would have an incentive to choose $e = 0$. There are two possibilities to improve the situation: The employers may agree to offer w_{eff} to only one of the two workers: Whenever a new match is formed, the employer randomizes whether to employ type-A or type-B with equal probability. Furthermore, each employer threatens to dissolve the relationship with the employed worker at the end of the period if she does not exert effort. From (2.25) it follows that if δ is sufficiently high and σ is sufficiently low, then this worker would exert effort in every period. Note that by (2.26) no employer has an incentive to hire more than one worker.

Alternatively, a *two-stage strategy* can be played by all matches with the Nash equilibrium of the stage game as the first stage profile and $y = \{(w_p, w_p), 1, 1\}$ as the second stage profile. Again, for appropriate values of δ and σ, the latter profile is implementable. Both options result in (involuntary) unemployment as predicted by the Stiglitz-Shapiro model.

Consider now a situation with finitely many agents. Now, it is not certain that an agent finds a new employer immediately after she quits the relationship. From Theorem 2 it follows that a simple grim-trigger strategy resolves the problem: The employer offers w_p to both workers in each period. If one of them fails to exert effort at some point in time, then she will not be hired as long as the match exists in subsequent periods. As $w_p - 1 > 0$, in these periods the workers payoff is strictly smaller compared to the case of employment. If σ is sufficiently small and δ is high enough, then the gain from not exerting effort is negative as the expected time of unemployment is too long. Thus, we would never observe unemployment in the population.

2.4.3 Extension: On the dependence of cooperative equilibria on m when δ and σ are fixed

Assume that the values of δ and σ are fixed. Consider for a given stage game g a cooperative profile $y \notin \Sigma^{NE}$ such that

- y is implementable by the strategies of theorem 2 when the number of groups is below m^* and
- y is implementable by a *two-stage strategy* using x as the first stage profile in the case of $m = \infty$.

Now increase m: From (2.21) we can see that for $m \to \infty$ we get $A(\sigma) \to 0$. Thus, the repeated play of y in all periods by all groups no longer can be supported as an equilibrium. However, the *two-stage strategy* remains capable to implement y as it does not rely on punishment by a certain group of opponents. For finite m, there is still a positive chance of being matched with the same group after a dissolution. Therefore, we have to specify the reaction to this event when an opponent of class j has deviated in the previous period: A combination of the strategies in chapters 3 and 4 can implement y for $m > m^*$: Let $z^1, ..., z^n$ be a collection of (possibly mixed) Nash profiles such that $\pi_i(x) < \pi_i(y)$ and $\pi_i(z^i) < \pi_i(y)$ holds for all i. Then, take the following strategy for a player of class j:

Phase I: When a new match is formed, play x_j for T periods and M in every period. If in this period there is a unilateral deviation from x by a player of class i, switch to phase III_i. Whenever the match dissolves in this phase start in phase I again.

Phase II: Play y_j and M in every period. If there was a unilateral deviation from y by player i in the previous period, switch to phase III_i. Whenever the match dissolves in this phase start in phase I again.

Phase III_i: Play z^i_j and Q in every period as long the same group r is matched together. If matched to a new group start in phase I again.

By using the same methods as in the proofs of lemma 1a and theorem 2, one can establish that this strategy supports a subgame perfect equilibrium for appropriate values of δ and σ if T is chosen sufficiently high. With this strategy, deviators are punished twice: By the current opponents as long they are matched together, and in the new group when profile x is played. T is chosen such that any gain is compensated by the expected losses of these two punishments. Note that T can be shorter than in a pure *two-stage strategy*. In equilibrium, the third phase is never carried out.

With above strategy, the expected duration of phase III_j decreases in m. This requires phase I to be extended when the population grows. Thus, the maximal sum of expected discounted payoffs depends indirectly on the number of players.

2.5 Overview of the related literature

The strategies used in the third chapter to establish cooperation in an environment with infinitely many players was subject of various papers. Datta (1993) and Kranton (1996) use a similar model and apply it to two player versions of the PD. Ray and Gosh (1996) also consider a modified version of the PD as the stage game, but their agents are either myopic or patient. Thus, it is rational for the latter group to “test” their opponent in the first period of the relationship by playing cooperative profiles with lower payoffs. Strategies are similar to ours, but the rationale for them is imperfect information about the opponents’ time preferences.

Similar strategies for heterogenous agents are obtained in the papers by Watson (1999, 2002) and Sobel (1985). These papers do not model the population of agents explicitly. Matsushima (1990) obtains a Folk Theorem for the infinite setting, but he allows agents to have more information about their opponents. Casas-Arce (2005) considers a principal-agent relationship, in which every party has the option to quit the relationship and the outside options are given exogenously. Differently to this, in our setting the value of the outside option is given by the equilibrium strategies.

Our approach establishes the difference between models with finitely and infinitely many agents. Furthermore, we focus on the equilibrium set of average payoff profiles and take up a general perspective: We allow for any stage game with complete information and any size of the population, while maintaining the game-theoretic terminology. For the case of infinitely many agents, we explicitly showed the optimality of the considered strategy for many stage games of interest.

2.6 Conclusion

With theorems 1 and 2, we obtain a theory of implementation of profiles $y \notin \Sigma^{NE}$ in games $\Gamma(g, \delta, \sigma, m)$ for every population size: The repeated play of profiles y with $\pi(y) \in V^{NE}$ is possible for any size of the population given that players both discount future gains sufficiently little and remain in the set of players for a sufficiently long time. Nevertheless, the structure of supporting strategies is quite different.

We first considered the case when players have the opportunity to switch to other opponents immediately. Before choosing a mutually beneficial profile, agents may first play a suboptimal action for considerable time. Although it is possible to approximate many efficient payoff vectors arbitrarily close, it is not possible to establish a complete Folk Theorem.

This changed when we considered the case of finitely many agents. We showed that by a small modification of well-known strategies, we obtain again a version of the standard Folk Theorems: If the rate of change in the set of players is small, then an agent cannot be sure to find new opponents immediately. Punishment by the same opponents becomes available to some extent. Thus, for appropriate values of δ and σ, cooperative profiles can be implemented without any efficiency loss when m is finite.

The role of the additional option to quit or to maintain the relationship in the supporting strategies was quite different in the two settings: While in the infinite version of the model it constituted the only way to punish deviant behavior, it provided a possibility to stop the punishment phase in the finite setting. With an example, we showed that this difference may have an impact on the equilibrium outcome of markets.

2.7 Appendix

Proof of Lemma 1a

In the first step we will treat y as pure action profile, and then extend the result to mixed profiles. If $y \in \Sigma^{NE}$, the claim holds trivially. Suppose this is not the case: As $\pi(y) \in V^{NE}$, there is a (possibly mixed) Nash profile x such that for all i

$$\pi_i(y_i^b, y_{-i}) \geq \pi_i(y) > \pi_i(x). \tag{2.28}$$

The supporting strategy for every player of class i is as follows:

Phase I**:** If your match is formed in period t, play the profile x_i and M for T periods. Switch to phase II in period $t+T$. If the player of class j of this group deviated in period $\tau \in [t, t+T-1]$ from x_j, choose Q in τ. If the group dissolves in this phase, start in phase I again.

Phase II**:** Choose y_i and M in every period. If any player of class j of the group deviated in a preceding period from y_j, chose Q. If the group dissolves in this phase, start in phase I again.

We apply the one stage deviation principle to show that this strategy is played by all agents in a subgame perfect equilibrium of $\Gamma(g, \delta, \sigma, \infty)$ if δ is sufficiently high, σ is low enough and T is chosen sufficiently large.

In phase I, all agents play a best response. Any deviation from the prescribed action profile would restart phase I. Because of (2.28), this can never be profitable.

Consider now a period t in phase II. Set $t = 0$. The gain from playing a best response in this period is

$$\Delta_{0,i} = \pi_i(y_i^b, y_{-i}) - \pi_i(y). \tag{2.29}$$

If i deviates in 0, she will be matched to a new group which again starts to play x for T periods. Denote the minimal difference between the sums of discounted expected payoffs of a conforming and a deviating player with

$$\Delta_i = E_i(a_i^0 = y_i) - E_i(a_i^0 = y_i^b). \tag{2.30}$$

To compute Δ_i, we have to sum up over all differences $\Delta_{tT,i}$ and $\Delta_{T,i}$ which can occur due to a (involuntary) dissolution in the periods $t \in \{1, 2, ..T\}$ and $t \in \{T+1, T+2, ...\}$. Given that the match of the conforming player does not dissolve in period 0, the probability of this event for a period $t > 0$ is equal for both players. Thus, we have

$$\Delta_i = \Delta_{0,i} + (1-\sigma)^n[\Delta_{tT,i} + \Delta_{T,i}]. \tag{2.31}$$

Given that the match of the conforming player does not dissolve in period 0, the probability that r is dissolved in period t with $t > 0$ is given by

$$(1-\sigma)^{(t-1)n}[1-(1-\sigma)^n]. \tag{2.32}$$

In the event of a dissolution of r in t with $0 < t \leq T$, the sum of expected discounted payoffs from period 1 onwards of a conforming player is

$$\sum_{\tau=1}^{t} \delta^\tau \pi_i(y) + \delta^{t+1}(1-\sigma)E_i((x,y),T), \tag{2.33}$$

while for a deviating player it is given by

$$\sum_{\tau=1}^{t} \delta^\tau \pi_i(x) + \delta^{t+1}(1-\sigma)E_i((x,y),T). \tag{2.34}$$

The difference is therefore

$$\sum_{\tau=1}^{t} \delta^\tau (\pi_i(y) - \pi_i(x)). \tag{2.35}$$

When adding this difference up over all periods $0 < t \leq T$, we get

$$\Delta_{tT,i} = [\pi_i(y) - \pi_i(x)](1-(1-\sigma)^n)\left(\sum_{t=1}^{T}(1-\sigma)^{(t-1)n}\sum_{\tau=1}^{t}\delta^\tau\right) =$$
$$= [\pi_i(y) - \pi_i(x)]\left(\frac{\delta - \delta^{T+1}(1-\sigma)^{Tn}}{1-\delta(1-\sigma)^n} - (1-\sigma)^{Tn}\frac{\delta-\delta^{T+1}}{1-\delta}\right) \quad (2.36)$$

For a dissolution of r in $t > T$ the difference in payoffs between the deviating and the conforming player is equal to $[\pi_i(y) - \pi_i(x)]\frac{\delta-\delta^{T+1}}{1-\delta}$. The probability of a dissolution after period T is given by $(1-\sigma)^{Tn}$ and hence,

$$\Delta_{T,i} = [\pi_i(y) - \pi_i(x)](1-\sigma)^{Tn}\left(\frac{\delta-\delta^{T+1}}{1-\delta}\right). \quad (2.37)$$

Thus, we get

$$\Delta_i = [\pi_i(y) - \pi_i(y_i^b, y_{-i})] + (1-\sigma)^n[\pi_i(y) - \pi_i(x)]\frac{\delta-\delta^{T+1}(1-\sigma)^{Tn}}{1-\delta(1-\sigma)^n}. \quad (2.38)$$

It is not profitable for any player to deviate in the second phase if $\Delta_i \geq 0$ for all i. Therefore, it must be the case that for any class i

$$(1-\sigma)^n\frac{\delta-\delta^{T+1}(1-\sigma)^{Tn}}{1-\delta(1-\sigma)^n} \geq \frac{\pi_i(y_i^b, y_{-i}) - \pi_i(y)}{\pi_i(y) - \pi_i(x)}. \quad (2.39)$$

Taking the limits on the left-hand side, we get

$$\lim_{\sigma\to 0}(1-\sigma)^n\frac{\delta-\delta^{T+1}(1-\sigma)^{Tn}}{1-\delta(1-\sigma)^n} = \frac{\delta-\delta^{T+1}}{1-\delta} \quad (2.40)$$

and

$$\lim_{\delta\to 1}\frac{\delta-\delta^{T+1}}{1-\delta} = T. \quad (2.41)$$

Thus, there is a $T < \infty$, such that $\Delta_i \geq 0$ holds for all classes i. Choose a T^* such that

$$T^* > \max_i\left(\frac{\pi_i(y_i^b, y_{-i}) - \pi_i(y)}{\pi_i(y) - \pi_i(x)}\right). \quad (2.42)$$

As δ enters the left-hand side of (2.41) continuously, there is a $\delta' < 1$ such that for all $\delta \geq \delta'$

$$\frac{\delta-\delta^{T^*+1}}{1-\delta} \geq \max_i\left(\frac{\pi_i(y_i^b, y_{-i}) - \pi_i(y)}{\pi_i(y) - \pi_i(x)}\right). \quad (2.43)$$

Fix a $\bar{\delta} \in (\delta', 1)$. As σ enters the left-hand side of (2.40) continuously, there is a $\bar{\sigma} > 0$ such that (2.39) holds for all i whenever $T = T^*$, $\delta \geq \bar{\delta}$ and $\sigma \leq \bar{\sigma}$. This completes the argument for pure profiles x and y.

Consider now a mixed profile y with support $Y \subseteq \Sigma$ and $\pi_i(y) > \pi_i(x)$ for all classes i. Choose now a T^* with

$$T^* > \max_{y^l\in Y}\max_i\left[\frac{\pi_i(y_i^{b,l}, y_{-i}^l) - \pi_i(y)}{\pi_i(y) - \pi_i(x)}\right] \quad (2.44)$$

and follow the same steps as in the first part to obtain the same result.
Q.E.D.

Proof of Lemma 1b

With the one stage deviation principle we prove that under assumption (2.10), a deviation from the first stage profile x is not profitable if δ is high and σ is low enough. Then from lemma 1a, the same follows for the second stage profile y when T is chosen sufficiently high.

Let $t = 0$ be the period in which a player of class i deviates. By going through the same steps as in the proof of lemma la, we get that the difference in the sum of expected discounted payoffs between a conforming and a deviating player, $\Delta_i = E_i(a_i^0 = x_i) - E_i(a_i^0 = x_i^b)$, is given by

$$\Delta_i = [\pi_i(x) - \pi_i(x_i^b, x_{-i})] + (1-\sigma)^{Tn}\delta^T[\pi_i(y) - \pi_i(x)]. \tag{2.45}$$

With (2.10) this becomes positive when δ is close enough to 1 and σ is close enough to 0. The rest is analogous to the proof of lemma 1a. For mixed profiles, the condition in (2.12) ensures that

$$\Delta_i = [\pi_i(z) - \pi_i(z_i^b, z_{-i})] + (1-\sigma)^{Tn}\delta^T[\pi_i(y) - \pi_i(x)] \tag{2.46}$$

is positive for all $z \in X$ when δ is close enough to 1 and σ is close enough to 0.
Q.E.D.

Proof of Theorem 1

The expected utility $E_i((x,y),T)$ of a new agent from a *two-stage strategy* with x as the (possibly mixed) Nash profile, y as the (possibly mixed) second stage profile and T as the duration of the first stage, is given by

$$\begin{aligned}
&(1-(1-\sigma)^n)\sum_{t=0}^{T-1}\left[(1-\sigma)^{tn}\sum_{\tau=0}^{t}\delta^\tau\pi_i(x)\right] + \\
&+(1-(1-\sigma)^n)\sum_{t=T}^{\infty}\left[(1-\sigma)^{tn}\left(\sum_{\tau=0}^{T-1}\delta^\tau\pi_i(x) + \sum_{\tau=T}^{t}\delta^\tau\pi_i(y)\right)\right] + \\
&+(1-\sigma)(1-(1-\sigma)^{n-1})\sum_{t=1}^{\infty}\delta^t(1-\sigma)^{(t-1)n}E_i((x,y),T),
\end{aligned} \tag{2.47}$$

where we set the intial period $t = 0$. This expression simplifies to

$$E_i((x,y),T) = \frac{\pi_i(x) + \delta^T(1-\sigma)^{Tn}[\pi_i(y) - \pi_i(x)]}{1-\delta(1-\sigma)}. \tag{2.48}$$

By multiplying (2.48) with $1 - \delta(1-\sigma)$, we get the expected average payoff of a new player of class i:

$$\bar{\pi}_i = \pi_i(x) + \delta^T(1-\sigma)^{Tn}[\pi_i(y) - \pi_i(x)]. \tag{2.49}$$

By taking the limit, we get

$$\lim_{\sigma\to 0}\bar{\pi}_i = (1-\delta^T)\pi_i(x) + \delta^T\pi_i(y). \tag{2.50}$$

For each $\bar{\pi} \in int(V^{NE})$, we can find (possibly mixed) profiles y and x such that $\bar{\pi}_i > \pi_i(x)$ and $\bar{\pi}_i < \pi_i(y)$ hold for all i and x is a Nash profile. From lemma 1a we know that a *two-stage strategy* with x as the first and y as the second stage profile constitudes a subgame perfect equilibrium if δ is sufficiently large, σ is small and T is chosen large enough. As σ enters (2.49) continuously, the claim follows from (2.50).
Q.E.D.

Proof of Lemma 3b

First, denote $y^i = (y_i^b, y_{-i})$ for $i \in \{1,2\}$. The logic ot the proof is the same as in lemma 3a. We construct a mixed profile x such that for $i \in \{1,2\}$,

$$\pi_i(x) = \pi_i(y) - c[\pi_i(y^i) - \pi_i(y)] \tag{2.51}$$

holds for some $c > 0$. From (2.15) it then follows that we can choose T such that both players are indifferent in the second stage between deviating or not.

As there is a Nash profile x^{NE} with $\pi_i(x^{NE}) < \pi_i(y)$ for both players, we find a unique $\bar{c} > 0$ with

$$\pi_1(y) - \bar{c}[\pi_1(y^1) - \pi_1(y)] = \pi_1(x^{NE}). \tag{2.52}$$

We distinguish between two cases:

$$\pi_2(y) - \bar{c}[\pi_2(y^2) - \pi_2(y)] \quad > \quad \pi_2(x^{NE}), \tag{2.53}$$
$$\pi_2(y) - \bar{c}[\pi_2(y^2) - \pi_2(y)] \quad < \quad \pi_2(x^{NE}). \tag{2.54}$$

If (2.54) holds consider a profile x which specifies to play x^2 with probability β and x^{NE} with probability $1 - \beta$ in every period of the first stage. Rewrite (2.51) as

$$\pi_i(y) - c[\pi_i(y^i) - \pi_i(y)] = \beta\pi_i(x^2) + (1-\beta)\pi_i(x^{NE}). \tag{2.55}$$

Set $i = 1$ and solve this expression for c. Then, using the resulting term in the same equation for player 2 yields us

$$\pi_2(y) - \frac{\pi_1(y) - \beta\pi_1(x^2) - (1-\beta)\pi_1(x^{NE})}{\pi_1(y^1) - \pi_1(y)}[\pi_2(y^2) - \pi_2(y)] = \beta\pi_2(x^2) + (1-\beta)\pi_2(x^{NE}). \tag{2.56}$$

With (2.17) and (2.54) one can verify that for $\beta \to 1$, the right-hand side of (2.56) exceeds the left-hand side and for $\beta \to 0$, the opposite is true. Hence, there is a $\beta \in \,]0,1[$ such that (2.56) holds and (2.51) is true for both players and for some c. By following the same steps one can establish the same result when (2.54) is true by using the same Nash profile x^{NE} and x^1. Condition (2.18) ensures that the resulting mixed profile x satisfies the requirements of lemma 1b. The claim then follows from lemma 1b.
Q.E.D.

Proof of Theorem 2

(a) Let $y \in \Sigma$ be a profile with $\pi(y) \in V^{NE}$ and x a (possibly mixed) Nash profile with $\pi(x) \in U^{NE}$ and $\pi_i(y) > \pi_i(x)$ for all i. The supporting strategy to establish $\pi(y)$ as an expected average payoff is as follows:

Phase I**:** Choose y_i and M in every period as long as no player of the current match r ever deviated from y. Whenever a player of class j of r deviated from y_j in a previous period, switch to phase II. If r is dissolved in this phase, start in phase I again.

Phase II**:** Choose x_i and Q in every period. Retain this strategy until matched to a new group. If r is dissolved in this phase, start in phase I again.

Again we use the one stage deviation principle to show that there is no opportunity to deviate profitably in phase I or phase II if δ is sufficiently high and σ sufficiently low. For phase II this is obvious, as every player chooses a best response in every period. As $\pi_i(y) > \pi_i(x)$ for all i, there is no incentive to maintain the match.

To establish the same for phase I, assume that a player of class i deviates in period t and set $t = 0$. We compare the utility of a conforming player $E_i(a_i^0 = y_i)$ and the maximal utility of a deviating player $E_i(a_i^0 = y_i^b)$. We treat y as a pure action profile, but the result easily extends to mixed profiles.

The sum of expected discounted payoffs of a conforming player is given by

$$E_i(a_i^0 = y_i) = \sum_{t=0}^{\infty} \sigma(1-\sigma)^t \sum_{\tau=0}^{t} \delta^\tau \pi_i(y), \tag{2.57}$$

which can be written as

$$E_i(a_i^0 = y_i) = \frac{\pi_i(y)}{1 - \delta(1-\sigma)}. \tag{2.58}$$

Consider now the deviating player: When she is matched to a new group in the next period, her continuation values is given by (2.58). Let E_i^{yc} be her continuation value after the deviation. Then

$$E_i(a_i^0 = y_i^b) = \pi_i(y_i^b, y_{-i}) + \delta(1-\sigma)E_i^{yc}, \tag{2.59}$$

where

$$E_i^{yc} = A(\sigma)[\pi_i(x) + \delta(1-\sigma)E_i^{yc})] + (1 - A(\sigma))E_i(a_i^0 = y_i). \tag{2.60}$$

Solving this for E_i^{yc} and using the resulting expression in (2.59) yields us

$$E_i(a_i^0 = y_i^b) = \pi_i(y_i^b, y_{-i}) + \frac{\delta(1-\sigma)}{1 - A(\sigma)\delta(1-\sigma)} \times \Big[A(\sigma)\pi_i(x) + (1 - A(\sigma))E_i(a_i^0 = y_i)\Big]. \tag{2.61}$$

Thus, both $E_i(a_i^0 = y_i)$ and $E_i(a_i^0 = y_i^b)$ are continuous in δ and σ. From (2.21) we get

$$\lim_{\sigma \to 0} A(\sigma) = 1. \tag{2.62}$$

With this we calculate for all $\delta < 1$

$$\lim_{\sigma \to 0} E_i(a_i^0 = y_i) = \pi_i(y) + \frac{\delta}{1-\delta}\pi_i(y), \tag{2.63}$$

$$\lim_{\sigma \to 0} E_i(a_i^0 = y_i^b) = \pi_i(y_i^b, y_{-i}) + \frac{\delta}{1-\delta}\pi_i(x). \tag{2.64}$$

Now fix a $\bar{\delta} < 1$ such that

$$\pi_i(y) + \frac{\bar{\delta}}{1-\bar{\delta}}\pi_i(y) > \pi_i(y_i^b, y_{-i}) + \frac{\bar{\delta}}{1-\bar{\delta}}\pi_i(x). \tag{2.65}$$

Then, from the continuity of $E_i(a_i^0 = y_i)$ and $E_i(a_i^0 = y_i^b)$ in σ, it follows that there exists a $\bar{\sigma} > 0$, such that for all $\sigma \leq \bar{\sigma}$ we have

$$E_i(a_i^0 = y_i) \geq E_i(a_i^0 = y_i^b) \tag{2.66}$$

whenever $\delta \geq \bar{\delta}$. It remains to mention that no agent can gain by choosing Q in phase I, as all matches play profile y in this phase.

(b) Now assume that $dim(V) = n$. Let $\pi(y) \in V^*$. Then there are (possibly mixed) profiles $(x^1, ..., x^n)$ such that

$$\pi_i(\underline{x}^i) < \pi_i(x^i) < \pi_i(y) \tag{2.67}$$

for all classes i and

$$\pi(x^i) = \big(\pi_1(x^1) + \epsilon, ..., \pi_{i-1}(x^{i-1}) + \epsilon, \pi_i(x^i), \pi_{i+1}(x^{i+1}) + \epsilon, ..., \pi_n(x^n) + \epsilon\big) \tag{2.68}$$

is an element of V. Further choose a $T \in \mathbb{N}$, such that for all i

$$\max_{a \in A} \pi_i(a) + T\pi_i(\underline{x}^i) < \min_{a \in A} \pi_i(a) + T\pi_i(x^i). \tag{2.69}$$

For simplicity we assume that the minimax profile for each player is a pure action profile. See Fudenberg and Tirole (1991, p. 159 - 160) for details how to drop this assumption.

The supporting strategy is as follows:

Phase I: Play y_i and M as long as there was no unilateral deviation from y by a member of class j of the current group r in the previous period. In that case switch to phase II_j. If r is dissolved in this phase, start in phase I again.

Phase II_j: Play $\underline{x}_i^j$ for T periods and choose Q in every period. Then, switch to phase III_j when there was no unilateral deviation from $\underline{x}_l^j$ by the player of class l in r. In that case switch to phase II_l. If r is dissolved in this phase, start in phase I again.

Phase III_j: Play x_i^j and Q in every period. If a single player of class l in r deviates from x_l^j, then switch to phase II_l. If r is dissolved in this phase start, in phase I again.

We check that in no phase a player can gain by deviating whenever δ is high and σ is low enough. In phase I player i receives $E_i(y) = E_i(a_i^0 = y_i)$ from in equation (2.58) by conforming and

$$\begin{aligned} \pi_i(y_i^b, y_{-i}) \quad &+ \quad \delta(1-\sigma)\left[\frac{A(\sigma)-\delta^T(1-\sigma)^T A(\sigma)^{T+1}}{1-A(\sigma)\delta(1-\sigma)}\pi_i(\underline{x}^i)+\right. \\ &+ \quad \left.\frac{\delta^T(1-\sigma)^T A(\sigma)^{T+1}}{1-A(\sigma)\delta(1-\sigma)}\pi_i(x^i)+\frac{1-A(\sigma)}{1-A(\sigma)\delta(1-\sigma)}E_i(y)\right] \end{aligned} \tag{2.70}$$

by deviating once. By taking the limits and applying the usual continuity argument, one can verify that this expression is smaller than $E_i(y)$ whenever δ is close enough to 1 and σ close enough to 0.

In phase II_j player i gets by conforming at least

$$\begin{aligned} \pi_i(\underline{x}^j) \quad &+ \quad \delta(1-\sigma)\left[\frac{A(\sigma)-\delta^T(1-\sigma)^T A(\sigma)^{T}}{1-A(\sigma)\delta(1-\sigma)}\pi_i(\underline{x}^j)+\right. \\ &+ \quad \left.\frac{\delta^T(1-\sigma)^T A(\sigma)^{T}}{1-A(\sigma)\delta(1-\sigma)}(\pi_i(x^i)+\epsilon)+\frac{1-A(\sigma)}{1-A(\sigma)\delta(1-\sigma)}E_i(y)\right], \end{aligned} \tag{2.71}$$

while a deviation yields her at most

$$\begin{aligned} \max_a \pi_i(a) \quad &+ \quad \delta(1-\sigma)\left[\frac{A(\sigma)-\delta^T(1-\sigma)^T A(\sigma)^{T+1}}{1-A(\sigma)\delta(1-\sigma)}\pi_i(\underline{x}^i)+\right. \\ &+ \quad \left.\frac{\delta^T(1-\sigma)^T A(\sigma)^{T+1}}{1-A(\sigma)\delta(1-\sigma)}\pi_i(x^i)+\frac{1-A(\sigma)}{1-A(\sigma)\delta(1-\sigma)}E_i(y)\right]. \end{aligned} \tag{2.72}$$

One can establish that for adequate values of δ and σ the former expression is larger than the latter one.

In phase III_j player i gets by conforming

$$\pi_i(x^i)+\epsilon+\delta(1-\sigma)\left[\frac{A(\sigma)}{1-A(\sigma)\delta(1-\sigma)}(\pi_i(x^i)+\epsilon)+\frac{1-A(\sigma)}{1-A(\sigma)\delta(1-\sigma)}E_i(y)\right], \tag{2.73}$$

while a deviation in this phase yields at most the level of expected utility given in equation (2.72). Again, for adequate values of δ and σ a deviation in phase III_j of a player of class i is not profitable.

Finally, in phase III_i player i gets by conforming

$$\pi_i(x^i)+\delta(1-\sigma)\left[\frac{A(\sigma)}{1-A(\sigma)\delta(1-\sigma)}\pi_i(x^i)+\frac{1-A(\sigma)}{1-A(\sigma)\delta(1-\sigma)}E_i(y)\right], \tag{2.74}$$

while a deviation yields at most the level of utility given in equation (2.72). The condition in (2.69) ensures that any gain is wiped out by the succeeding period of minimax payoffs for adequate values of δ and σ. Because of (2.67), no class i can gain by choosing Q in phase I or by playing M in phase II_i/II_j or III_i/III_j.
Q.E.D.

Chapter 3

Imitating Cooperation and the Formation of long-term Relationships

3.1 Introduction

In this paper, we study the following game: each agent of an infinite population plays the prisoner's dilemma with some opponent in each period. After observing the partner's action choice, each player has the option to maintain or to quit the current relationship. If the latter action is chosen by both agents, they play the game together in the next period, otherwise they return to a "market for long-term relationships" and are matched randomly with another opponent. The population therefore consists of long-term relationships and the market. Matching in the market is global and non-assortative: everybody can be matched with anybody else and own behavior does not affect the probability of being paired up with an agent who plays a certain strategy. Furthermore, there are no information flows between pairs.

This game is a general description of markets for private or professional relationships. However, only few studies have dealt with it in the literature on repeated games. For a summary, see Mailath and Samuelson (2006), chapter 5.2. In order to implement cooperation in this game, the literature suggests to "start small": at the beginning of a new relationship both agents defect and start to cooperate in later periods. Whenever a player deviates from this path of play, her opponent chooses to quit the relationship. Thus, any gain from deviation is wiped out by the subsequent phase of low payoffs in the new relationship.

In the considered framework, this solution has serious shortcomings. Without further assumptions, such a strategy is not robust against communication: given that all other players in the population stick to the described pattern, it is optimal for two agents who meet for the first time in the market, to start the relationship with cooperation. This would not violate any incentive constraint. However, if all pairs act in this way, we are no longer in equilibrium. Gosh and Ray (1996) solve this problem by introducing heterogeneous time preferences: a fixed share of players is myopic while the rest of the population is patient to some extent. Thus, a period of less cooperation at the beginning of a new relationship serves to "test" the opponent's patience. As long as we do not make this assumption, we do not have a convincing solution for the implementation of cooperation in the framework above.

Where this criticism is specific to the considered game, another disadvantage of the existing solution is more fundamental to game theory: the equilibrium concept requires common knowledge of strategies. Especially for large populations, this assumption is unrealistic. Conventional solutions leave open the question how agents manage to reach an equilibrium in the absence of abundant sophistication. Also note that in this game, there are many equilibria, including a situation in which agents never cooperate.

We avoid both problems by adopting an evolutionary approach. Players are boundedly-rational such that—according to Kandori et. al. (1993)—(i) not all agents adjust their behavior instantaneously, (ii) agents react myopically and (iii) with small probability, agents change their behavior randomly. Players sometimes have access to information about what strategy currently is on average the most successful one in the population. The mechanism which creates this information can be any kind of mass-communication. The goal is to show that cooperative behavior can prevail in the population if agents imitate successful behavior.

According to the concept of bounded rationality, we assume that players follow simple strategies which are not history dependent, like "never cooperate and quit every relationship" or "behave cooperatively and maintain the relationship if and only if your opponent cooperated as well". Cooperative long-term relationships immediately start with cooperation. The only "punishment" of non-cooperation is the breakup of a relationship. With positive probability, agents imitate the strategy which is currently the most successful one in the population. Additionally, they change their behavior randomly with small probability, i.e. they choose to "experiment" and quit their current relationship.

Our focus lies on the behavior of aggregate shares of agents who follow a certain strategy. With the law of large numbers in mind, we concentrate our analysis on deterministic approximations. The first result is that in almost all periods, all learning agents adapt to the cooperative strategy if only a small fraction of agents imitates in each period and the experimentation rate is sufficiently small. The intuition for this is the following: The share of cooperative long-term relationships only increases if the share of cooperators in the market is not too small over a larger number of periods, i.e. if cooperators in the market do not switch to the non-cooperative strategy too fast. Additionally, if too many agents imitate in each period, the share of cooperative players in the market—and therefore the average utility of non-cooperative players—is sufficiently high, such that cooperative long-term relationships are broken up, as many individuals switch to the non-cooperative strategy.

The result implies global convergence to a unique distribution of strategies in the population, where the cooperative strategy is played by most agents. It is independent of the size of payoffs and thus resembles the folk theorem for repeated games. We conclude that an environment with global interaction and non-assortative matching is less hostile to cooperation if agents change their behavior only infrequently. It is important to note that in a setting with myopic agents, a best response dynamic—like, for example, fictitious play or regret matching—does not solve the social dilemma, as the best response in the stage game is not to cooperate.

In order to relax the assumption on the imitation rate, we then introduce heterogeneity into the population: an agent does not want to interact with any agent of the population, but prefers to meet individuals with certain characteristics or manners. Long-term relationships only can exist between agents who "get along with each other". The period payoff a player receives increases by a fixed amount if she gets along with her current opponent, regardless of the strategies played by agents. Thus, the strategy "never cooperate and maintain the relationship if and only if you get along with your opponent", becomes attractive if the probability of meeting such an opponent in the market is sufficiently small. Under the assumption of heterogeneity of players, we maintain the same result as in the homogeneous setting, but without a boundary on the imitation rate: even though almost all agents switch to the most successful strategy in each period, learning agents adapt to the cooperative strategy in almost all periods if the experimentation rate is sufficiently small.

The dynamics in the market now are different: In a heterogeneous population, agents also maintain non-cooperative long-term relationships if the payoff from the fact that one is paired up with an opponent with whom she gets along, is high. The size of the market converges to a small value which in turn decreases in the experimentation rate. One can then derive a lower bound on the share of agents in the market which plays according to the cooperative strategy. This implies that the probability of meeting a cooperative player in the market, with whom one gets along, remains above a positive threshold for all values of the experimentation and imitation rate. Thus, for a sufficiently small experimentation rate, the number of cooperative long-term relationships increases until the average utility of cooperators is higher than the average utility of other strategies in all subsequent periods. Again we obtain global convergence to a unique distribution of strategies in the population.

This paper does not only provide an evolutionary solution for the repeated prisoner's dilemma with the option to maintain or to quit relationships, but also contributes to the literature on large-scale cooperation in social dilemmas. The work closest to ours is by Eshel et. al. (1998), where each agent has a number of neighbors and imitates the behavior of the most successful one whenever this neighbor's payoff is larger than own utility. The learning rule thus is similar to ours. Social interaction is structured and agents do not have the opportunity to change their neighbors. They show that if the experimentation rate is small, then the proportion of cooperating agents is between 70 and 87 percent after sufficiently many periods—independent of the initial distribution of strategies. Another model, in which cooperative behavior is established by imitation, is presented by Levine and Pesendorfer (2007). In their approach, players know to some extend the strategy her opponent will play, before choosing their own action. Thus, agents may condition their action on the information about their current opponent. In our approach, we abstract from any information flows and assume anonymity among players.

There are some evolutionary models in which cooperation is at least an evolutionary stable outcome. The considered mechanisms are green beards, kin-based selection, reciprocity, indirect reciprocity and punishment, see Hendrich (2004) for a summary. In most of these models, a certain degree of assortative matching is required. An alternative approach, to explain cooperation in the PD in a large population with random matching, can be taken by assuming that agents have certain flexible aspirations regarding their payoffs. Palomino and Vega-Redondo (1999) show that there will be a significant share of cooperating players if agents become dissatisfied with the non-cooperative outcome.

However, we are confident that the option to maintain or to quit relationships provides an even more convincing answer to the question why cooperation in a large population emerges, as it can be observed quite often that individuals search for cooperative long-term relationships.

The rest of the paper is organized as follows: in the next chapter, we develop the model and state the result for the homogeneous population. We also show that it is dependent on the option to maintain or to quit relationships. Chapter 3 extends the model to a heterogeneous population and states the corresponding result. To illustrate the dynamics and to find out the parameter values which lead to the cooperative outcome, we run a number of simulations for both the homogeneous and heterogeneous setting. In chapter 4, we discuss the robustness of the model, in particular the assumptions on information, strategies and learning. The last chapter summarizes. All proofs and figures can be found in the appendix.

3.2 Cooperation in a population with homogeneous agents

3.2.1 Framework of the Model

We consider an infinitely repeated two-player normal form game which is played simultaneously by infinitely many pairs. Time is discrete and denoted by $t \in \{0, 1, 2, ...\}$. The population is a continuum of agents.

Every agent plays the stage game in each period with some opponent. An agent has the choice between the actions "cooperate" (C) and "do not cooperate" (D). Payoffs are given 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 stage game is a version of the prisoner's dilemma (PD) where the sum of payoffs is maximal at the action profile (C, C).

After the game has been played, each agent observes her opponent's action choice and has the option to maintain (M) or to quit (Q) the relationship with her current opponent. If both players choose the first option, they play the game together again in the next period. We call the link between these two players a "long-term relationship". If at least one player chooses Q, then both are not in a long-term relationship. Agents who are not in a long-term relationship at the end of period t, will be matched together randomly at the beginning of period $t+1$ before the stage game is played. The pool of agents who are not in a long-term relationship at the beginning of period t, will be called the "market" in this period. Thus, agents are in one of two "physical states": either "in the market" or "in a long-term relationship". We assume that agents are aware of their current physical state.

Attention will be restricted to the following kinds of behavior: an agent who follows strategy (Q, D) never cooperates and quits every relationship. We will call these agents "cheaters". Strategy (M, C) describes cooperative behavior: cooperate in every period and maintain the relationship if and only if the opponent cooperates as well. Hence, long-term relationships exist only between agents with strategy (M, C). In chapter 4, we discuss this assumption in detail.

Define

$$X = \{(Q, D), (M, C)\}. \tag{3.1}$$

Let $y_x(t)$, $x \in X$, be the share of agents following strategy x at the beginning of period t. Those who play strategy (M, C), are either part of a long-term relationship or in the market. Denote the respective shares with $y^m_{M,C}(t)$ for "matched" players and $y^u_{M,C}(t)$ for the "unmatched" ones. Thus, we have

$$y_{M,C}(t) = y^m_{M,C}(t) + y^u_{M,C}(t). \tag{3.2}$$

In contrast, agents who play strategy (Q, D) are always in the market. The size of the market, $y^u(t)$, is therefore given by

$$y^u(t) = y_{Q,D}(t) + y^u_{M,C}(t). \tag{3.3}$$

The probability of meeting a cooperative player in the market in period t is given by

$$s(t) = \frac{y^u_{M,C}(t)}{y^u(t)}. \tag{3.4}$$

Furthermore, denote the share of cooperators in long-term relationships relative to all cooperators in period t by

$$h(t) = \frac{y^m_{M,C}(t)}{y_{M,C}(t)}. \tag{3.5}$$

Let Δ_{HOM} be the two-dimensional simplex. Then, the distribution of physical states and strategies in period t is given by $Y(t) \in \Delta_{HOM}$ with

$$Y(t) = \begin{pmatrix} y^m_{M,C}(t) \\ y^u_{M,C}(t) \end{pmatrix}, \tag{3.6}$$

as we have $y_{Q,D}(t) = 1 - y_{M,C}(t)$.

Decisions of agents about the imitation of a certain strategy will be based upon the average payoffs of the respective strategies in the population. The average payoffs can be calculated as follows: the payoff structure from above implies that agents in a long-term relationship receive a payoff of G. The average utility in period t of those who play (Q, D) is given by

$$\bar{U}_{Q,D}(t) = s(t)H + 1 - s(t), \tag{3.7}$$

while for those who play (M, C) and who are in the market at the beginning of period t it is

$$\bar{U}^u_{M,C}(t) = s(t)G. \tag{3.8}$$

Thus, the average utility of the agents who play (M, C) is given by

$$\bar{U}_{M,C}(t) = h(t)G + (1 - h(t))s(t)G. \tag{3.9}$$

Denote by "order in period t" the ordering of the numbers $\bar{U}_{M,C}(t)$, $\bar{U}_{Q,D}(t)$ and G. Combined with the learning rule given below, the order characterizes which agents adapt to which strategy.

From (3.7) to (3.9) we see that the success of a certain strategy depends on the current distribution of physical states and strategies. With probability $\sigma > 0$, an agent receives information about what strategy is currently the most successful one and potentially adapts to it. We will call σ the "imitation rate". Define that strategy (Q, D) has rank 1 and strategy (M, C) has rank 2. Then, we can specify:

Learning Rule (L)

1. [Agent is in a long-term relationship] Switch to the strategy with the highest average payoff if and only if this payoff exceeds your payoff in the current relationship (if more than one strategy attains the highest average utility, then choose the strategy with the lowest rank). If this is not the case, then do not change anything.

2. [Agent is in the market] Switch to the strategy with the highest average payoff (if more than one strategy attains the highest average utility, then choose the strategy with the lowest rank).

An agent who changes her strategy reverses her decision to Q, i.e. she leaves the relationship regardless of what has happened before. This convention is made for simplicity and would not change the results (if an agent changes her strategy, she would loose her relationship anyway in the next period).

The asymmetry in **(L)** reflects the behavior in two different physical states: agents, who are not in a long-term relationship at the end of period t, are aware of the fact that in period $t+1$ they will play the game against another opponent in any case. Therefore, they switch to the strategy which works best in the population. However, an agent in a long-term relationship will lose the link to her opponent if she changes her behavior. Therefore, the average utility of the new strategy should exceed a certain level (which in our case is the payoff in the relationship).

In the next chapter, we will consider the same learning rule with an enlarged set of strategies. Thus, **(L)** is stated with more generality than needed in the current setup.

Additionally, players quit the current relationship and select their strategy randomly with small probability $\epsilon > 0$. Justifications for this can be

- idiosyncratic reasons of breakup (if the current relationship is long-term),
- experimentation or innovation,
- wrong action choice by mistake ("trembling hands") or
- a player dies and is replaced by a new agent who chooses her strategy randomly.

For simplicity, we treat these sources uniformly and call ϵ the "experimentation rate". We assume that agents, who select behavior randomly, choose each strategy with equal probability. Thus, there will be a share of players at the end of every period who lost their long-term relationship because they (or their opponent) have terminated the relationship due to experimentation. The sequence of play in every period is as follows:

(i) Agents in the market are paired up randomly (matching-phase).

(ii) Agents choose actions C or D simultaneously according to their strategy.

(iii) Payoffs are realized and agents observe the action choice of their opponent.

(iv) Each agent chooses Q or M according to her strategy and observes the respective choice of her opponent. If both choose M, then the link between the agents is a long-term relationship.

(v) With probability ϵ, an agent changes her strategy randomly and reverses her decision to Q (experimentation-phase).

(vi) If an agent has not changed her strategy in the experimentation-phase, she learns with probability σ according to (**L**). If she switches to a strategy different from the current one, she reverses her decision to Q (learning-phase).

Finally, we introduce the following definitions: we say that strategy $x \in X$ becomes the "norm" if this strategy is chosen by all learning players in almost all periods, i.e. from some point in time onwards no other strategy is chosen by learning agents any more. Furthermore, strategy $x \in X$ is said to be "temporarily attractive" if it does not become the norm, but is chosen by some learning players in infinitely many periods.

3.2.2 Analytical Results

We now show that in the above framework, a large share of the population may play the cooperative strategy after a number of periods, even though the initial share of those players, $y_{M,C}(0)$, is small or equal to zero.

The following trade-off has to be solved: cheaters benefit from the presence of many cooperative players in the market. Their average utility increases linearly in $s(t)$. On the other hand, cooperative players rely on the option to maintain a relationship, as they do strictly worse in the market than non-cooperating players. Due to experimentation, the share of these relationships in the population—and therefore the average utility of cooperative players—only grows if $s(t)$ is not too small over a larger number of successive periods. There are three possible orders:

Order A	$\bar{U}_{Q,D}(t) > G > \bar{U}_{M,C}(t)$
Order B	$G \geq \bar{U}_{Q,D}(t) \geq \bar{U}_{M,C}(t)$
Order C	$G \geq \bar{U}_{M,C}(t) > \bar{U}_{Q,D}(t)$

Assume first that $\epsilon = 0$, such that under the orders B and C no long-term relationships are dissolved. From (3.7) we can derive that if $s(t) \leq \frac{G-1}{H-1}$ holds, then $\bar{U}_{Q,D}(t) \leq G$ and the order in period t cannot be A. If σ is small, then $s(t)$ never exceeds $\frac{G-1}{H-1}$ in later periods. For any

$$Y(0) \in \Delta_{HOM} - \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{3.10}$$

we get adjustments similar to those plotted in figure (I): As long as non-cooperative strategies dominate, the share of cooperative individuals in the market decreases, either because cooperative agents find a proper partner, or because they switch to strategy (Q, D). The average utility of cooperating players converges to a value above 1, while for cheaters it decreases to 1. In some finite period t, the ratio $h(t)$ is sufficiently high and $s(t)$ sufficiently small, such that $\bar{U}_{M,C}(t) > \bar{U}_{Q,D}(t)$ and learning players choose (M, C): if σ is sufficiently small, the share of cooperating players in a long-term relationship increases in the following periods, while in the market the share of cooperating agents relative to the non-cooperating ones remains small. Thus, we get $\lim_{t\to\infty} y^m_{M,C}(t) = 1$.

For positive values of ϵ, we can show that adjustments are similar if σ and ϵ are sufficiently small. Then, cooperative behavior prevails in the population after a finite number of periods:

Theorem 1 [Cooperation in a homogeneous population]

If agents follow (**L**)*, then there is a* $\bar{\sigma} > 0$*, such that if* $\sigma \leq \bar{\sigma}$ *and* ϵ *is sufficiently small, strategy* (M, C) *becomes the norm.*

The transition function between two successive distributions of states and strategies, $Y(t)$ and $Y(t + 1)$, induced by (**L**), is neither continuous nor a contraction. Therefore, we cannot apply a standard fixed-point theorem. However, for each order the evolution of the distribution of physical states and strategies can be described by a system of non-linear difference equations. We do not solve these equations directly, but for given $Y(t)$ and $T \in \mathbb{N}$ we can estimate an upper and a lower bound on the elements in $Y(t+T)$, given that the order does not change in the interval $[t, t+T]$. In particular, we show that for σ and ϵ sufficiently small the following holds:

(i) The number of successive periods with order A or B is finite and the order does not switch from B to A. Thus, order C is reached within finite time.

(ii) The ratio s decreases until $s < \bar{s}$ in all remaining periods, where $\bar{s}$ is small enough such that the average utility of cheaters is well below G whenever $s < \bar{s}$.

(iii) After order C was reached for the first time, the number of periods between two points in time with order C, is finite.

(iv) The share of long-term relationships and the ratio h rise until the average utility of cooperating agents exceeds the average utility of cheaters in all future periods.

The upper bound on σ is needed for two reasons: firstly, it rules out that there are too many cooperating players in the market in later periods. Otherwise, the average utility of cheaters would be higher than G, which implies that also players in long-term relationships switch to (Q, D) and return to the market. For example, an imitation rate close to 1 may cause the strategies (Q, D) and (M, C) to replace each other as the most successful strategy from time to time and therefore, the claim of theorem 1 would no longer hold. Secondly, with a small imitation rate it is more likely for a cooperative player in the market that she finds another cooperative agent and does not adapt to (Q, D): if $y_{M,C}(0) = 0$, ϵ is very small and σ too high, then the ratio between cooperative long-term relationships and cooperators in the population, h, is always too small to make the average utility of strategy (M, C) exceed the one of cheaters. To illustrate this point, consider the following example:

Example. Set $\sigma = 1$, $y_{M,C}(0) = 0$ and $G < 4$. Assume that the order in all periods is B. Then, in each period $\tau \in \{1, 2, ...\}$ we have $y^u_{M,C}(\tau) = \frac{1}{2}\epsilon$. Furthermore, assume that in all periods $\tau \in \{1, 2, ...\}$ we also have

$$2y^m_{M,C}(\tau) \leq y^u_{M,C}(\tau), \tag{3.11}$$

which implies that

$$s(\tau) \leq \frac{\frac{1}{2}\epsilon}{1-\frac{1}{4}\epsilon}. \tag{3.12}$$

Under all orders it holds that

$$y^m_{M,C}(\tau+1) = (1-\epsilon)^2 y^m_{M,C}(\tau) + (1-\epsilon)^2 s(\tau) y^u_{M,C}(\tau). \tag{3.13}$$

With (3.13), it follows from $y^m_{M,C}(0) = 0$ that in all periods $\tau \in \{0, 1, ...\}$ we have

$$y^m_{M,C}(\tau) \leq \lim_{t\to\infty} y^m_{M,C}(t). \tag{3.14}$$

Thus, we can calculate that

$$\lim_{\epsilon\to 0}\lim_{t\to\infty} h(t) \leq \frac{1}{5}. \tag{3.15}$$

Therefore, the assumption in (3.11) is justified for sufficiently small ϵ. We get that the average utility of cooperative players in the market is close to 0 and overall average utility of cooperative players is always below 1 if ϵ is very small. The average utility for cheaters is always equal to or above 1, but well below G if ϵ is sufficiently small. Thus, the assumption on the order in all periods is justified as well and (M, C) will never prevail in the population. [**End of example**]

Note that theorem 1 is independent of the values H, G and of the distribution of physical states and strategies in the first period. For σ^* and ϵ^* small enough, the distribution $\lim_{t\to\infty} Y(t)$ is independent of $Y(0)$ as (M, C) becomes the norm. If order C prevails in t, then we have

$$y_{Q,D}(t+1) = (1-\epsilon^*)(1-\sigma^*) y_{Q,D}(t) + \frac{1}{2}\epsilon^*. \tag{3.16}$$

This yields us global convergence to the following distribution of strategies:

$$\lim_{t\to\infty} y_{Q,D}(t) = \frac{\frac{1}{2}\epsilon^*}{1-(1-\epsilon^*)(1-\sigma^*)}, \tag{3.17}$$

$$\lim_{t\to\infty} y_{M,C}(t) = 1 - \frac{\frac{1}{2}\epsilon^*}{1-(1-\epsilon^*)(1-\sigma^*)}. \tag{3.18}$$

For small values of σ and ϵ we therefore obtain a unique outcome in the repeated prisoner's dilemma with the option to maintain or to quit relationships.

3.2.3 The model without the option to maintain or to quit a relationship

The essence of theorem 1 is the option to maintain or to quit a relationship. Consider a version of the model where this option does not exist: agents can choose among the strategies (DD) "never cooperate" and (CC) "always cooperate". Let $y_{DD}(t)$ be the share of agents who play the first strategy in period t, and $1 - y_{DD}(t)$ the share of agents who play the second one. The respective average utilities therefore are

$$\bar{U}_{DD}(t) = (1 - y_{DD}(t))H + y_{DD}(t), \quad (3.19)$$

$$\bar{U}_{CC}(t) = (1 - y_{DD}(t))G. \quad (3.20)$$

Thus, in each period we have $\bar{U}_{DD}(t) > \bar{U}_{CC}(t)$, i.e. learning players always choose (DD). We therefore get for each pair of values ϵ^* and σ^* that

$$\lim_{t\to\infty} y_{DD}(t) = 1 - \frac{\frac{1}{2}\epsilon^*}{1 - (1 - \epsilon^*)(1 - \sigma^*)}, \quad (3.21)$$

$$\lim_{t\to\infty} y_{CC}(t) = \frac{\frac{1}{2}\epsilon^*}{1 - (1 - \epsilon^*)(1 - \sigma^*)}. \quad (3.22)$$

The share of cooperative players in the population can never exceed 50 percent and with $\sigma^* > 0$ and ϵ^* small, strategy (DD) prevails in the population. This is analogous to a standard result in evolutionary models, see Bergstrom (2002) for example: the only stable outcome with random matching and the PD as stage game, is a population consisting entirely of cheaters.

3.2.4 Simulation

Let us briefly look at absolute numbers to see what values of σ and ϵ support cooperation in the population. All graphs and detailed parameter values can be found in the appendix. For the following simulations, we fix $G = 2$ and $H = 3.8$. In the first period, 99 percent of the agents play strategy (Q, D). The upper graphs show the share of cooperating players in long-term relationships, the lower ones display the corresponding distribution of strategies in the market.

Figure (II) plots the adjustments for $\sigma = 0.02$ and $\epsilon = 0.01$ in the first 300 periods. We see that the share of cooperating players rises smoothly, as there are never sufficiently many cooperating players in the market to make strategy (Q, D) attractive for those who are in a long-term relationship. (M, C) also becomes the norm. However, in the market most players are cheaters, while in the overall population cooperative behavior dominates in later periods. Finally, we can check that these observations remain unchanged if we decrease ϵ even further.

Figure (III) shows the evolution if we increase σ to 0.05 and leave all other parameters—including $Y(0)$—unchanged: in infinitely many periods, the share of cooperative players in the market becomes sufficiently high to make strategy (Q, D) attractive for learning agents, except for those who are in a long-term relationship. (M, C) no longer becomes the norm, but remains temporarily attractive. Nevertheless, in later periods the share of cooperative long-term relationships in the population is even higher than in the first scenario. This is due to the increased speed of learning which leads to a higher share of cooperative agents in the market and thus a higher probability of meeting a cooperative player.

Finally, we set $\sigma = 0.15$ while again all other parameters remain the same. From figure (IV) we can observe that with these values, there are enough cooperating players in the market in some periods, such that the average utility of (Q, D) exceeds G. Then, even those players who are in a long-term relationship, switch to (Q, D). We can check that this holds for all positive values of ϵ and even bigger values of σ.

If there are almost no cooperative players in the population, it can take some time until (M, C) becomes attractive. In figure (V), the adjustments are displayed if we set $\sigma = 0.01$, $\epsilon = 0.1 \times 10^{-6}$ and the initial distribution of physical states and strategies is given by $y^u_{M,C}(0) = 0.1 \times 10^{-15}$, $y^u_{M,C}(0) = 0$. Note that learning agents adapt to (M, C) just after 200.000 periods.

3.3 Cooperation in a population with heterogeneous agents

3.3.1 Modification of the Framework

Whenever information about the average success of strategies is publicly available, we should expect that a large share of agents adapts to the most successful strategy in each period (in Eshel et. al. (1998), the reference case is $\sigma = 1$). In order to drop the restriction on the imitation rate of theorem 1, we now consider a population with heterogeneous players: an agent does not want to interact with any agent of the population, but prefers to meet individuals with certain characteristics or manners. Long-term relationships only can exist between agents who "get along with each other". This additional feature of relationships is independent of the strategies played by agents. We introduce the probability $s_T(t)$ that two agents who meet in the market, get along with each other. This probability must not be constant, but can take on values in the interval $[\underline{s}, \bar{s}]$, where

$$0 < \underline{s} \leq \bar{s} \leq 1. \tag{3.23}$$

With probability $1 - s_T(t)$, two agents who meet in the market, do not get along with each other and the relationship is broken up, regardless of the played strategies.

The difference to the previous setting is as follows: if both players of a given match get along with each other, then their utility increases by $Z \in \mathbb{R}_+$, regardless of the action choice. Hence, overall payoffs in this case are given by the following payoff matrix (where player 1 chooses rows and player 2 chooses columns):

	D	C
D	$1+Z, 1+Z$	$H+Z, Z$
C	$Z, H+Z$	$G+Z, G+Z$

If the two players do not get along with each other, their payoff is the same as in the last section, i.e. the payoff matrix is

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

Therefore, the value of a relationship consists of two parts: the utility generated by the opponent's action choice, and by the nature of their interaction. Both values are observed by an agent after the game has been played.

Now we consider the following strategies: an agent who follows (Q, D) never cooperates and quits every relationship. Strategy (M, D) prescribes never to cooperate, but to maintain a relationship if and only if she gets along with her opponent—regardless of her action choice. (M, C) now represents the following behavior: cooperate in every period and maintain the relationship if and only if the opponent cooperates as well and you get along with her. Now there are two types of long-term relationships: "cooperative" relationships, in which both agents play strategy (M, C), and "non-cooperative" ones, in which both agents play (M, D). In both cases, the agents have to get along with each other.

Note that in a homogeneous population there would be no scope for strategy (M, D): while cheating agents profit from the presence of cooperative players in the market in every period, this advantage of non-cooperative behavior ceases to exist when players maintain relationships with non-cooperative agents. Therefore, learning agents never would switch to strategy (M, D) in the homogeneous setting.

Denote for

$$X = \{(Q, D), (M, D), (M, C)\} \tag{3.24}$$

the share of agents playing strategy $x \in X$ in period t, by $y_x(t)$. Agents, who play a strategy $w \in W$ with

$$W = \{(M, D), (M, C)\}, \tag{3.25}$$

are either part of a long-term relationship or in the market. The respective shares are $y_w^m(t)$ and $y_w^u(t)$. Accordingly, we have

$$y_w(t) = y_w^m(t) + y_w^u(t). \tag{3.26}$$

The size of the market is then given by

$$y^u(t) = y_{Q,D}(t) + y_{M,D}^u(t) + y_{M,C}^u(t). \tag{3.27}$$

Define $s_{M,C}(t)$ $[s_{M,D}(t)]$ as the probability in period t that an agent with strategy (M, C) $[(M, D)]$ in the market is matched to a player with the same strategy and with whom she gets along, i.e.

$$s_{M,C}(t) = s_T(t)\frac{y_{M,C}^u(t)}{y^u(t)}, \tag{3.28}$$

$$s_{M,D}(t) = s_T(t)\frac{y_{M,D}^u(t)}{y^u(t)}. \tag{3.29}$$

Furthermore, we introduce the notation

$$\begin{aligned} h_{M,C}(t) &= \frac{y_{M,C}^m(t)}{y_{M,C}(t)}, \\ h_{M,D}(t) &= \frac{y_{M,D}^m(t)}{y_{M,D}(t)}. \end{aligned} \tag{3.30}$$

Let Δ_{HET} be the four-dimensional simplex. Then, the distribution of physical states and strategies in period t is given by $Y(t) \in \Delta_{HET}$ with

$$Y(t) = \begin{pmatrix} y^m_{M,C}(t) \\ y^u_{M,C}(t) \\ y^m_{M,D}(t) \\ y^u_{M,D}(t) \end{pmatrix}, \tag{3.31}$$

as we have $y_{Q,D}(t) = 1 - y_{M,C}(t) - y_{M,D}(t)$.

As in the case of a homogeneous population, we assume that learning agents imitate according to (**L**), where (Q, D) has rank 1, (M, D) has rank 2 and (M, C) has rank 3.

The average payoffs can be computed as follows: each partner in a cooperative long-term relationship receives utility of $G+Z$, while in a non-cooperative long-term relationship players receive $1+Z$. The average utility from strategy (Q, D) is given by

$$\begin{aligned} \bar{U}_{Q,D}(t) &= \frac{1}{y^u(t)}[y^u_{M,C}(t)s_T(t)(H+Z) + (y^u_{M,D}(t) + y_{Q,D}(t))s_T(t)(1+Z) + \\ &\quad + y^u_{M,C}(t)(1-s_T(t))H + (y^u_{M,D}(t) + y_{Q,D}(t))(1-s_T(t))], \end{aligned} \tag{3.32}$$

while for a player with strategy (M, C) in the market, it is

$$\begin{aligned} \bar{U}^u_{M,C}(t) &= \frac{1}{y^u(t)}[y^u_{M,C}(t)s_T(t)(G+Z) + (y^u_{M,D}(t) + y_{Q,D}(t))s_T(t)Z + \\ &\quad + y^u_{M,C}(t)(1-s_T(t))G]. \end{aligned} \tag{3.33}$$

This yields us the average payoffs for the remaining strategies:

$$\bar{U}_{M,C}(t) = h_{M,C}(t)(G+Z) + (1-h_{M,C}(t))\bar{U}^u_{M,C}(t), \tag{3.34}$$

$$\bar{U}_{M,D}(t) = h_{M,D}(t)(1+Z) + (1-h_{M,D}(t))\bar{U}_{i,Q,D}(t). \tag{3.35}$$

Again, denote by "order in period t" the ordering of the numbers $\bar{U}_{M,C}(t)$, $\bar{U}_{M,D}(t)$, $\bar{U}_{Q,D}(t)$, $G+Z$ and $1+Z$. Learning, experimentation and the sequence of events in each period are the same as in the last section.

3.3.2 Analytical Results

If the chance of meeting an opponent with whom one can get along, is small, and the importance of individual characteristics in social interaction—i.e. the payoff Z—is high, then there is an incentive to maintain a long-term relationship though it is non-cooperative. Thus, strategy (M, D) can be preferable to (Q, D), depending on $s_T(t) \in [\underline{s}, \bar{s}]$ and Z. The following assumption ensures that the maintenance of long-term relationships is strictly preferred to strategy (Q, D):

Assumption (A) *For G, Z and $\bar{s}$ it holds that $Z > \frac{2G-1}{1-\bar{s}}$.*

Using (3.32) and the fact $H < 2G$, one can the show that if (**A**) is fulfilled, we have

$$\bar{U}_{Q,D}(\tau) < \bar{U}_{M,D}(\tau) < 1 + Z \tag{3.36}$$

in all periods in which there are non-cooperative long-term relationships ($y^m_{M,D} > 0$), i.e. at least in the periods $\tau \in \{2, 3, ...\}$. By assuming (**A**), it is possible to show the second result:

Theorem 2 [Cooperation in a heterogeneous population]

Assume that agents follow (**L**) *and assumption* (**A**) *is fulfilled. If ϵ is sufficiently small, then strategy (M, C) becomes the norm.*

The proof works as follows: If learning players never adapt to strategy (Q, D), then there remain three orders for the periods $t \in \{2, 3, ...\}$:

Order A	$G + Z > 1 + Z \geq \bar{U}_{M,D}(t) \geq \bar{U}_{M,C}(t) > \bar{U}_{Q,D}(t)$
Order B	$G + Z > 1 + Z \geq \bar{U}_{M,C}(t) > \bar{U}_{M,D}(t) > \bar{U}_{Q,D}(t)$
Order C	$G + Z \geq \bar{U}_{M,C}(t) > 1 + Z \geq \bar{U}_{M,D}(t) > \bar{U}_{M,D}(t)$

If ϵ is sufficiently small, we can show:

(i) Given that the order in all periods is A or B, the number of periods between two points in time with order B, is bounded after finite time.

(ii) The statement in (i) implies that if the order is only A or B, the ratio $h_{M,C}$, and therefore the average utility of cooperative players increases until $\bar{U}_{M,C} > 1 + Z$, such that order C is reached after finitely many periods.

(iii) After order C has been reached the first time, the number of periods between two points in time with order C is bounded. Thus, $y^m_{M,C}$ rises until $\bar{U}_{M,C} > 1 + Z$ in all future periods. As the average utility of agents playing the strategies (Q, D) or (M, D) is always below $1 + Z$, strategy (M, C) becomes the norm.

The restriction on σ in theorem 1 can be dropped, as the average utility of players in the market is always below $1 + Z$, regardless of how many cooperators there are. However, we also found for the homogeneous population that the restriction on σ is necessary, because otherwise cooperators disappear from the market before they are matched to a cooperative agent and therefore the share of cooperative long-term relationships remains too small. In the heterogeneous setup this is no longer the case: as long as not all learning agents adapt to (M, C), i.e. the order is either A or B, the size of the market converges to a small value which in turn decreases in the experimentation rate. Thus, for any σ and any ϵ, the probability of meeting a cooperative player with whom one can get along, $s_{M,C}$, remains above some positive threshold after sufficiently many periods.

The assumption on Z, given in (**A**), replaces any restriction on σ. Smaller values of Z do not exclude that (M, C) becomes the norm as long as it is ruled out by a small σ—as in theorem 1—that the average utility of cheaters rises above $1 + Z$ in later periods, such that the size of the market remains small.

As mentioned in the introduction, heterogeneity of the population was used already in some models, in order to make it unattractive to return to the market after non-cooperative play. In an evolutionary setting, heterogeneity additionally gives rise to a small market which increases the probability that two agents who choose the cooperative strategy through experimentation, are matched together.

The result in theorem 2 is again independent of the payoffs H, G and of the distribution of physical states and strategies in the first period. If for fixed values of σ^* and ϵ^* the cooperative strategy becomes the norm, then $\lim_{t\to\infty} Y(t)$ does not depend on $Y(0)$. If order C prevails in period t, we have

$$y_{Q,D}(t+1) \quad = \quad (1-\epsilon^*)(1-\sigma^*)y_{Q,D}(t) + \frac{1}{3}\epsilon^*, \qquad (3.37)$$

$$y_{M,D}(t+1) \quad = \quad (1-\epsilon^*)(1-\sigma^*)y_{M,D}(t) + \frac{1}{3}\epsilon^*. \qquad (3.38)$$

We therefore get global convergence to the following distribution of strategies in the population:

$$\lim_{t\to\infty} y_{Q,D}(t) \quad = \quad \frac{\frac{1}{3}\epsilon^*}{1-(1-\epsilon^*)(1-\sigma^*)}, \qquad (3.39)$$

$$\lim_{t\to\infty} y_{M,D}(t) = \frac{\frac{1}{3}\epsilon^*}{1-(1-\epsilon^*)(1-\sigma^*)}, \quad (3.40)$$

$$\lim_{t\to\infty} y_{M,D}(t) = 1 - \frac{\frac{2}{3}\epsilon^*}{1-(1-\epsilon^*)(1-\sigma^*)}. \quad (3.41)$$

If the population is heterogeneous and ϵ is sufficiently small, we obtain a unique outcome in the repeated prisoner's dilemma with the option to maintain or to quit relationships.

3.3.3 Simulation

To illustrate the dynamics of the model, we again simulate it for several parameter values. For simplicity, we set $\underline{s} = \bar{s} = 0.15$ for the following but the last scenario. The payoffs are given by $H = 3.8$, $G = 2$ and $Z = 4$, such that assumption (**A**) is fulfilled. In the initial period, 99 percent of the players are cheaters.

Figure (VI) displays the adjustments if both the parameter values and the initial distribution of strategies is the same as in the third simulation of the previous chapter—with the exception that $\epsilon = 0.001$. Now we can observe that strategy (M, C) becomes the norm. In the first periods, strategy (M, D) works best due to the large share of non-cooperative players in the market. Thus, the share of non-cooperative long-term relationships rises. The share of cooperative relationships also rises, as cooperative players are matched to each other in a small market. In later periods, approximately 80 percent of the agents in the market behave cooperatively, while around 20 percent are non-cooperative. Thus, strategy (M, C) dominates even in the market which would of course not be possible in a homogenous population.

The next scenario shows that the evolution of strategies in the population is quite sensitive to changes in the experimentation rate. Figure (VII) displays the case when we set $\epsilon = 0.002$ and all other parameters remain unchanged compared to the scenario in figure (VI). The ratio $h_{M,C}$ remains low, as too many relationships are broken up in each period. Thus, strategy (M, C) never yields a higher average utility than (M, D), as (M, D) performs better in the market. (M, D) becomes the norm and most long-term relationships are non-cooperative. We see that a small change in the behavior of players may have a huge impact on the organization of the society.

In figure (VIII) we are confronted with the same situation when the imitation rate is equal to $\sigma = 0.99$ and $\epsilon = 0.001$. Thus, most agents imitate in every period. Strategy (M, D) prevails for very long time, as the share of cooperators in the market remains small, due to the high imitation rate. Nevertheless, (M, C) does not only become the norm, but also dominates in the market in later periods.

The last simulation, displayed in figure (IX), shows what happens if there is too little heterogeneity in the population: we set $\underline{s} = \bar{s} = 0.70$. Then we have a situation similar to the one presented in figure (III), i.e. in a homogeneous population when the imitation rate is too high. As for an agent the probability of meeting an opponent in the market with whom she gets along, is high, strategy (Q, D) remains temporarily attractive and (M, C) cannot become the norm. This observation does not change if we decrease the experimentation rate even further.

3.4 Discussion of assumptions and robustness

In the setup of the model we made a number of assumptions on the informational structure, the learning rule and the set of strategies. This section is devoted to their justification.

3.4.1 Information about average payoffs

We assumed that learning agents have global information about what strategy is doing best in the moment. This is plausible for a world with mass communication where a "public opinion" results from this process. For example, if some companies have success with a certain employment policy, they will be imitated quickly by other firms.

If we change the setting, such that agents base their decision upon a limited number of observations, then we have to introduce a so-called "reference network" for each agent, as in Cartwright (2003, 2007). We would have to make assumptions about whether and how an agent's reference network changes over time. Also, we should introduce the possibility that agents can choose to be matched with observed players directly, or justify why this is not possible. Thus, we end up in a spatial model—something we wanted to avoid in the present approach: the goal was to show that cooperative behavior can prevail if we have global interaction and non-assortative matching.

Apart of that, the basic trade-offs should remain unchanged: cooperative behavior can prevail if there are never too many cooperative players in the market or if there is sufficient heterogeneity in the market. However, it is not for sure that we obtain global convergence in such a setting.

The assumption that players base their decisions on the average success of strategies was also used by Robson et. al. (1996) for a finite, but arbitrary large set of agents. In Eshel et. al. (1998), agents observe the success of their neighbors and therefore have complete information about strategy and utility of players with whom they interact.

A further critique with respect to the informational assumption was stated by Henrich (2004): the costs of gathering and distributing the information about average payoffs, is not accounted for. We argue that these costs are negligible relative to the payoff from social interaction and agents derive utility from purchasing the information from time to time.

3.4.2 The learning rule

Imitation is well-established in both the theoretical—see Schlag (1998), Selten and Ostmann (2001)—and experimental literature—see Apesteguia et. al. (2007) and Offerman et. al. (2002)—as a form of learning in the presence of bounded rationality and imperfect information. The considered learning rule (**L**) is called "imitate the best average"—see, for example, Ellison and Fudenberg (1995) or Schlag (1999). It satisfies a number of plausible properties which were used in Apesteguia et. al. (2007), to evaluate the plausibility of imitation rules in experiments:

(i) If all strategies are distinct, the more successful strategies are imitated with higher probability.

(ii) Never switch to a strategy with an average payoff lower than the average payoff of the own action.

(iii) Imitate the action with the highest average payoff with strictly positive probability (unless the current match allows for an even higher period payoff).

(iv) Never switch to an action with average payoff below the average payoff in the population.

Property (iii) is slightly modified in order to take care of the fact that agents in long-term relationships may have a high payoff, although the average utility of their strategy is below the average utility of another strategy.

There are other accepted forms of naive learning: the most common ones are "fictitious play"—see Fudenberg et. al. (1998) or Young (1998)—"reinforcement learning", introduced by Roth and Erev (1995) and "regret matching" which is due to Hart and Mas-Colell (2000). These learning rules do not solve the social dilemma in the considered game without further restrictions: fictitious play prescribes to play a best response against past average behavior of agents. As cooperation is strictly dominated, the outcome always would be non-cooperative as long as agents are myopic.

The remaining two types of naive learning specify to play an action with a probability proportional to its (potential) success in past periods in terms of payoffs. This randomization among different actions is not consistent with the idea of a long-term relationship. Additionally, as long as a large share of the population is not in a long-term relationship, strategy (M, C) will do strictly worse than (Q, D) for most agents who try out this strategy and for quite a number of periods: cooperative behavior will not emerge if purely self-interested and myopic agents stick to these learning rules in a setting of global interaction and non-assortative matching.

Therefore, a learning rule which leads to cooperative behavior has to be based on imitation at least partially: not only individual experiences are informative, but also those of other players. The opportunities to generalize learning rule (**L**), are limited.

3.4.3 Strategies

In both the homogeneous and the heterogeneous setting, we restricted the set of strategies. However, we considered all strategies which fulfill the following criteria:

- The strategy is not history dependent.
- The respective long-term relationships are symmetric.
- The strategy is not strictly dominated in terms of average payoff by another strategy for all distributions of physical states and strategies and in the presence of all physical states and strategies (this is ensured by the positive experimentation rate).

The first criterion rules out that a cooperative long-term relationship starts with non-cooperation. As mentioned in the introduction, both agents then can improve the outcome by starting with cooperation immediately.

Our results are robust with respect to the inclusion of strictly dominated strategies. Consider first the homogeneous setting in which the remaining two strategies (M, D), "defect in every period and maintain the relationship", and (Q, C), "cooperate in every period and quit the relationship", are strictly dominated. Thus, they are never chosen by learning players as long as (M, C) and (Q, D) are present in the population. By the methods used in the proof of theorem 1, one can show that for sufficiently small values of σ and ϵ, the share of cooperative players in the market relative to the share of non-cooperative players remains small after finitely many periods, and that the ratio h increases until strategy (M, C) dominates (Q, D).

In the heterogeneous setting, only strategy (Q, C) is strictly dominated for all values of s_T. As long as assumption (**A**) is fulfilled, the adjustments would not change if we include (Q, C) into the setting: (M, D) strictly dominates (Q, D) even if there are only cooperative agents in the market. Therefore, the size of the market converges to a small value and the share of cooperative relationships rises until (M, C) is chosen by all learning players in all periods.

3.5 Concluding remarks

In this paper, we provided an evolutionary solution for the repeated prisoner's dilemma with the option to maintain or to quit relationships: We considered boundedly-rational agents who imitate the strategy which is on average the most successful one. Strategies are simple and independent of the history of past interactions. In particular, there is no phase of small payoffs at the beginning of a long-term relationship which is difficult to justify if each agent faces the same trade-off. The evolutionary approach also has the advantage that knowledge of aggregate play of agents is not required.

We showed that with the option to maintain or to quit relationships, cooperative behavior emerges in a large population with non-assortative matching and global interaction. For a homogeneous population, we have to place a severe restriction on the imitation and experimentation rate. With heterogeneous agents, this restriction becomes disposable if the utility from meeting opponents with whom one can get along, is sufficiently large, and the probability of meeting such an opponent, is sufficiently small. We intensely discussed the dynamics in the market for long-term relationships. The remarkable feature of the model is global convergence to a unique outcome in a game which exhibits a large number of equilibria.

Though cooperation may prevail in the long run, agents always make mixed experiences with cooperative behavior: In the market, they possibly will suffer from the interaction with cheaters and players they do not want to be matched with. In a long-term relationship, they benefit from cooperation and the repeated interaction with someone they accept as partner. Further research may concentrate on the combination of imitation and learning by experience in this or related games. A combination of several learning forms—similar to the EWA model of Camerer and Ho (1998, 1999)—would contribute to the generality of the model and lead to new insights.

3.6 Appendix

Proof of Theorem 1

Denote by Δ^A those elements of Δ_{HOM} for which the corresponding order is A and accordingly, Δ^B and Δ^C. Obviously, we have

$$\Delta^A \cup \Delta^B \cup \Delta^C = \Delta_{HOM}, \tag{3.42}$$

$$\Delta^j \cap \Delta^l = \emptyset, \tag{3.43}$$

for $l, j \in \{A, B, C\}$ and $l \neq j$. After some transformations of (3.7), one can see that $\bar{U}_{Q,D}(t) \leq G$ if and only if

$$s(t) \leq \frac{G-1}{H-1}. \tag{3.44}$$

If (3.44) holds, then $Y(t) \notin \Delta^A$. As $G > 1$, we can find a $\lambda > 0$, such that

$$\frac{1}{1+\lambda}G - 1 > 0. \tag{3.45}$$

Then, from (3.7) we can derive that $\bar{U}_{Q,D}(\tau) < 1 + \lambda$ if

$$s(\tau) < \frac{\lambda}{H-1}, \tag{3.46}$$

and from (3.9) we see that $\bar{U}_{M,C}(\tau) > 1 + \lambda$ if

$$\frac{y^u_{M,C}(\tau)}{y^m_{M,C}(\tau)} < \frac{1}{1+\lambda}G - 1. \tag{3.47}$$

Choose $\bar{s} > 0$ such that

$$\bar{s} < \min\left\{\frac{\lambda}{H-1}, \frac{G-1}{H-1}, \frac{1}{4}\right\}, \tag{3.48}$$

and $\bar{y} < 1$ such that

$$\frac{\bar{s}(1-\bar{y})}{\bar{y}} < \frac{1}{1+\lambda}G - 1. \tag{3.49}$$

Regardless of the order in period t, it holds that

$$y_{Q,D}(t+1) \geq (1-\epsilon)(1-\sigma)y_{Q,D}(t) + \frac{1}{2}\epsilon. \tag{3.50}$$

Thus, for each $\epsilon > 0$ there is a finite period t^*_ϵ, such that in all periods $\tau \in \{t^*_\epsilon, t^*_\epsilon + 1, ...\}$ we have

$$y_{Q,D}(\tau) \geq \frac{\frac{1}{3}\epsilon}{1-(1-\epsilon)(1-\sigma)}. \tag{3.51}$$

In the following, we consider adjustments after period t^*_ϵ.

Lemma 1.1 *Assume that $Y(t) \in \Delta^A$. Then, for each $\sigma > 0$ there is a finite period $t+T$ with $Y(t+T) \in \Delta^B \cup \Delta^C$ if ϵ is sufficiently small.*

Proof: Assume that the claim does not hold and $Y(\tau) \in \Delta^A$ all periods $\tau \in \{t, t+1, ...\}$. Then it holds that

$$y_{M,C}(\tau+1) = (1-\epsilon)(1-\sigma)y_{M,C}(\tau) + \frac{1}{2}\epsilon. \tag{3.52}$$

Hence, we have

$$\lim_{t\to\infty} y_{M,C}(t) = \frac{\frac{1}{2}\epsilon}{1-(1-\epsilon)(1-\sigma)}, \tag{3.53}$$

$$\lim_{t\to\infty} y_{Q,D}(t) = 1 - \frac{\frac{1}{2}\epsilon}{1-(1-\epsilon)(1-\sigma)}. \tag{3.54}$$

It follows that

$$\lim_{\epsilon\to 0}\lim_{t\to\infty} s(t) = 0. \tag{3.55}$$

Thus, there is a finite period $t+T$ with $s(t+T) \leq \frac{G-1}{H-1}$ and therefore $Y(t+T) \in \Delta^B \cup \Delta^C$ if ϵ is sufficiently small.
Q.E.D.

Lemma 1.2 *Suppose that $t \geq t^*_\epsilon$. Then there is a $\hat{\sigma} > 0$, such that from $\sigma \in (0, \hat{\sigma}]$ and $Y(t) \in \Delta^B \cup \Delta^C$ it follows that $Y(t+1) \notin \Delta^A$, and there is a finite period t^{**}_ϵ for which it holds that in all periods $\tau \geq t^{**}_\epsilon$ we have $s(\tau) \leq \bar{s}$ if ϵ is sufficiently small.*

Proof: If $Y(t) \in \Delta^B$, then we have

$$s(t+1) = \frac{(1-\sigma)(1-s(t))y^u_{M,C}(t) + (1-\sigma)\epsilon y^m_{M,C}(t) + \frac{1}{2}\frac{\epsilon}{1-\epsilon}}{(1-s(t))y^u_{M,C}(t) + y_{Q,D}(t) + \epsilon y^m_{M,C}(t) + \frac{\epsilon}{1-\epsilon}}, \tag{3.56}$$

while if $Y(t) \in \Delta^C$ we have

$$s(t+1) = \frac{(1-s(t))y^u_{M,C}(t) + \sigma y_{Q,D}(t) + \epsilon y^m_{M,C}(t) + \frac{1}{2}\frac{\epsilon}{1-\epsilon}}{(1-s(t))y^u_{M,C}(t) + y_{Q,D}(t) + \epsilon y^m_{M,C}(t) + \frac{\epsilon}{1-\epsilon}}. \tag{3.57}$$

Note that under $Y(t) \in \Delta^C$ the value $s(t+1)$ is larger than under $Y(t) \in \Delta^B$. By simplifying the right-hand side of (3.57) we get

$$s(t+1) < \frac{(1-s(t))(s(t)+\sigma) + \frac{1}{y^u(t)}\left(\epsilon + \frac{1}{2}\frac{\epsilon}{1-\epsilon}\right)}{1-s(t)^2 + \frac{1}{y^u(t)}\left(\epsilon + \frac{\epsilon}{1-\epsilon}\right)}. \tag{3.58}$$

As $t \geq t^*_\epsilon$, we have

$$y^u(t) > y_{Q,D}(t) \geq \frac{\frac{1}{3}\epsilon}{1-(1-\epsilon)(1-\sigma)}. \tag{3.59}$$

With (3.58) and (3.59), we get

$$\lim_{\epsilon\to 0} s(t+1) < \frac{(1-s(t))(s(t)+\sigma) + \frac{9}{2}\sigma}{1-s(t)^2 + 6\sigma} \equiv ub(s(t)). \tag{3.60}$$

Then, we can estimate

$$\lim_{\epsilon \to 0} [s(t) - s(t+1)] > \frac{s(t)^2 - s(t)^3 - \frac{11}{2}\sigma}{1 - s(t)^2 + 6\sigma}. \tag{3.61}$$

Note that the term on the right-hand side of (3.61) strictly decreases in σ. Thus, we can set $\hat{\sigma} < \frac{1}{2}$ such that for all $\sigma \leq \hat{\sigma}$ from $s(t) \in \left[\bar{s}, \frac{G-1}{H-1}\right]$ it follows that $s(t+1) + d < s(t)$ for a small $d > 0$. From the first derivative of $ub(s(t))$, we see that

$$s(t) > \frac{1}{2(1-\sigma)} \tag{3.62}$$

implies

$$\frac{d\ ub(s(t))}{d\ s(t)} > 0. \tag{3.63}$$

Recall from (3.48) that $\bar{s} < \frac{1}{4}$. Thus, if $s(t) < \bar{s}$, then

$$\lim_{\epsilon \to 0} s(t+1) < ub(s(t)) < ub(\bar{s}) < \bar{s}. \tag{3.64}$$

Now fix any $\sigma \in (0, \hat{\sigma}]$. As ϵ enters (3.58) continuously, there is a $\epsilon' > 0$, such that for all ϵ with $0 < \epsilon < \epsilon'$ the following holds: from $s(t) \in \left[\bar{s}, \frac{G-1}{H-1}\right]$ it follows that $s(t+1) + \frac{d}{2} < s(t)$ and from $s(t) \in [0, \bar{s})$ we get $s(t+1) < \bar{s}$. Thus, both claims follow.
Q.E.D.

Fix a $\hat{\sigma}$ such that for $\bar{s}$ the claim of lemma 1.2 holds. In the following, we consider adjustments after period t^{**}_ϵ. Furthermore, fix a $\sigma^* \in (0, \hat{\sigma}]$ and a $\hat{s} > 0$ such that

$$12\left[\sigma^* + 2\hat{s} + \frac{\hat{s}^2}{\sigma^*}\right] < \frac{1}{1+\lambda}G - 1. \tag{3.65}$$

Further, fix an $\alpha > 0$ with $\alpha < (1-\alpha)\hat{s}^2$ and $\alpha < (\sigma^*)^2$. Finally, choose $k \in \mathbb{N}$ such that

$$(1-\sigma^*)^k \frac{1-\alpha}{\alpha} < \frac{1}{1+\lambda}G - 1, \tag{3.66}$$

$$(1-\sigma^*)^k \frac{G-1}{H-G} < \frac{\lambda}{H-\lambda-1}. \tag{3.67}$$

Lemma 1.3a *Assume that $Y(t) \in \Delta^B$, $\sigma = \sigma^*$, $y^m_{M,C}(t) > \alpha$ and $t \geq t^{**}_\epsilon$. If ϵ is sufficiently small, then $Y(\tau) \in \Delta^C$ for $\tau \leq t + k$.*

Proof: We can estimate

$$y^m_{M,C}(t+k) > (1-\epsilon)^{2k} y^m_{M,C}(t), \tag{3.68}$$

$$y^u_{M,C}(t+k) < (1-\epsilon)^k (1-\sigma^*)^k y^u_{M,C}(t) + k\left(\frac{3}{2}\epsilon - \epsilon^2\right), \tag{3.69}$$

$$y_{Q,D}(t+k) > (1-\epsilon)^k y_{Q,D}(t). \tag{3.70}$$

By assumption we have

$$\frac{y^u_{M,C}(t)}{y^m_{M,C}(t)} < \frac{1-\alpha}{\alpha} \tag{3.71}$$

and from $Y(t) \in \Delta^B$ it follows that

$$s(t) \leq \frac{G-1}{H-1}. \tag{3.72}$$

This can be transformed to

$$\frac{y^u_{M,C}(t)}{y_{Q,D}(t)} \leq \frac{G-1}{H-G}. \tag{3.73}$$

Thus, we get

$$\lim_{\epsilon\to 0}\frac{y^u_{M,C}(t+k)}{y^m_{M,C}(t+T)} < (1-\sigma^*)^k\frac{(1-\alpha)}{\alpha} < \frac{1}{1+\lambda}G-1, \tag{3.74}$$

$$\lim_{\epsilon\to 0}\frac{y^u_{M,C}(t+k)}{y_{Q,D}(t+T)} < (1-\sigma^*)^k\frac{G-1}{H-G} < \frac{\lambda}{H-\lambda-1}, \tag{3.75}$$

which implies that if ϵ is sufficiently small, then (3.46) and (3.47) are fulfilled in a period $\tau \leq t+k$ and $Y(\tau)\in\Delta^C$.
Q.E.D.

Lemma 1.3b *Assume $Y(t)\in\Delta^B$, $\sigma=\sigma^*$, $y^m_{M,C}(t)\leq\alpha$ and $t\geq t^{**}_\epsilon$. If ϵ is sufficiently small, then there is a period $t+T^{**}_\epsilon$, such that $Y(\tau)\in\Delta^C$ for $\tau\geq t+T^{**}_\epsilon$.*

Proof: From Lemma 1.2 we know that in all periods $\tau\geq t^{**}_\epsilon$, we have $s(\tau)\leq\bar{s}$. Assume that $y^m_{M,C}(\tau)\leq\alpha$ and $Y(\tau)\in\Delta^B$ in all periods $\tau\in\{t,t+1,...\}$. Then, it holds in all periods $\tau\in\{t,t+1,...\}$ that

$$y^u_{M,C}(\tau+1)=(1-\epsilon)(1-\sigma^*)(1-s(\tau))y^u_{M,C}(\tau)+\epsilon(1-\epsilon)y^m_{M,C}(\tau)+\frac{1}{2}\epsilon. \tag{3.76}$$

It follows that

$$y^u_{M,C}(\tau+1)<(1-\epsilon)(1-\sigma^*)y^u_{M,C}(\tau)+\frac{3}{2}\epsilon-\epsilon^2. \tag{3.77}$$

Thus, for each $\epsilon>0$ there is a finite period $t+T_\epsilon$, such that in all periods $\tau\geq t+T_\epsilon$, we have

$$y^u_{M,C}(\tau)<\frac{\frac{3}{2}\epsilon-\epsilon^2}{1-(1-\epsilon)(1-\sigma^*)} \tag{3.78}$$

and $\bar{U}_{Q,D}(\tau)<1+\lambda$, as $s(\tau)\leq\bar{s}$.

Assume further that in all periods $\tau\in\{t,t+1,...\}$ it holds that $s(t)<\hat{s}$. Using (3.76), we get for each $\epsilon>0$ that there is a finite period $t+T'_\epsilon$, such that in all periods $\tau\geq t+T'_\epsilon$ we have

$$y^u_{M,C}(\tau)>\frac{\frac{1}{2}\epsilon}{1-(1-\epsilon)(1-\sigma^*)(1-\hat{s})}. \tag{3.79}$$

This implies that in all periods $\tau\geq t+T'_\epsilon$, we also have

$$s(\tau)>\frac{\frac{1}{2}\epsilon}{1-(1-\epsilon)(1-\sigma^*)(1-\hat{s})}. \tag{3.80}$$

From (3.79), (3.80) and

$$y^m_{M,C}(\tau+1)=(1-\epsilon)^2y^m_{M,C}(\tau)+s(\tau)(1-\epsilon)^2y^u_{M,C}(\tau), \tag{3.81}$$

it follows that for each $\epsilon>0$, there is a finite period $t+T''_\epsilon$, such that in all periods $\tau\geq t+T''_\epsilon$ we have

$$y^m_{M,C}(\tau)>\frac{(1-\epsilon)^2}{1-(1-\epsilon)^2}\frac{\frac{1}{4}\epsilon^2}{[1-(1-\epsilon)(1-\sigma^*)(1-\bar{s})]^2}. \tag{3.82}$$

Fix $T^{**}_\epsilon=\max\{T_\epsilon,T'_\epsilon,T''_\epsilon\}$. Using (3.78) and (3.82), we therefore get

$$\lim_{\epsilon\to 0}\frac{y^u_{M,C}(t+T^{**}_\epsilon)}{y^m_{M,C}(t+T^{**}_\epsilon)}<12\left[\sigma^*+2\hat{s}+\frac{\hat{s}^2}{\sigma^*}\right]. \tag{3.83}$$

From (3.65) it follows that in some period $\tau \leq t + T_\epsilon^{**}$ the conditions (3.46) and (3.47) are fulfilled if ϵ is sufficiently small and therefore $Y(\tau) \in \Delta^C$.

Assume now that in at least one period $t' \in \{t, t+1, ...\}$ it holds that $s(t') > \hat{s}$. Then, we have

$$y^u_{M,C}(t') > \hat{s}(1-\alpha) \tag{3.84}$$

and

$$y^m_{M,C}(t'+1) > \hat{s}^2(1-\epsilon)^2(1-\alpha), \tag{3.85}$$

where by the specification of α, it holds that the right-hand side of (3.85) exceeds α if ϵ is sufficiently small. We have $Y(t'+1) \in \Delta^B \cup \Delta^C$. If $Y(t'+1) \in \Delta^B$, then the result follows from lemma 1.3a.
Q.E.D.

Lemma 1.3c *Assume that $t \geq t_\epsilon^{**}$ and $\sigma = \sigma^*$. From $Y(t) \in \Delta^C$ it follows that $Y(\tau) \in \Delta^C$ with $\tau \leq t + k + 1$ if ϵ is sufficiently small.*

Proof: If $Y(t) \in \Delta^C$ we can estimate

$$\begin{aligned} y^u_{M,C}(t+1) &> (1-\epsilon)\sigma^*(1-y^m_{M,C}(t)), &(3.86)\\ s(t+1) &> (1-\epsilon)\sigma^*(1-y^m_{M,C}(t)), &(3.87)\end{aligned}$$

and therefore

$$y^m_{M,C}(t+2) = (1-\epsilon)^2 y^m_{M,C}(t+1) + (1-\epsilon)^2 s(t+1) y^u_{M,C}(t+1) > (1-\epsilon)^4(\sigma^*)^2. \tag{3.88}$$

We have $Y(t+2) \in \Delta^B \cup \Delta^C$. If $Y(t+2) \in \Delta^B$, then, by the specification of α, the result follows from lemma 1.3a.
Q.E.D.

Lemma 1.4 *Assume $Y(t) \in \Delta^C$, $\sigma = \sigma^*$ and $t \geq t_\epsilon^{**}$. If ϵ is sufficiently small, then there is a finite period $t + T^{***}$, such that $y^m_{M,C}(t + T^{***}) > \bar{y}$.*

Proof: Consider the subsequence $\{t^p\}_{p\in\mathbb{N}}$ of $\{t, t+1, ...\}$ with $t^0 = t$ and $Y(t^p) \in \Delta^C$ for all $p \in \mathbb{N}$. From lemma 1.3c we know that $t^{p+1} \leq t^p + k + 1$. We then can estimate

$$y^m_{M,C}(t^{p+1}+2) > (1-\epsilon)^{2(k+1)} y^m_{M,C}(t^p+2) + (1-\epsilon)^{2(k+3)}(\sigma^*)^2(1-y^m_{M,C}(t^p))^2.$$

If ϵ is sufficiently small, then

$$(1-\epsilon)^{2(k+1)} - 1 + (1-\epsilon)^{2(k+3)}(\sigma^*)^2(1-\bar{y})^2 > d \tag{3.89}$$

for a small $d > 0$. The inequality in (3.89) ensures that for some $\bar{p} < \infty$ the finite sequence $\{y^m_{M,C}(t^p + 2)\}_{p\in\{0,1,...,\bar{p}\}}$ increases monotonically. The claim holds for some $t + T^{***} \leq t^{\bar{p}} + k + 3$, as $\bar{p}$ can be chosen sufficiently large.
Q.E.D.

With the results in the lemmas 1.1 to 1.4, it follows that there is a finite period $\bar{t}_\epsilon$, such that $s(\tau) \leq \bar{s}$ for all periods $\tau \geq \bar{t}_\epsilon$ and $y^m_{M,C}(\bar{t}_\epsilon) \geq \bar{y}$. Then, from the specifications in (3.46) to (3.49), it follows that $Y(\bar{t}_\epsilon) \in \Delta^C$ and

$$y^m_{M,C}(\bar{t}_\epsilon + 2) > (1-\epsilon)^4 y^m_{M,C}(\bar{t}_\epsilon) + (1-\epsilon)^4(\sigma^*)^2(1-y^m_{M,C}(\bar{t}_\epsilon)). \tag{3.90}$$

Thus, if ϵ is sufficiently small, it holds that in all periods $\tau \geq \bar{t}_\epsilon$ we have $y^m_{M,C}(\tau) \geq \bar{y}$ and therefore $Y(\tau) \in \Delta^C$. The same obviously holds for all $\sigma^* \in (0, \hat{\sigma}]$ which proofs the result.
Q.E.D.

Proof of Theorem 2

Fix $\sigma > 0$, $\epsilon > 0$ and the interval $[\underline{s}, \bar{s}]$ according to (3.23). Assume that (**A**) holds. We first make the following observations:

(1) As $s_T(t) \in [\underline{s}, \bar{s}]$ for all periods t, from (3.32) it follows that

$$\bar{U}_{Q,D}(t) < \bar{s}(H+Z) + (1-\bar{s})(H) < 1 + Z, \tag{3.91}$$

where the last inequality is implied by the assumptions on H and Z. Thus, we have in all periods $t \in \{2, 3, ...\}$

$$\bar{U}_{Q,D}(t) < 1 + Z, \tag{3.92}$$

$$\bar{U}_{Q,D}(t) < \bar{U}_{M,D}(\tau), \tag{3.93}$$

such that in all these periods learning players never switch to strategy (Q, D).

(2) From (3.34) it follows that

$$h_{M,C}(t) > \frac{1+Z}{G+Z} \tag{3.94}$$

implies $\bar{U}_{M,C}(t) > 1 + Z$. From observation **(1)**, we get that in this period τ all learning players with the strategies (Q, D) and (M, D) switch to (M, C). Fix a $\bar{y}$ with $1 > \bar{y} > \frac{1+Z}{G+Z} + d$ for a small $d > 0$. Then, $y^m_{M,C}(\tau) \geq \bar{y}$ implies that in period τ all learning players with the strategies (Q, D) and (M, D) switch to (M, C).

(3) For $t \in \{2, 3, ...\}$ the following orders can occur:

Order A	$G+Z > 1+Z \geq \bar{U}_{M,D}(t) \geq \bar{U}_{M,C}(t) > \bar{U}_{Q,D}(t)$
Order B	$G+Z > 1+Z \geq \bar{U}_{M,C}(t) > \bar{U}_{M,D}(t) > \bar{U}_{Q,D}(t)$
Order C	$G+Z \geq \bar{U}_{M,C}(t) > 1+Z \geq \bar{U}_{M,D}(t) > \bar{U}_{M,D}(t)$

Denote by Δ^A those elements of Δ_{HET}, for which the corresponding order is A for all $s_T \in [\underline{s}, \bar{s}]$, and accordingly Δ^B and Δ^C. Again, we have $\Delta^j \cap \Delta^l = \emptyset$ for $l, j \in \{A, B, C\}$ and $l \neq j$.

Set

$$s^* = \frac{1}{24} \frac{\underline{s}}{\left(1 + \frac{2-2(1-\bar{s})(1-\sigma)}{\sigma}\right)^2 (1-(1-\bar{s})(1-\sigma))}. \tag{3.95}$$

Choose $\bar{k} \in \mathbb{N}$ such that

$$\frac{1}{4}\bar{k}s^*(1-(1-\sigma)(1-s^*)) > \frac{1+Z}{G-1}, \tag{3.96}$$

$$\frac{s^*}{(1-\sigma)^{\bar{k}}(1-s^*)^{\bar{k}}} > \frac{1+Z}{G-1}, \tag{3.97}$$

$$\frac{\underline{s}\sigma^2}{\underline{s}\sigma^2 + (1-\sigma)^{\bar{k}-1}(1-\underline{s}\sigma^2)} > \frac{1+Z}{G+Z}. \tag{3.98}$$

Lemma 2.1 *Assume that $Y(\tau) \in \Delta^A \cup \Delta^B$ in all periods $\tau \in \{t, t+1, ...\}$. If ϵ is sufficiently small, then there is a finite period $t + T^*_\epsilon$, such that for all $\tau \geq t + T^*_\epsilon$ from $Y(\tau) \in \Delta^A$ it follows $Y(\tau + k) \in \Delta^B$ for some $k \leq \bar{k}$.*

Proof: As $Y(\tau) \in \Delta^A \cup \Delta^B$, for all $\tau \in \{t, t+1, ...\}$ we can estimate

$$y^u_{M,C}(\tau+1) > (1-\epsilon)(1-\sigma)(1-s_{M,C}(\tau))y^u_{M,C}(\tau) + \frac{1}{3}\epsilon, \tag{3.99}$$

$$y^u_{M,D}(\tau+1) > (1-\epsilon)(1-\sigma)(1-s_{M,D}(\tau))y^u_{M,D}(\tau) + \frac{1}{3}\epsilon, \tag{3.100}$$

$$y_{Q,D}(\tau+1) = (1-\epsilon)(1-\sigma)y_{Q,D}(\tau) + \frac{1}{3}\epsilon. \tag{3.101}$$

From these equations it follows that for any $\epsilon > 0$, there is a finite period $t + T_\epsilon$, such that in all periods $\tau \geq t + T_\epsilon$, we have

$$
\begin{aligned}
y^u_{M,C}(\tau) &> \frac{\frac{1}{6}\epsilon}{1-(1-\bar{s})(1-\epsilon)(1-\sigma)}, & (3.102)\\
y^u_{M,D}(\tau) &> \frac{\frac{1}{6}\epsilon}{1-(1-\bar{s})(1-\epsilon)(1-\sigma)}, & (3.103)\\
y_{Q,D}(\tau) &< \frac{\frac{2}{3}\epsilon}{1-(1-\epsilon)(1-\sigma)}. & (3.104)
\end{aligned}
$$

Define

$$
\kappa(\epsilon) = \frac{1}{4}\frac{1-(1-\epsilon)(1-\sigma)}{1-(1-\bar{s})(1-\epsilon)(1-\sigma)}. \tag{3.105}
$$

Then, in all periods $\tau \geq t + T_\epsilon$, we have

$$
\begin{aligned}
\frac{y^u_{M,C}(\tau)}{y_{Q,D}(\tau)} &> \kappa(\epsilon), & (3.106)\\
\frac{y^u_{M,D}(\tau)}{y_{Q,D}(\tau)} &> \kappa(\epsilon), & (3.107)
\end{aligned}
$$

and therefore

$$
y^u(\tau) < \left(1 + \frac{1}{2\kappa(\epsilon)}\right)(y^u_{M,C}(\tau) + y^u_{M,D}(\tau)). \tag{3.108}
$$

Define $s(t)$ as the average probability in period t that an agent in the market enters a long-term relationship, i.e.

$$
s(t) = \frac{y^u_{M,C}(t)}{y^u(t)} s_{M,C}(t) + \frac{y^u_{M,D}(t)}{y^u(t)} s_{M,D}(t). \tag{3.109}
$$

With this, we can estimate

$$
s(\tau) \geq \underline{s}\frac{y^u_{M,C}(\tau)^2 + y^u_{M,D}(\tau)^2}{y^u(\tau)^2} \geq \frac{\frac{1}{2}\underline{s}}{\left(1 + \frac{1}{2\kappa(\epsilon)}\right)^2} \equiv \bar{s}(\epsilon) > 0. \tag{3.110}
$$

for all $\tau \geq t + T_\epsilon$. Note that

$$
\lim_{\epsilon \to 0} \bar{s}(\epsilon) = \frac{\frac{1}{2}\underline{s}}{\left(1 + \frac{2-2(1-\bar{s})(1-\sigma)}{\sigma}\right)^2}. \tag{3.111}
$$

From

$$
y^m(t+1) = (1-\epsilon)^2 y^m(t) + (1-\epsilon)^2 s(t) y^u(t) \tag{3.112}
$$

it follows that there is a finite period $t + T'_\epsilon$ with $T'_\epsilon \geq T_\epsilon$, such that in all periods $\tau \geq t + T'_\epsilon$, we have

$$
y^u(\tau) \leq \frac{1-(1-\epsilon)^2}{1-(1-\epsilon)^2(1-\bar{s}(\epsilon))}. \tag{3.113}
$$

Using (3.102), (3.111) and (3.113), we than can calculate that in all periods $\tau \geq t + T'_\epsilon$, we have

$$
\lim_{\epsilon \to 0} s_{M,C}(\tau) > \frac{1}{24}\frac{\underline{s}}{\left(1 + \frac{2-2(1-\bar{s})(1-\sigma)}{\sigma}\right)^2 (1-(1-\bar{s})(1-\sigma))} \equiv \bar{s}_{M,C} > 0, \tag{3.114}
$$

where the last inequality holds as σ is positive. Set $T'_\epsilon = T^*_\epsilon$.

Assume now that $Y(\tau) \in \Delta^A$ for all $\tau \in \{t', ..., t'+k\}$ with $t' \geq t + T^*_\epsilon$. Then, we have

$$
y^u_{M,C}(\tau+1) < (1-\epsilon)(1-\sigma)(1-s_{M,C}(\tau))y^u_{M,C}(\tau) + \frac{4}{3}\epsilon - \epsilon^2 \tag{3.115}
$$

and it holds true for some $\tilde{s}_{M,C} \in (0,1)$ that

$$y^m_{M,C}(t'+k) > \frac{1}{3}\epsilon\tilde{s}_{M,C}\sum_{i=0}^{k-1}(1-\epsilon)^{2i} + (1-\epsilon)^{2k}\tilde{s}_{M,C}y^u_{M,C}(t'), \tag{3.116}$$

$$\begin{aligned} y^u_{M,C}(t'+k) &< \left(\frac{4}{3}\epsilon-\epsilon^2\right)\sum_{i=0}^{k-1}(1-\epsilon)^i(1-\sigma)^i(1-\tilde{s}_{M,C})^i + \\ &+(1-\epsilon)^k(1-\sigma)^k(1-\tilde{s}_{M,C})^k y^u_{M,C}(t'), \end{aligned} \tag{3.117}$$

as (3.116) is strictly increasing in $\tilde{s}_{M,C}$ while (3.117) is strictly decreasing in $\tilde{s}_{M,C}$. These inequalities can be simplified to

$$y^m_{M,C}(t'+k) > \frac{1-(1-\epsilon)^{2k}}{1-(1-\epsilon)^2}\left(\frac{1}{3}\tilde{s}_{M,C}\epsilon\right) + (1-\epsilon)^{2k}\tilde{s}_{M,C}y^u_{M,C}(t'), \tag{3.118}$$

$$\begin{aligned} y^u_{M,C}(t'+k) &< \frac{1-(1-\epsilon)^k(1-\sigma)^k(1-\tilde{s}_{M,C})^k}{1-(1-\epsilon)(1-\sigma)(1-\tilde{s}_{M,C})}\left(\frac{4}{3}\epsilon-\epsilon^2\right) \\ &+(1-\epsilon)^k(1-\sigma)^k(1-\tilde{s}_{M,C})^k y^u_{M,C}(t'). \end{aligned} \tag{3.119}$$

We therefore get with l'Hospitals rule

$$\lim_{\epsilon\to 0}\frac{y^m_{M,C}(t'+k)}{y^u_{M,C}(t'+k)} > \min\left\{\frac{1}{4}k\bar{s}_{M,C}(1-(1-\sigma)(1-\bar{s}_{M,C})), \frac{\bar{s}_{M,C}}{(1-\sigma)^k(1-\bar{s}_{M,C})^k}\right\}. \tag{3.120}$$

From observation **(2)** and the specifications in (3.95), (3.96) and (3.97), we get

$$h_{M,C}(t'+k) > \frac{1+Z}{G+Z} \tag{3.121}$$

for some $k \le \bar{k}$, such that $Y(t'+k) \in \Delta^B$ if ϵ is sufficiently small.
Q.E.D.

Lemma 2.2 *If ϵ is sufficiently small, then there is a finite period $t+T^{**}_\epsilon$, such that $Y(t+T^{**}_\epsilon) \in \Delta^C$.*

Proof: Assume that the claim does not hold. Then, from lemma 2.1 it follows that we have an infinite subsequence $\{t^p\}_{p\in\mathbb{N}}$ of $\{t, t+1, ...\}$ with $Y(t^p) \in \Delta^B$ for all $p \in \mathbb{N}$ and $t^{p+1} \le t^p + \bar{k} + 1$. We can estimate

$$y^u_{M,C}(t^p+1) > (1-\bar{s})(1-\epsilon)\sigma y^u(t^p), \tag{3.122}$$

$$s_{M,C}(t^p+1) \ge \frac{\underline{s}(1-\bar{s})(1-\epsilon)\sigma y^u(t^p)}{2y^u(t^p)} = \frac{\underline{s}(1-\bar{s})(1-\epsilon)\sigma}{2}, \tag{3.123}$$

using the fact that we have $y^u(t^p+1) < 2y^u(t^p)$, as in the learning-phase no long-term relationships are broken up under the orders A and B, and $y^u(\tau) > \epsilon$ for $\tau \in \{1,2,...\}$. With (3.122) and (3.123), it follows from

$$y^m_{M,C}(t^p+2) > (1-\epsilon)^4 y^m_{M,C}(t^p) + (1-\epsilon)^2 s_{M,C}(t^p+1)y^u_{M,C}(t^p+1) \tag{3.124}$$

that

$$y^m_{M,C}(t^{p+1}) > (1-\epsilon)^{2(\bar{k}+1)}y^m_{M,C}(t^p) + (1-\epsilon)^{2(\bar{k}+1)}\frac{\underline{s}(1-\bar{s})^2\sigma^2}{2}y^u(t^p). \tag{3.125}$$

With (3.125) and $y^u(\tau) > \epsilon$ for $\tau \in \{1,2,...\}$, it follows that for each $\epsilon > 0$, there is a finite period $t+T''_\epsilon$, such that for all $t^p \ge t + T'_\epsilon$, $p \in \mathbb{N}$, we have

$$y^m_{M,C}(t^p) > \frac{(1-\epsilon)^{2(\bar{k}+1)}\frac{\underline{s}(1-\bar{s})^2\sigma^2}{2}\epsilon}{1-(1-\epsilon)^{2(\bar{k}+1)}}. \tag{3.126}$$

As in the proof of lemma 2.1, we get that for each $\epsilon > 0$, there is a finite period $t + T_\epsilon'''$ with $T_\epsilon''' \geq T_\epsilon''$, such that in all periods $\tau \geq t + T_\epsilon'''$, we have

$$y^u(\tau) \leq \frac{1-(1-\epsilon)^2}{1-(1-\epsilon)^2(1-\bar{s}(\epsilon))}, \tag{3.127}$$

where

$$\lim_{\epsilon\to 0} \bar{s}(\epsilon) = \frac{\frac{1}{2}\underline{s}}{\left(1+\frac{2-2(1-\bar{s})(1-\sigma)}{\sigma}\right)^2}. \tag{3.128}$$

Combining (3.126) and (3.127), we get for any $t^p \geq t + T_\epsilon'''$ that

$$\lim_{\epsilon\to 0} \frac{y^m_{M,C}(t^p)}{y^u(t^p)} = \infty. \tag{3.129}$$

It follows from observation **(2)** that there is a finite period $t + T_\epsilon^{**}$, such that $Y(t + T_\epsilon^{**}) \in \Delta^C$ if ϵ is sufficiently small.
Q.E.D.

Lemma 2.3 *Assume that $Y(t) \in \Delta^C$. If ϵ is sufficiently small, then $Y(t+k+1) \in \Delta^B \cup \Delta^C$ for some $k \leq \bar{k}$.*

Proof: Under all orders, we have

$$y^m_{M,C}(t+2) = (1-\epsilon)^2 y^m_{M,C}(t+1) + (1-\epsilon)^2 s_{M,C}(t+1) y^u_{M,C}(t+1). \tag{3.130}$$

If $Y(t) \in \Delta^C$, then we can estimate

$$y^u_{M,C}(t+1) > (1-\epsilon)\sigma(1-y^m_{M,C}(t)), \tag{3.131}$$

$$s_{M,C}(t+1) \geq \underline{s}\frac{(1-\epsilon)\sigma(1-y^m_{M,C}(t))}{y^u(t+1)} > \underline{s}(1-\epsilon)\sigma(1-y^m_{M,C}(t)). \tag{3.132}$$

Thus, from (3.130) it follows that we have

$$y^m_{M,C}(t+2) > \underline{s}(1-\epsilon)^4\sigma^2 \tag{3.133}$$

and

$$y^u_{M,C}(t+2) < 1 - \underline{s}(1-\epsilon)^4\sigma^2. \tag{3.134}$$

If $Y(\tau) \in \Delta^A$, then we have

$$\begin{aligned} y^u_{M,C}(\tau+1) &< (1-\epsilon)(1-\sigma)y^u_{M,C}(\tau) + \frac{4}{3}\epsilon - \epsilon^2, &(3.135)\\ y^m_{M,C}(\tau+1) &> (1-\epsilon)^2 y^m_{M,C}(\tau). &(3.136)\end{aligned}$$

Assume that $Y(\tau) \in \Delta^A$ for $\tau \in [t+1, t+k]$. We then can estimate

$$\begin{aligned} y^u_{M,C}(t+k+1) &< (1-\epsilon)^{k-1}(1-\sigma)^{k-1}\left(1-\underline{s}(1-\epsilon)^4\sigma^2\right) + k\left(\frac{4}{3}\epsilon - \epsilon^2\right), &(3.137)\\ y^m_{M,C}(t+k+1) &> (1-\epsilon)^{2(k-1)}\underline{s}(1-\epsilon)^4\sigma^2. &(3.138)\end{aligned}$$

This yields us

$$\lim_{\epsilon\to 0} \frac{y^m_{M,C}(t+k+1)}{y_{M,C}(t+k+1)} = \frac{\underline{s}\sigma^2}{\underline{s}\sigma^2 + (1-\sigma)^{k-1}(1-\underline{s}\sigma^2)}. \tag{3.139}$$

From observation **(2)** and the specification in (3.98) the result follows.
Q.E.D.

Lemma 2.4 *Assume that $Y(t) \in \Delta^C$. If ϵ is sufficiently small, then there is a finite period $t+T^{***}$ with $y^m_{M,C}(t+T^{***}) > \bar{y}$.*

Proof: As shown in the proof of lemma 2.3, from $Y(t) \in \Delta^C$ it follows that

$$y^m_{M,C}(t+2) > \underline{s}(1-\epsilon)^4\sigma^2. \tag{3.140}$$

Consider the subsequence $\{t^p\}_{p\in\mathbb{N}}$ of $\{t, t+1, ...\}$ with $t^0 = t$ and $Y(t^p) \in \Delta^B \cup \Delta^C$ for all $p \in \mathbb{N}$. With (3.140) we can estimate

$$y^m_{M,C}(t^1) > (1-\epsilon)^{2(\bar{k}+1)}\underline{s}\sigma^2. \tag{3.141}$$

Assume that in all periods $\tau \in \{t+2, t+3, ...\}$ it holds that

$$y^m_{M,C}(\tau) > (1-\epsilon)^{2(\bar{k}+1)}\underline{s}\sigma^2. \tag{3.142}$$

Choose an $\alpha \in (0, 1-\bar{y})$ small enough such that

$$\frac{\underline{s}\sigma^2}{\underline{s}\sigma^2+\alpha} > \frac{1+Z}{G+Z}. \tag{3.143}$$

From observation **(2)** and the assumption in (3.142), it follows that $Y(t^p) \in \Delta^C$ whenever $y^u(t^p) \leq \alpha$ and ϵ is sufficiently small. If the two conditions

$$(1-\epsilon)^{2(\bar{k}+1)} - 1 + (1-\epsilon)^{2(\bar{k}+1)}\underline{s}\sigma^2\alpha^2 > d, \tag{3.144}$$

$$(1-\epsilon)^{2(\bar{k}+1)} - 1 + (1-\epsilon)^{2(\bar{k}+1)}\underline{s}\sigma^2(1-\bar{y})^2 > d \tag{3.145}$$

hold for a small $d > 0$, then we have

$$y^m_{M,C}(t^{p+1}+2) > y^m_{M,C}(t^p+2) + d \tag{3.146}$$

for all $p \in \mathbb{N}$, as long as $y^m_{M,C}(t^p+2) \leq \bar{y}$, regardless of the order in the periods $\{t^p\}_{p\in\mathbb{N}}$. Thus, if ϵ is sufficiently small, then the assumption in (3.142) is justified, and there is a finite period $t+T^{***}$, such that $y^m_{M,C}(t+T^{***}) > \bar{y}$.
Q.E.D.

From observation **(2)**, the lemmas 2.2, 2.4 and the specification of $\bar{y}$ we know that for sufficiently small ϵ, there is a period $\bar{t} < \infty$, such that in all periods $t \geq \bar{t}$, we have $Y(t) \in \Delta^C$.
Q.E.D.

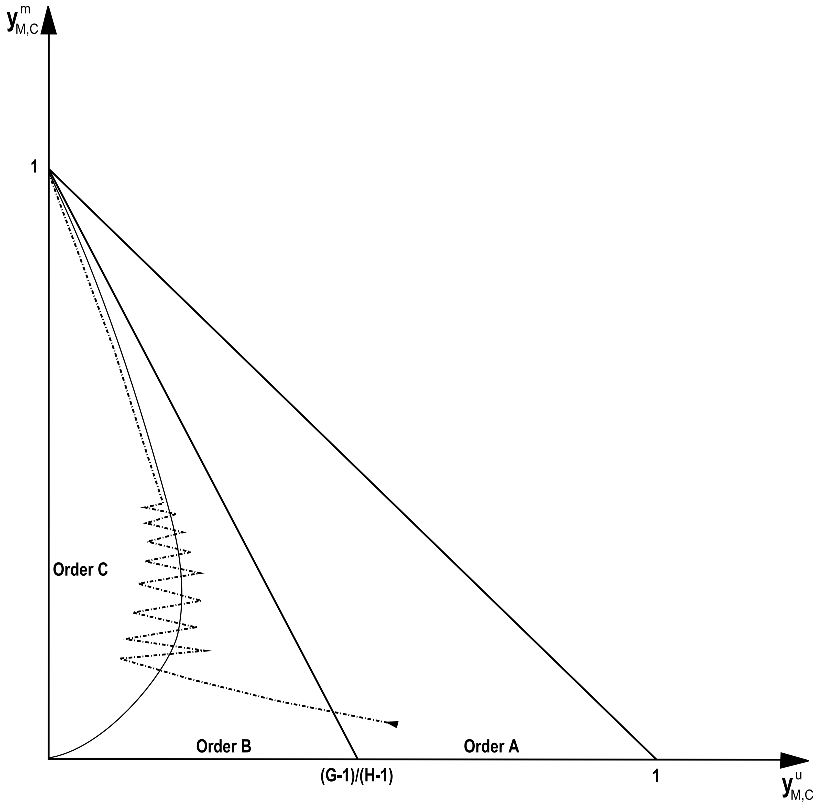

FIGURE I

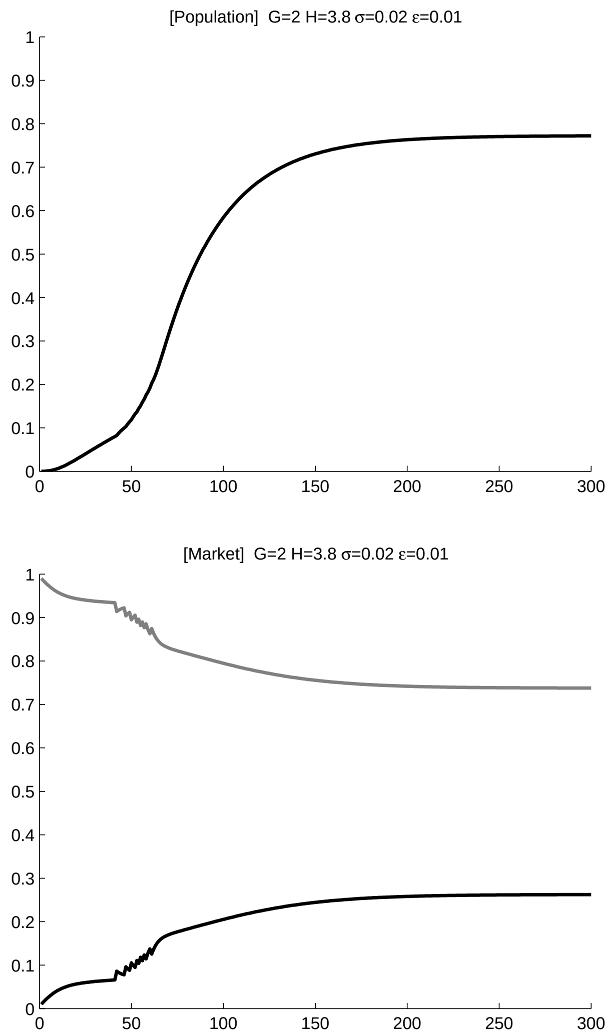

FIGURE II

Caption: Top: [Black] Share of players in cooperative long-term relationships, $y^m_{M,C}$. **Bottom: [Black]** Share of (M,C)-players in the market, $\frac{y^u_{M,C}}{y^u}$. **[Gray]** Share of (Q,D)-players in the market, $\frac{y_{Q,D}}{y^u}$. **Parameters:** $H = 3.8$, $G = 2$, $\sigma = 0.02$, $\epsilon = 0.01$.

Starting values: $y^m_{M,C}(0) = 0.000$ - $y^u_{M,C}(0) = 0.010$ - $y_{Q,D}(0) = 0.990$.

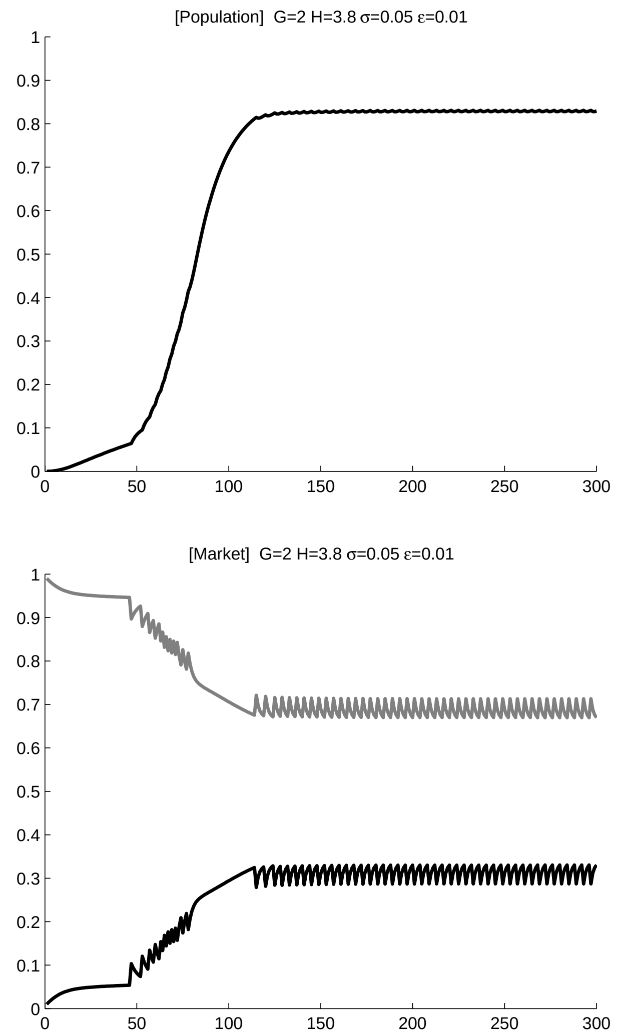

FIGURE III

Caption: Top: [Black] Share of players in cooperative long-term relationships, $y^m_{M,C}$. **Bottom: [Black]** Share of (M, C)-players in the market, $\frac{y^u_{M,C}}{y^u}$. **[Gray]** Share of (Q, D)-players in the market, $\frac{y_{Q,D}}{y^u}$. **Parameters:** $H = 3.8$, $G = 2$, $\sigma = 0.05$, $\epsilon = 0.01$.

Starting values: $y^m_{M,C}(0) = 0.000$ - $y^u_{M,C}(0) = 0.010$ - $y_{Q,D}(0) = 0.990$.

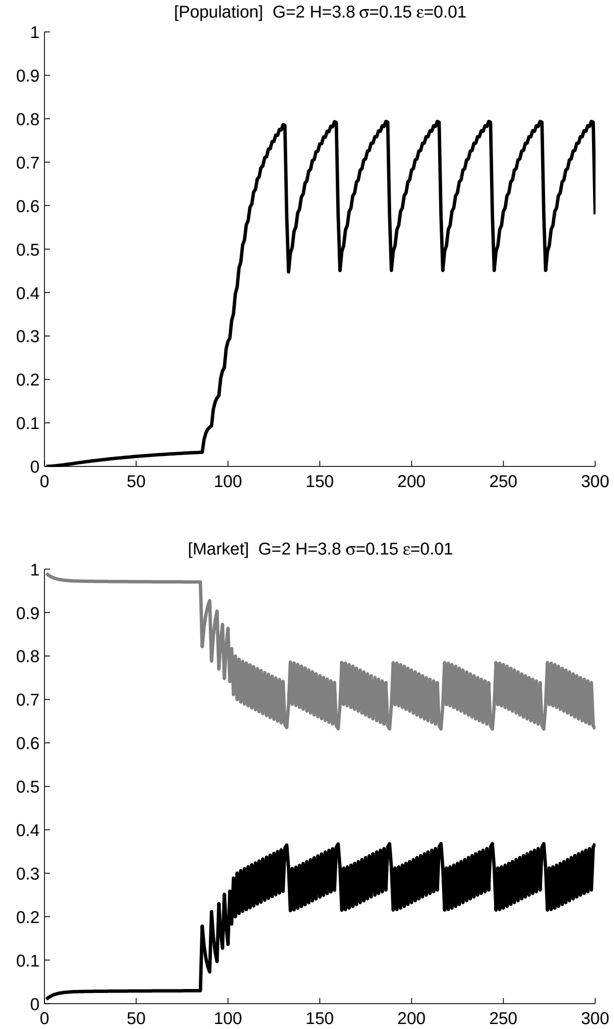

FIGURE IV

Caption: Top: [Black] Share of players in cooperative long-term relationships, $y^m_{M,C}$. **Bottom: [Black]** Share of (M,C)-players in the market, $\frac{y^u_{M,C}}{y^u}$. **[Gray]** Share of (Q,D)-players in the market, $\frac{y_{Q,D}}{y^u}$. **Parameters:** $H = 3.8$, $G = 2$, $\sigma = 0.15$, $\epsilon = 0.01$.

Starting values: $y^m_{M,C}(0) = 0.000$ - $y^u_{M,C}(0) = 0.010$ - $y_{Q,D}(0) = 0.990$.

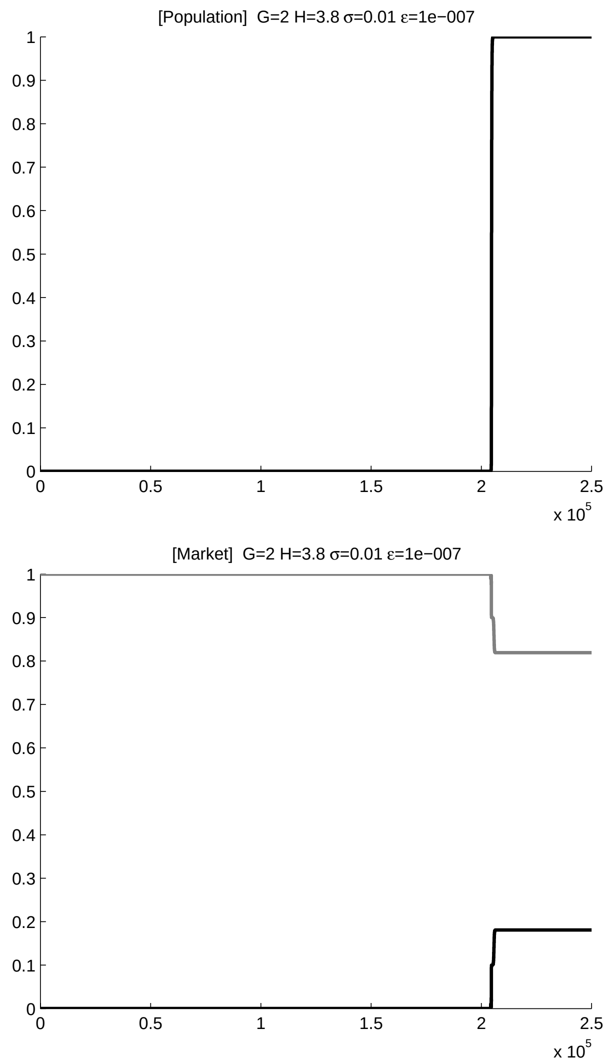

FIGURE V

Caption: Top: [Black] Share of players in cooperative long-term relationships, $y^m_{M,C}$. **Bottom: [Black]** Share of (M, C)-players in the market, $\frac{y^u_{M,C}}{y^u}$. **[Gray]** Share of (Q, D)-players in the market, $\frac{y_{Q,D}}{y^u}$. **Parameters:** $H = 3.8$, $G = 2$, $\sigma = 0.01$, $\epsilon = 0.1 \times 10^{-6}$.

Starting values: $y^m_{M,C}(0) = 0.00$ - $y^u_{M,C}(0) = 0.1 \times 10^{-15}$ - $y_{Q,D}(0) = 1 - 0.1 \times 10^{-15}$.

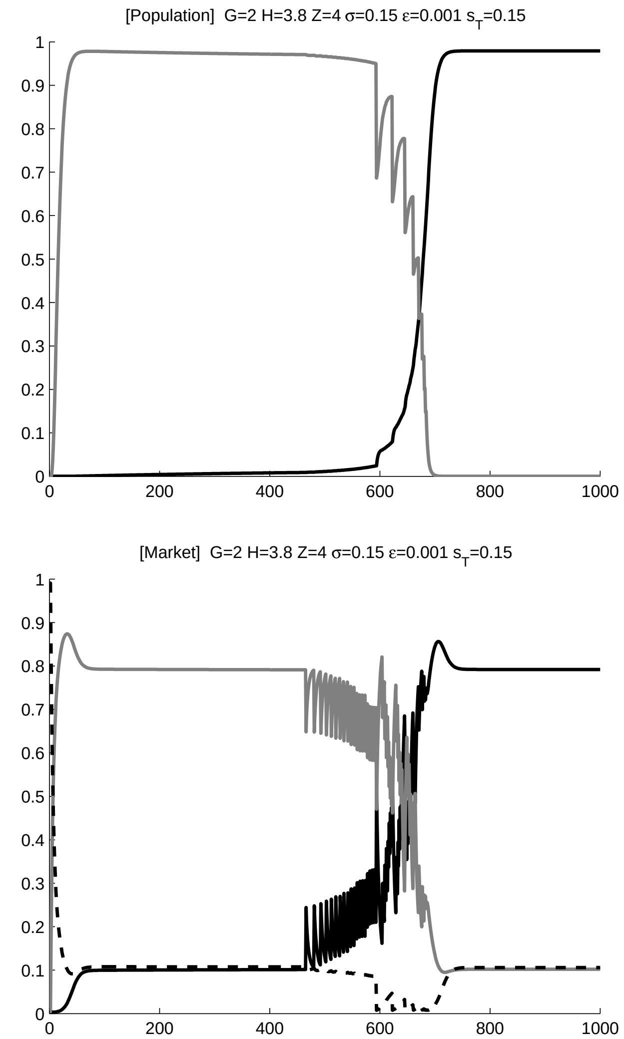

FIGURE VI

Caption: Top: [Black] Share of players in cooperative long-term relationships, $y^m_{M,C}$. **[Gray]** Share of players in non-cooperative long-term relationships, $y^m_{M,D}$. **Bottom: [Black]** Share of (M, C)-players in the market, $\frac{y^u_{M,C}}{y^u}$. **[Gray]** Share of (M, D)-players in the market, $\frac{y_{M,D}}{y^u}$. **[Black-dotted]** Share of (Q, D)-players in the market, $\frac{y_{Q,D}}{y^u}$. **Parameters:** $H = 3.8$, $G = 2$, $Z = 4$, $\sigma = 0.15$, $\epsilon = 0.001$, $s = 0.15$.

Starting values: $y^m_{M,C}(0) = 0.000$ - $y^u_{M,C}(0) = 0.005$ - $y^m_{M,D}(0) = 0.000$ - $y^u_{M,D}(0) = 0.005$ - $y_{Q,D}(0) = 0.990$.

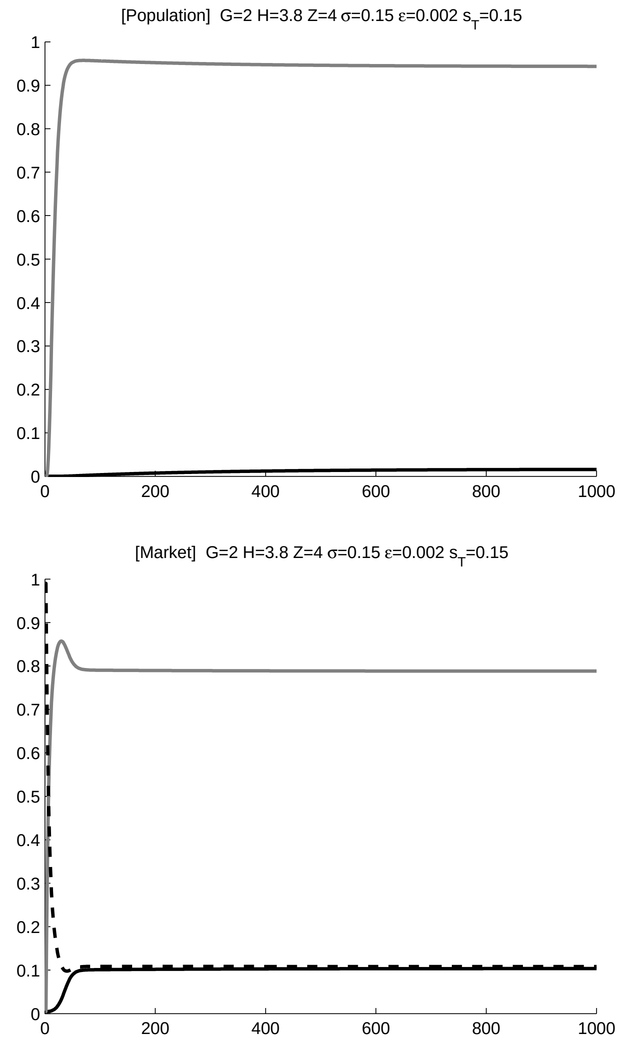

FIGURE VII

Caption: Top: [Black] Share of players in cooperative long-term relationships, $y^m_{M,C}$. **[Gray]** Share of players in non-cooperative long-term relationships, $y^m_{M,D}$. **Bottom: [Black]** Share of (M,C)-players in the market, $\frac{y^u_{M,C}}{y^u}$. **[Gray]** Share of (M,D)-players in the market, $\frac{y_{M,D}}{y^u}$. **[Black-dotted]** Share of (Q,D)-players in the market, $\frac{y_{Q,D}}{y^u}$. **Parameters:** $H = 3.8$, $G = 2$, $Z = 4$, $\sigma = 0.15$, $\epsilon = 0.002$, $s = 0.15$.

Starting values: $y^m_{M,C}(0) = 0.000$ - $y^u_{M,C}(0) = 0.005$ - $y^m_{M,D}(0) = 0.000$ - $y^u_{M,D}(0) = 0.005$ - $y_{Q,D}(0) = 0.990$.

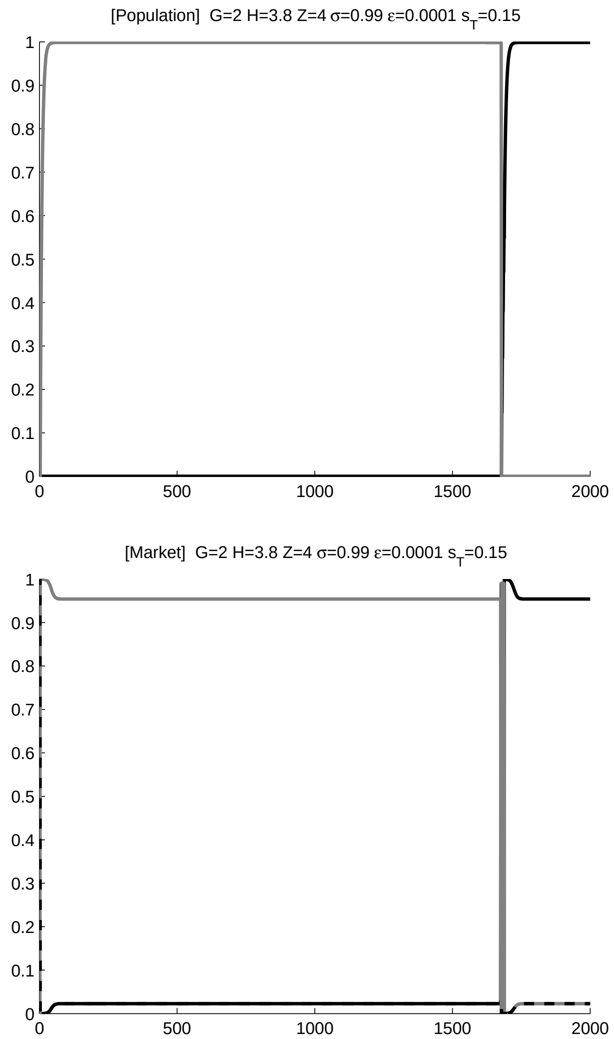

FIGURE VIII

Caption: Top: [Black] Share of players in cooperative long-term relationships, $y^m_{M,C}$. **[Gray]** Share of players in non-cooperative long-term relationships, $y^m_{M,D}$. **Bottom: [Black]** Share of (M,C)-players in the market, $\frac{y^u_{M,C}}{y^u}$. **[Gray]** Share of (M,D)-players in the market, $\frac{y_{M,D}}{y^u}$. **[Black-dotted]** Share of (Q,D)-players in the market, $\frac{y_{Q,D}}{y^u}$. **Parameters:** $H = 3.8$, $G = 2$, $Z = 4$, $\sigma = 0.99$, $\epsilon = 0.0001$, $s = 0.15$.

Starting values: $y^m_{M,C}(0) = 0.000$ - $y^u_{M,C}(0) = 0.005$ - $y^m_{M,D}(0) = 0.000$ - $y^u_{M,D}(0) = 0.005$ - $y_{Q,D}(0) = 0.990$.

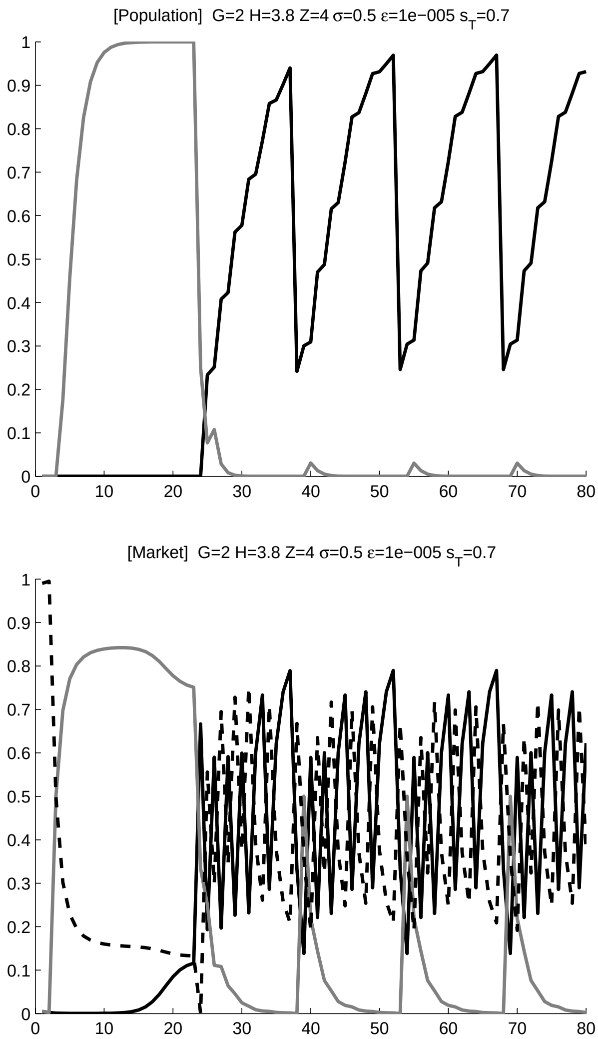

FIGURE IX

Caption: Top: [Black] Share of players in cooperative long-term relationships, $y^m_{M,C}$. **[Gray]** Share of players in non-cooperative long-term relationships, $y^m_{M,D}$. **Bottom: [Black]** Share of (M,C)-players in the market, $\frac{y^u_{M,C}}{y^u}$. **[Gray]** Share of (M,D)-players in the market, $\frac{y_{M,D}}{y^u}$. **[Black-dotted]** Share of (Q,D)-players in the market, $\frac{y_{Q,D}}{y^u}$. **Parameters:** $H = 3.8$, $G = 2$, $Z = 4$, $\sigma = 0.50$, $\epsilon = 0.00001$, $s = 0.70$.

Starting values: $y^m_{M,C}(0) = 0.000$ - $y^u_{M,C}(0) = 0.005$ - $y^m_{M,D}(0) = 0.000$ - $y^u_{M,D}(0) = 0.005$ - $y_{Q,D}(0) = 0.990$.

Chapter 4

On the Dynamics in a Market for long-term Relationships

4.1 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,

[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).

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 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 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 his 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.

4.2 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".

4.2.1 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

$$g(h_{T^i}) = \begin{cases} 1 & \text{if } T^i > T^* \text{ and } h_{T^i} = h^c_{T^i} \\ 0 & \text{if } T^i \leq T^* \text{ and } h_{T^i} = h^c_{T^i} \\ -1 & \text{otherwise} \end{cases}. \tag{4.1}$$

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.

4.2.2 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

$$hs^{c(t)} = \{g(h_{T^{-n+1}}), ..., g(h_{T^{c(t)}})\} \tag{4.2}$$

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

$$\{g(h_{T^{-n+1}}), ..., g(h_{T^0})\}, \tag{4.3}$$

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

$$HS^c(n) = \big\{\{g_i\}_{i\in\{1,...,n+c\}} \mid g_i \in \{-1, 0, 1\}\big\} \tag{4.4}$$

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

$$HS(n) = \bigcup_{c>0} HS^c(n) \tag{4.5}$$

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

$$\tilde{\mu} : HS(n) \to [0, 1]. \tag{4.6}$$

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

4.2.3 Sequence of events and strategy choice

The sequence of events in each period is 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.

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

$$E[f^c_{T^*}, \tilde{\mu}(t)] = \frac{1 + \delta^{T^*}(1-\sigma)^{T^*}(\tilde{\mu}(t)G - 1)}{(1-\delta)(1-\delta^{T^*+1}(1-\sigma)^{T^*+1}(1-\tilde{\mu}(t)))} \tag{4.7}$$

and

$$E[f^d_{T^*}, \tilde{\mu}(t)] = \frac{1 + \delta^{T^*}(1-\sigma)^{T^*}[(1-\delta(1-\sigma))(\tilde{\mu}(t)H + 1 - \tilde{\mu}(t)) - 1]}{(1-\delta)(1-\delta^{T^*+1}(1-\sigma)^{T^*+1})}. \tag{4.8}$$

We summarize the collection of parameters of the game by $\Gamma = \{H, G, \delta, \sigma\}$. With equations (4.7) and (4.8), 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 $\bar{\mu}_{\Gamma_1,2} = 1$.

4.2.4 Nash equilibria in symmetric strategies

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

$$0 < \underline{\mu}_{\Gamma,T^*} < \bar{\mu}_{\Gamma,T^*} \leq 1. \tag{4.9}$$

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.

4.2.5 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

$$Y(t) \in \Delta(\mathcal{B}^{T^*+2}) \tag{4.10}$$

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

$$Y^{HS}(t) \in \Delta(HS(n)^{T^*+2}) \tag{4.11}$$

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

$$\mathcal{Y} = \{Y(t)\}_{t \geq 0} \tag{4.12}$$

and

$$\mathcal{Y}^{HS} = \left\{Y^{HS}(t)\right\}_{t \geq 0} \tag{4.13}$$

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

$$\tau_{\tilde{\mu}} : \Delta(HS(n)^{T^*+2}) \to \Delta(\mathcal{B}^{T^*+2}), \tag{4.14}$$

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

$$\mathcal{Y} = \{\tau_{\tilde{\mu}^{FP}}(Y^{HS}(t))\}_{t \geq 0}. \tag{4.15}$$

4.2.6 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 $Y(t)$.

4.3 On the dynamics under fictitious play

4.3.1 Infinite memory

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:

$$\kappa_c(t) = \sum_{i=-n+1}^{c(t)} \mathbf{1}_{\{g(h_{Ti})=1\}}, \tag{4.16}$$

$$\kappa_d(t) = \sum_{i=-n+1}^{c(t)} \mathbf{1}_{\{g(h_{Ti})=-1\}}, \tag{4.17}$$

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

$$\tilde{\mu}^{FP}(t) = \frac{\kappa_c(t)}{\kappa_c(t) + \kappa_d(t)}. \tag{4.18}$$

The subjective belief in the first period, $\tilde{\mu}^{FP}(0)$, therefore is determined by (4.3): the more elements in (4.3) 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

$$\mathcal{B} = \mathbb{R} \cap [0, 1]. \tag{4.19}$$

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.

4.3.2 Finite memory

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

$$\mathcal{B} = \left\{0, \frac{1}{n}, ..., \frac{n-1}{n}, 1\right\}. \tag{4.20}$$

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

$$k(t) = \arg_{k\in\mathbb{N}} \left\{ \sum_{i=k}^{c(t)} \mathbf{1}_{\{g(h_{T^i})\neq 0\}} = n \right\}. \tag{4.21}$$

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

$$\tilde{\mu}^A(t) = \frac{1}{n} \sum_{i=k(t)}^{c(t)} \mathbf{1}_{\{g(h_{T^i})=1\}}. \tag{4.22}$$

The subjective belief in the first period is again determined by the vector given in (4.3).

As $\tilde{\mu}(t)$ can take on only finitely many values given in (4.20), we introduce the following notation related to the boundaries $\underline{\mu}_{\Gamma,T^*}$ and $\bar{\mu}_{\Gamma,T^*}$:

$$\underline{k}_{\Gamma,T^*,n} = [n\underline{\mu}_{\Gamma,T^*}]_+, \tag{4.23}$$
$$\bar{k}_{\Gamma,T^*,n} = [n\bar{\mu}_{\Gamma,T^*} - 1]_+, \tag{4.24}$$

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

$$y_C(t) = \sum_{i=\underline{k}}^{\bar{k}} y_i(t), \tag{4.25}$$

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

$$y_D(t) = \sum_{i=0}^{\underline{k}} y_i(t) + \sum_{i=\bar{k}}^{n} y_i(t). \tag{4.26}$$

With these specifications, we get

$$\mu(t) = \frac{y_C(t)}{y_D(t) + y_C(t)} \tag{4.27}$$

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 $y^l_i(t)$. 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

$$y_i^{l+1}(t) = (1-\sigma)y_i^l(t-1) \tag{4.28}$$

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.

4.4 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 the authors' webpage.

- 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.

4.5 Extension: 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

[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.

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 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 [Some 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

$$\mathcal{X}^{HS}_N = \left\{X^{HS}(t)\right\}_{t\geq 0} \tag{4.29}$$

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* $X_N^{HS}(0) = Y^{HS}(0)$, *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* $t > 0$.

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,

$$\begin{aligned} B(Y^*) \;=\; & \left\{Y^{HS} \in \Delta(HS(n)^{T^*+2}) \,:\, Y^{HS}(0) = Y^{HS} \Rightarrow \right. \\ & \left. d(Y(t), Y^*) \to 0 \text{ as } t \to \infty\right\}. \end{aligned} \tag{4.30}$$

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

[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 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.

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^*.

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.

To illustrate the importance of this result, consider scenario Γ^1 from the simulations with $n = 20$ (where we had $\underline{k} = 2$): Under $\tilde{\mu}^A$ we always observed convergence to a steady state with cooperation. However, Proposition 5 says that cooperation brakes down in finite time. This happens after a disadvantagous matching of cooperative and non-cooperative agents, such that the subjective beliefs of the fist group worsen (and therefore these agents switch to the non-cooperative strategy), while the subjective beliefs of the second group remain low. If we consider any scenario with $\underline{k} = 1$, cooperation cannot disappear.

4.6 Conclusion

In this paper we analyzed the dynamics on a market for long-term relationships when agents learn from their observations about aggregate play. 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 result is a population in which different agents make different experiences in the market and therefore act differently even if aggregate play remains constant. 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.

4.7 Appendix

Calculation of (4.7) and (4.8)

We calculate these two terms from the expressions

$$\begin{aligned} E[f^c_{T^*}\tilde{\mu}(t)] &= \tilde{\mu}(t)\left[\sigma\sum_{T=0}^{T^*-1}\left[(1-\sigma)^T\sum_{\tau=0}^{T}\delta^\tau\right]+\sigma\sum_{T=T^*}^{\infty}\left[(1-\sigma)^T\left(\sum_{\tau=0}^{T^*-1}\delta^\tau+\sum_{\tau=T^*}^{T}\delta^\tau G\right)\right]\right. \\ &\left.+\sigma\sum_{\tau=0}^{\infty}\delta^{\tau+1}(1-\sigma)^\tau E[f^c_{T^*},\tilde{\mu}(t)]\right]+(1-\tilde{\mu}(t))\left[\sigma\sum_{T=0}^{T^*-1}\left[(1-\sigma)^T\sum_{\tau=0}^{T}\delta^\tau\right]\right. \\ &\left.+\sigma\sum_{\tau=0}^{T^*-1}\delta^{\tau+1}(1-\sigma)^\tau E[f^c_{T^*},\tilde{\mu}(t)]+\delta^{T^*+1}(1-\sigma)^{T^*}E[f^c_{T^*}\tilde{\mu}(t)]\right], \end{aligned} \tag{4.31}$$

and

$$\begin{aligned} E[f^d_{T^*}\tilde{\mu}(t)] &= \sigma\sum_{T=0}^{T^*-1}\left[(1-\sigma)^T\sum_{\tau=0}^{T}\delta^\tau\right]+\tilde{\mu}(t)\delta^{T^*}(1-\sigma)^{T^*}H+(1-\tilde{\mu}(t))\delta^{T^*}(1-\sigma)^{T^*} \\ &+\sigma\sum_{\tau=0}^{T^*-1}\delta^{\tau+1}(1-\sigma)^\tau E[f^d_{T^*},\tilde{\mu}(t)]+\delta^{T^*+1}(1-\sigma)^{T^*}E[f^d_{T^*},\tilde{\mu}(t)]. \end{aligned} \tag{4.32}$$

Proof of Lemma 1

First consider the ratio

$$E_{T^*}(\tilde{\mu})=\frac{E[f^c_{T^*},\tilde{\mu}]}{E[f^d_{T^*},\tilde{\mu}]}. \tag{4.33}$$

Taking the limit yields us

$$\lim_{\sigma\to 0}E_{T^*}(\tilde{\mu})=\frac{[1+\delta^{T^*}(\tilde{\mu}G-1)][1-\delta^{T^*+1}]}{[1-(1-\tilde{\mu})\delta^{T^*+1}][1+\delta^{T^*}((1-\delta)(\tilde{\mu}H+1-\tilde{\mu})-1)]}, \tag{4.34}$$

and with l'Hospitals rule

$$\lim_{\delta\to 1}\lim_{\sigma\to 0}E_{T^*}(\tilde{\mu})=\frac{G(T^*+1)}{(T^*+1)+\tilde{\mu}(H-1)}, \tag{4.35}$$

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 (4.35) is larger than 1 if

$$\tilde{\mu}<(T^*+1)\frac{G-1}{H-1}. \tag{4.36}$$

As δ and σ enter (4.33) continuously, part (b) of the result follows from (4.35).

Now fix a $\tilde{\mu}^*$ such that the right-hand side of (4.35) 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

$$E[f^c_{T^*}, \tilde{\mu}^*] > E[f^d_{T^*}, \tilde{\mu}^*] \tag{4.37}$$

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

$$\lim_{\tilde{\mu}\to 0} E_{T^*}(\tilde{\mu}) = \frac{1-\delta^{T^*}(1-\sigma)^{T^*}}{1-\delta^{T^*+1}(1-\sigma)^{T^*+1}} < 1. \tag{4.38}$$

Define

$$\nabla E_{T^*}(\tilde{\mu}) = E[f^c_{T^*}, \tilde{\mu}] - E[f^d_{T^*}, \tilde{\mu}]. \tag{4.39}$$

From (4.37) and (4.38) we know that $\nabla E_{T^*}(\tilde{\mu}^*) > 0$ and $\nabla E_{T^*}(0) < 0$. Further from

$$\frac{\partial \nabla E_{T^*}(\tilde{\mu})^2}{\partial^2 \tilde{\mu}} = \frac{-2G(1-\delta)^2\delta^{T^*+1}(1-\sigma)^{T^*+1}[A-\delta^{T^*+1}(1-\sigma)^{T^*+1}(1-\delta^{T^*}(1-\sigma)^{T^*})]}{[(1-\delta)(1-\delta^{T^*+1}(1-\sigma)^{T^*+1})+(1-\delta)\delta^{T^*+1}(1-\sigma)^{T^*+1}\tilde{\mu}]^3} < 0 \tag{4.40}$$

with

$$A = \delta^{T^*}(1-\sigma)^{T^*}G(1-\delta^{T^*+1}(1-\sigma)^{T^*+1}) \tag{4.41}$$

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:

$$\forall t > 0 : Y(t) \in Y_1 \Rightarrow \mu(t) = 0 \Rightarrow Y(t+1) \notin Y_1. \tag{4.42}$$

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

$$\max\{\mu(\tau) \mid \tau \in [t, t+n]\} \leq 2^n(1-y_0(t)-y_1(t)), \tag{4.43}$$

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

$$\nabla^1(t) > \sigma(1 - 2^n(1 - y_0(t) - y_1(t)))^n(1 - y_0(t) - y_1(t)). \tag{4.44}$$

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

$$\nabla^2(t) < n\left(2^{2n}(1 - y_0(t) - y_1(t))^2 y_0(t) + 2^n(1 - y_0(t) - y_1(t))y_1(t)\right). \tag{4.45}$$

We therefore get

$$\frac{\nabla^1(t)}{\nabla^2(t)} > \frac{\sigma(1 - y_0(t) - y_1(t))}{n(2^{2n}(1 - y_0(t) - y_1(t))y_0(t) + 2^n y_1(t))}, \tag{4.46}$$

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 $y_0(t)$

$$\mu(t+n-1) < \mu(t). \tag{4.47}$$

With

$$y_1(t+1) < (1 - \mu(t))y_2(t) + \mu(t) < (1 - \mu(t))\mu(t) + \mu(t) < 2\mu(t) \tag{4.48}$$

and (4.47) we get $y_1(t+n) < 2\mu(t+n-1) < 2\mu(t)$. Thus, we have

$$\frac{\nabla^1(t+n)}{\nabla^2(t+n)} > \frac{\sigma(1 - y_0(t) - y_1(t))}{n(2^{2n}(1 - y_0(t) - y_1(t))y_0(t) + 2^{n+1}\mu(t))}. \tag{4.49}$$

If the expressions on the right-hand side of (4.46) and (4.49) are larger than 1, this yields us

$$\frac{\nabla^1(t+in)}{\nabla^2(t+in)} > \frac{\sigma}{n2^n\mu(t)} > 1 \tag{4.50}$$

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

$$\lim_{t\to\infty} y_0(t) = 1, \tag{4.51}$$

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

$$\nabla_1(t) \leq (1-\sigma)^{T^*}(1 - \mu(t))y_1(t), \tag{4.52}$$
$$\nabla_2(t) = (1-\sigma)^{T^*}\mu(t)y_0(t). \tag{4.53}$$

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

$$\lim_{t\to\infty} y_0(t) = 0, \tag{4.54}$$

which implies that

$$\lim_{t\to\infty} \mu(t) = 1, \tag{4.55}$$

as

$$y_C(t) > \sigma(1 - y_0(t-1)). \tag{4.56}$$

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

Proof of Lemma 3

We proof statement (4.29) 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 (4.29)). 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:

$$L_c = L_{c,1} + L_{c,2}. \tag{4.57}$$

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

$$\mathbf{E}(L_A^2) = \frac{L(L-1)M/2(M/2-1)}{M(M-1)} + \frac{M/2L}{M} = \frac{L(L-1)(M-2)}{4 \cdot (M-1)} + \frac{L}{2} \tag{4.58}$$

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

$$\begin{aligned} \mathbf{E}(L_{c,1}) &= \mathbf{E}(\mathbf{E}(L_{c,1}|L_A)) && (4.59) \\ &= 2 \cdot \mathbf{E}\left(\frac{L_A(L-L_A)}{M/2}\right) && (4.60) \\ &= \frac{4}{M}\left(L\mathbf{E}(L_A) - \mathbf{E}(L_A^2)\right) && (4.61) \\ &= \frac{L(L-1)}{M-1} && (4.62) \end{aligned}$$

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

$$\begin{aligned} \mathbf{E}(L_{c,2}) &= \mathbf{E}(\mathbf{E}(L_{c,2}|L_{c,1})) && (4.63) \\ &= \mathbf{E}\left(\frac{(L-L_{c,1})(C-L)}{M-L}\right) && (4.64) \\ &= \frac{\left(L - \frac{L(L-1)}{M-1}\right)(C-L)}{M-L} && (4.65) \end{aligned}$$

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

$$\mathbf{E}\left(m^M(hs, f^c_{T^*})\right) = \frac{1}{L}\left(\mathbf{E}(L_{c,1}) + \mathbf{E}(L_{c,2})\right) \quad (4.66)$$

$$= \frac{1}{L}\left(\frac{L(C-1)}{M-1}\right) \quad (4.67)$$

$$= \frac{\mu M - 1}{M-1} \rightarrow \mu \text{ as } M \rightarrow \infty \quad (4.68)$$

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

$$\mathrm{Var}\left(m^M(hs, f^c_{T^*})\right) = \frac{\mathbf{E}(L_c^2) - \mathbf{E}^2(L_c)}{L^2}$$

$$= \frac{1}{L^2}\left[\mathbf{E}(L_{c,1}^2) + \mathbf{E}(L_{c,2}^2) + 2\mathbf{E}(L_{c,1}L_{c,2}) - \mathbf{E}^2(L_c)\right] \quad (4.69)$$

Using equation (4.58) we get

$$\mathbf{E}(L_{c,1}^2) = 4\cdot\mathbf{E}\left(\frac{(L-L_A)(L-L_A-1)L_A(L_A-1)}{\frac{M}{2}(\frac{M}{2}-1)}\right) + 2\frac{L(L-1)}{M-1}$$

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

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

$$\mathbf{E}(L_A^3) = \frac{L(L-1)(L-2)M/2(M/2-1)(M/2-2)}{M(M-1)(M-2)} + 3\cdot\mathbf{E}(L_A^2) - 2\cdot\mathbf{E}(L_A)$$

$$= \frac{(M-4)L(L-1)(L-2)}{8(M-1)} + 3\cdot\mathbf{E}(L_A^2) - 2\cdot\mathbf{E}(L_A) \quad (4.70)$$

and

$$\mathbf{E}(L_A^4) = \frac{L(L-1)(L-2)(L-3)M/2(M/2-1)(M/2-2)(M/2-3)}{M(M-1)(M-2)(M-3)}$$

$$+6\cdot\mathbf{E}(L_A^3) - 11\cdot\mathbf{E}(L_A^2) + 6\cdot\mathbf{E}(L_A)$$

$$= \frac{(M-4)(M-6)L(L-1)(L-2)(L-3)}{16(M-1)(M-3)} + 6\cdot\mathbf{E}(L_A^3) - 11\cdot\mathbf{E}(L_A^2) + 6\cdot\mathbf{E}(L_A) \quad (4.71)$$

Plugging equations (4.58), (4.70) and (4.71) into the expression for $\mathbf{E}(L_{c,1}^2)$, dividing by L^2 and replacing L by $y_{hs}M$ yields

$$\frac{\mathbf{E}(L_{c,1}^2)}{L^2} \rightarrow 16\cdot\left[0 + \frac{y_{hs}^2}{4} - \frac{2y_{hs}^2}{8} + \frac{y_{hs}^2}{16}\right] + 0 = y_{hs}^2, \text{ as } M \rightarrow \infty. \quad (4.72)$$

Furthermore,

$$\mathbf{E}(L_{c,2}^2) = \mathbf{E}\left(\frac{(C-L)(C-L-1)(L-L_{c,1})(L-L_{c,1}-1)}{(M-L)(M-L-1)} + \frac{(L-L_{c,1})(C-L)}{M-L}\right)$$

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

Using $\mathbf{E}(L_{c,1})/L \rightarrow y_{hs}$ and (4.72), we get from the previous equation

$$\frac{\mathbf{E}(L_{c,2}^2)}{L^2} \rightarrow \frac{(\mu - y_{hs})^2}{(1-y_{hs})^2}\left[y_{hs}^2 - 2y_{hs} + 1\right] + 0 = (\mu - y_{hs})^2, \text{ as } M \rightarrow \infty. \quad (4.73)$$

Finally, the third term of equation (4.69) 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:

$$\frac{\mathbf{E}(L_{c,1}L_{c,2})}{L^2} \to \frac{\mu-y_{hs}}{1-y_{hs}}\left[y_{hs}-y_{hs}^2\right] = y_{hs}(\mu-y_{hs}), \text{ as } M\to\infty. \tag{4.74}$$

Since all limits exist, sums and limits and interchangeable and we finally get from equation (4.69)

$$\mathrm{Var}\left(m^M(hs, f_{T^*}^c)\right) \to y_{hs}^2 + (\mu-y_{hs})^2 + 2y_{hs}(\mu-y_{hs}) - \mu^2 = 0, \text{ as } M\to\infty. \tag{4.75}$$

It follows from (4.68) and (4.75) 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

$$||Y^{HS}|| = \sum_{hs\in HS(n)}\sum_{s\in\{0,..,T^*+1\}} |Y_{hs,s}^{HS}|. \tag{4.76}$$

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 $s \le T^*$:

$$|X_{N,i,s}^{HS}(t+1) - Y_{X_N^{HS}(t),i,s}(1)| \xrightarrow{p} 0. \tag{4.77}$$

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

$$||Y_{X_N^{HS}(1)}(1) - Y^{HS}(2)|| \xrightarrow{p} 0 \tag{4.78}$$

$$\vdots$$

$$||Y_{X_N^{HS}(t)}(1) - Y^{HS}(t+1)|| \xrightarrow{p} 0, \tag{4.79}$$

and hence

$$\sum_{s=0}^t ||Y_{X_N^{HS}(s)}^{HS}(1) - Y^{HS}(s+1)|| \xrightarrow{p} 0. \tag{4.80}$$

This implies

$$\begin{aligned} &\Pr\left[||X_N^{HS}(t) - Y^{HS}(t)|| > \epsilon\right] &(4.81)\\ = \;&\Pr\left[||\sum_{s=0}^t Y_{X_N^{HS}(s)}^{HS}(1) - Y^{HS}(s+1) + U_s^N|| > \epsilon\right] &(4.82)\\ \le \;&\Pr\left[\sum_{s=0}^t ||Y_{X_N^{HS}(s)}^{HS}(1) - Y^{HS}(s+1)|| + ||\sum_{s=0}^t U_s^N|| > \epsilon\right] \to 0\,\forall\epsilon>0. &(4.83)\end{aligned}$$

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

$$\Pr\left[X_N^{HS}(t+s) \in \tau_{\tilde{\mu}}^{-1}(Y_0) | X_N^{HS}(t) \in \tau_{\tilde{\mu}}^{-1}(Y)\right] > 0. \tag{4.84}$$

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 (4.84).

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.

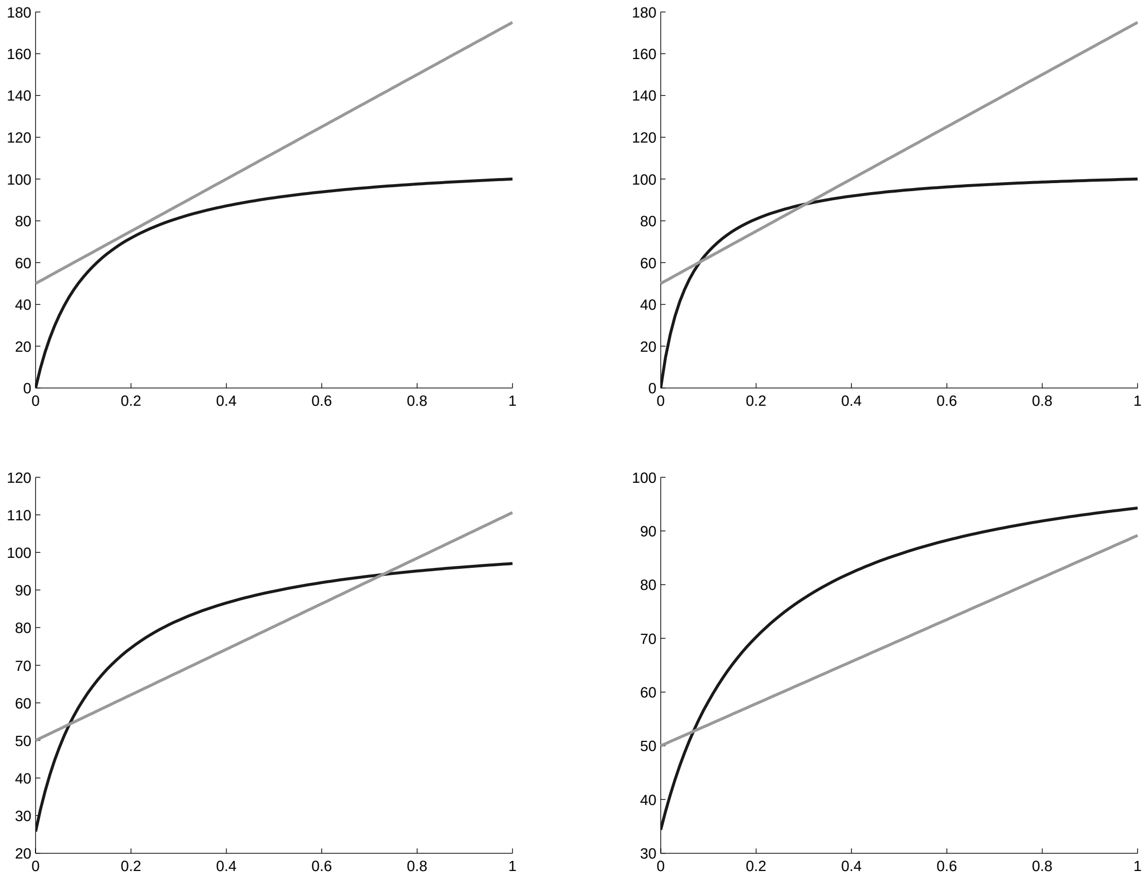

FIGURE I

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

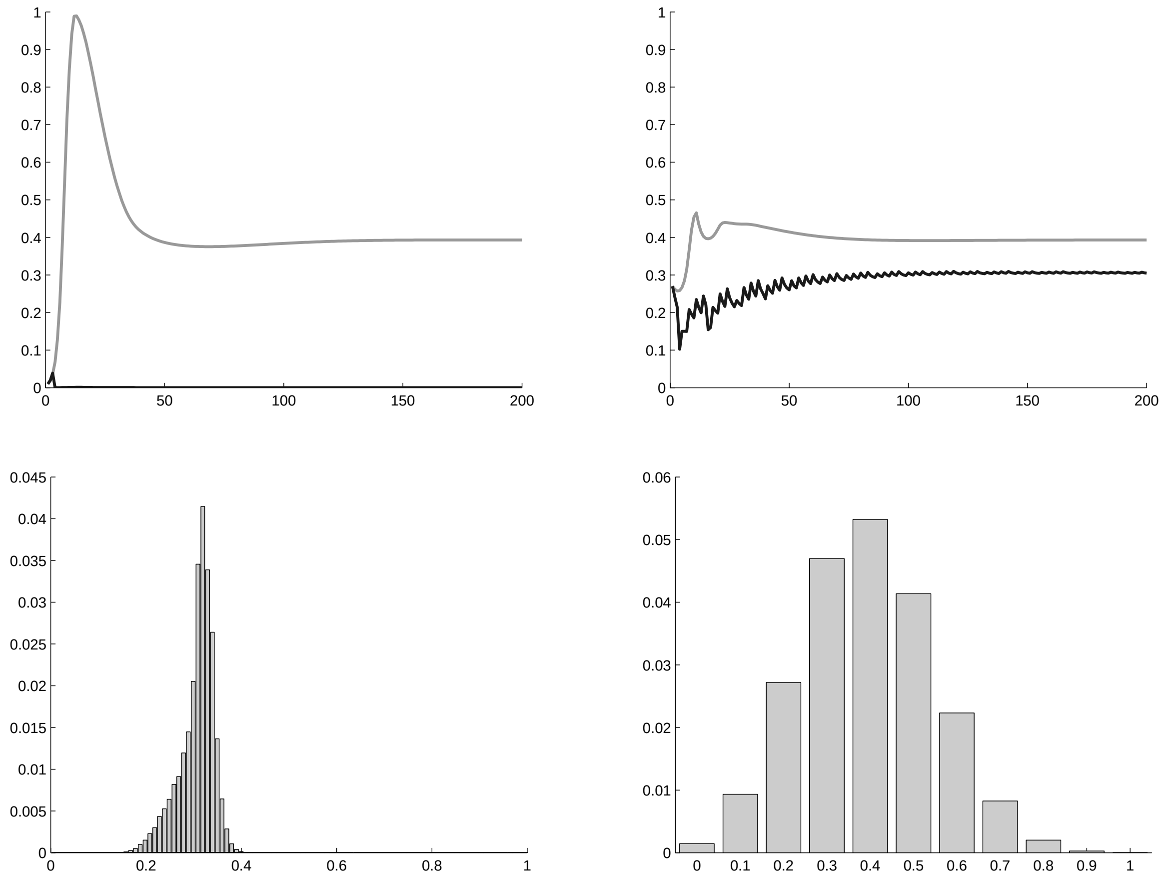

FIGURE II

[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.

Chapter 5

Literature

Aoyagi, M. and G. R. Fréchette (2003): *Collusion in Repeated Games with Imperfect Public Monitoring*, Harvard Business School, mimeo

Apesteguia, J., S. Huck and J. Oechssler (2007): *Imitation Theory and Experimental Evidence*, forthcoming J. Econ. Theory

Benaïm, M. and J. Weibull (2003): *Deterministic Approximation of Stochastic Evolution in Games*, Econometrica 71, 873 - 903

Bergstrom, T. C. (2002): *Evolution of Social Behavior: Individual and Group Selection*, J. Econ. Perspect. 16, 67 - 88

Boylan, R. (1995): *Continuous Approximation of Dynamical Systems with Randomly Matched Individuals*, J. Econ. Theory 66, 615 - 625

Brown, G. (1951): *Iterative solution of games by fictitious play*, in Koopmanns, T. C. (ed.): Activity Analysis of Production and Allocation, Wiley, New York

Brosig, J. (2002): *Identifying cooperative behavior: some experimental results in a prisoner's dilemma game*, J. Econ. Behav. Organ. 47, 275 - 290

Camerer, C. F. and T.-H. Ho (1999): *Experience-weighted attraction learning in normal-form games*, Econometrica 67, 827 - 874

Cartwright, E. (2003): *Learning to play approximate Nash equilibria in games with many players*, Warwick Economic Research Paper Series 671

Cartwright, E. (2007): *Imitation and the Emergence of Nash Equilibrium Play in Games with Many Players*, forthcoming Int. J. Game Theory

Casas-Arce, P. (2005): *Cooperation in Repeated Games with Dismissals and Quits*, mimeo, University of Oxford

Corradi, V. and R. Sarin (1999): *Continuous Approximation of Stochastic Evolutionary Game Dynamics*, forthcoming J. Econ. Theory

Dal Bó, P. and G. R. Fréchette (2007): *The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence*, Working Paper

Datta, S. (1993): *Building Trust*, mimeo, London School of Economics

Duffy, J. and J. Ochs (2003): *Cooperative Behavior and the Frequency of Social Interaction*, mimeo, University of Pittsburgh

Ellison, G. (1994): *Cooperation in the prisoner's Dilemma with Anonymous Random Matching*, Rev. Econ. Stud. 61, 567 - 588

Eshel, I., L. Samuelson und A. Shaked (1998): *Altruists, Egoists, and Hooligans in a Local Interaction Model*, Amer. Econ. Rev. 88, 157 - 179

Friedman, J. (1971): *A Noncooperative Equilibrium for Supergames*, Rev. Econ. Stud. 38, 1 - 12

Fudenberg, D. and J. Tirole (1991): *Game Theory*, MIT Press, Cambridge, Massachusetts

Fudenberg, D. and E. Maskin (1986): *The Folk Theorem in Repeated Games with Discounting or with Incomplete Information*, Econometrica 54, 533 - 554

Fudenberg, D. and E. Maskin (1991): *On the Dispensibility of Public Randomizations in Discounted Repeated Games*, J. Econ. Theory 53, 428 - 438

Fudenberg, D. and D. Kreps (1993): *Learning Mixed Equilibria*, Games Econ. Behav. 5, 320 - 367

Fudenberg, D. and D. K. Levine (1998): *The Theory of Learning in Games*, MIT Press, Cambridge, Massachusetts

Ghosh, P. and D. Ray (1996): *Cooperation in Community Interaction without Information Flows*, Rev. Econ. Stud. 63, 491 - 519

Hart, S. and A. Mas-Colell (2000): *A Simple Adaptive Procedure Leading to Correlated Equilibrium*, Econometrica 68, 1127 - 1150

Hendrich, J. (2004): *Cultural group selection, coevolutionary processes and large-scale cooperation*, J. Econ. Behav. Organ. 53, 3 - 35

Hopkins, E. (1999): *Learning, Matching, and Aggregation*, Games Econ. Behav. 26, 79 - 110

Kalai E. and E. Lehrer (1993): *Rational Learning Leads to Nash Equilibrium*, Econometrica 61, 1019 - 1045

Kandori, M., Mailath, G. and R. Rob (1993): *Learning, Mutation, and Long Run Equilibria in Games*, Econometrica 61, 29 - 56

Kandori, M. (1992): *Social Norms and Community Enforcement,* Rev. Econ. Stud. 59, 63 - 80

Kranton, R. E. (1996): *The Formation of Cooperative Relationships*, J. Law, Econ., Organ. 12, 214 - 233

Levine, D. and W. Pesendorfer (2007): *The Evolution of Cooperation Through Imitation*, Games Econ. Behav. 58, 293 - 315

Mailath, G. J. and L. Samuelson (2006): *Repeated Games and Reputations*, Oxford University Press, New York

Matsushima, H. (1990): *Long-term partnership in a repeated prisoner's dilemma with random matching*, Econ. Letters 34, 245 - 248

Murnighan, J. K. and A. E. Roth (1983): *Expecting Continued Play in Prisoner's Dilemma Games*, J. Conflict Resolution 27, 279 - 300

Offerman, T., Potters, J., and Sonnemans, J. (2002): *Imitation and Belief Learning in an Oligopoly Experiment*, Rev. Econ. Stud. 69, 973 - 997

Palomino, F. and F. Vega-Redondo (1999): *Convergence of aspirations and (partial) cooperation in the prisoner's dilemma*, Int. J. Game Theory 28, 465

- 488

Robson, A. and F. Vega-Redondo (1996): *Efficient Equilibrium Selection in Evolutionary Games with Random Matching*, J. Econ. Theory 70. 65 - 92

Roth, A. E. and J. K. Murnighan (1978): *Equilibrium Behavior and Repeated Play of the Prisoner's Dilemma*, J. Math. Psychology 17, 189 - 197

Roth, A. and I. Erev (1995): *Learning in Extensive Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term*, Games Econ. Behav. 8, 164 - 212

Schlag, K. (1998): *Why Imitate, and If So, How? A Boundedly Rational Approach to Multi-armed Bandits*, J. Econ. Theory 78, 130 - 56

Schlag, K. (1999): *Which One Should I Imitate?*, J. Math. Econ. 31, 493 - 522

Selten, R. and Ostmann, A. (2001): *Imitation Equilibrium*, Homo Oeconomicus 43, 111 - 149

Sobel, J. (1985): *A theory of credibility*, Rev. Econ. Stud. 52, 557 - 573

Sorin, S. (1986): *On Repeated Games with Complete Information*, Math. Operations Res. 11, 147 - 160

Stiglitz, J. E. and C. Shapiro (1984): *Equilibrium Unemployment as a Worker Discipline Device*, Amer. Econ. Rev. 74, 433 - 44

Watson, J. (1999): *Starting Small and Renegotiation*, J. Econ. Theory 85, 52 - 90

Watson, J. (2002): *Starting Small and Commitment*, Games Econ. Behav. 38, 176 - 199

Young, P. (1998): *Individual Strategy and Social Structure*, Princeton University Press, Princeton, New Jersey

Curriculum Vitae

Personal Details

Name	Heiner Schumacher
Born	23.02.1981 in Munich (Germany)

Education

Sep 2005 - Mar 2006	Affiliate Research Student, Department of Economics, University College London/United Kingdom (ENTER exchange)
Oct 2004 - Aug 2008	PhD Program in Economics, Center for Doctoral Studies in Economics (CDSE), University of Mannheim/Germany
Oct 2003 - present	Undergraduate Studies of Mathematics, Fern Universität Hagen/Germany
Apr 2002 - Jun 2004	Graduate Studies of Economics, Ludwig-Maximilians-University Munich/Germany, Diploma (M.Sc. equivalent)
Sep 2002 - Feb 2003	Graduate Studies of Economics, Carlos III de Madrid/Spain (ERASMUS exchange)
Oct 2000 - Mar 2002	Undergraduate Studies of Economics, Ludwig-Maximilians-University Munich, Vordiplom (B.A. equivalent)
Jun 2000	"Abitur", General Certificate of Education at Advanced Level, Munich

Eidesstattliche Erklärung

Hiermit erkläre ich, dass ich die Dissertation selbständig angefertigt und mich anderer als der in ihr angegebenen Hilfsmittel nicht bedient habe, insbesondere, dass aus anderen Schriften Entlehnungen, soweit sie in der Dissertation nicht ausdrücklich als solche gekennzeichnet und mit Quellenangaben versehen sind, nicht stattgefunden haben.

Mannheim, 5. Mai 2008

Heiner Schumacher