2012

Cooperative Strategies in Anonymous Economies: An Experiment

Gabriele Camera
Chapman University, camera@chapman.edu

Marco Casari
University of Bologna

Maria Bigoni
University of Bologna

Recommended Citation

Camera, G., M. Casari, and M. Bigoni (2012). Cooperative strategies in anonymous economies: An experiment. *Games and Economic Behavior* 75, 570–586. doi: 10.1016/j.geb.2012.02.009

Cooperative Strategies in Anonymous Economies: An Experiment

Comments
NOTICE: this is the author's version of a work that was accepted for publication in Games and Economic Behavior. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in *Games and Economic Behavior*, volume 75 (2012). DOI: 10.1016/j.geb.2012.02.009

Cooperative strategies in anonymous economies: an experiment

Gabriele Camera[a] | Marco Casari[b] | Maria Bigoni[b]

Purdue University | University of Bologna | University of Bologna

10 February 2012

Abstract

We study cooperation in economies of indefinite duration. Participants faced a sequence of prisoner's dilemmas with anonymous opponents. We identify and characterize the strategies employed at the individual level. We report that (i) grim trigger does not describe well individual play and there is wide heterogeneity in strategies; (ii) systematic defection does not crowd-out systematic cooperation; (iii) coordination on cooperative strategies does not improve with experience. We discuss alternative methodologies and implications for theory.

Keywords: Folk theorem, repeated games, equilibrium selection, finite automata, social dilemma, random matching.

JEL codes: C90, C70, D80

[a] Corresponding author: Dept. of Economics, 425 West State str., West Lafayette, IN 47907-2056, USA; e-mail: gcamera@purdue.edu; Tel. 765-494-8658.

[b] Piazza Scaravilli 2, 40126 Bologna, Italy; e-mails: marco.casari@unibo.it, maria.bigoni@unibo.it.

1. Introduction

Fostering cooperation in society can be problematic when agents who face a social dilemma do not know each other and cannot easily establish reputations. The individual appeal of opportunistic behavior is especially strong when it is difficult to communicate intentions, to maintain stable partnerships, or to monitor and to enforce cooperation of others. Yet, Folk Theorem-type results suggest that, over the long haul, none of these frictions present a fundamental obstacle to cooperation.[1] Groups of self-regarding agents can overcome the short-run temptation to cheat others by employing a social norm threatening permanent defection through a decentralized punishment scheme that spreads by contagion. Studying these dynamics is central because, as Binmore (2005, p.818) points out, "the ideas that motivate the folk theorem of repeated game theory remain our best hope of understanding how societies hold together and adapt to new challenges". The open question is how, in practice, groups of agents reach cooperation when theoretically feasible and what strategies they adopt to sustain it.[2]

We address this question through an experiment where subjects faced an indefinitely repeated prisoner's dilemma implemented through a random stopping rule (e.g., as in Palfrey and Rosenthal, 1994, Dal Bó, 2005). Each subject was assigned to a group of four. In each period, each group was randomly divided into two pairs that played the prisoner's dilemma. Subjects could not identify opponents and could only observe outcomes in their own pair. Each subject played a sequence of five supergames. In each supergame every subject was assigned to a different group formed according to a perfect stranger protocol.

[1] The foundation for this statement traces back to the Folk Theorem in Friedman (1971) and the random-matching extensions in Kandori (1992) and Ellison (1994).

[2] The relevance of such a question has been recently emphasized by Ostrom (2010, p. 660), who writes that "Simply assuming that humans adopt norms […] is not sufficient to predict behavior in a social dilemma, especially in very large groups with no arrangements for communication."

Our design makes it possible to empirically identify and characterize strategies employed by subjects. The empirical evidence about strategies adopted in indefinitely repeated games is still limited, and it is confined to two-person economies of short-duration with a subject pool of undergraduates (Engle-Warnick et al., 2004, Engle-Warnick and Slonim, 2006, Aoyagi and Fréchette, 2009, Dal Bó and Fréchette, 2011, Fudenberg et al., forthcoming). This study advances the understanding of how subjects play indefinitely repeated games by studying their behavior in four-person economies of substantially longer duration than in the literature, and with a varied subject pool (undergraduates, MBA students, white-collar workers).

We study economies where anonymous subjects are not in stable partnerships, which are the theoretical platform of an important segment of the economic literature. Consider for instance the matching models in macroeconomics as in Diamond (1982) or in Mortensen and Pissarides (1994), and the decentralized trading models in microeconomics as in Kandori (1992) and Milgrom, North, and Weingast (1990). A key feature of these models and of our experiment is that it is impossible for a subject to build a reputation. Clearly, no model can perfectly fit situations in the field and a similar argument generally applies to experiments. In our laboratory economies we control the informational flows and the matching process to capture essential "trade frictions" characteristic of larger economies. In particular: subjects may not know and may not trust each other, reputation is hard to establish, it is difficult or costly to monitor the actions of all other members of society, to communicate intentions, and institutions for enforcement have limitations.

Folk theorem-type results show that a long-run interaction can sustain a multiplicity of equilibrium outcomes but do not offer much guidance regarding which equilibrium will be selected. We report that full efficiency is rarely achieved in our experimental economies, which suggests that efficiency may not be the key equilibrium selection criterion (see also Duffy and Ochs, 2009, and Duffy et al., forthcoming). Such result contrasts with the assumption in many applications of theories of infinitely repeated games. Moreover, in these theories the efficiency frontier is often traced through a "grim trigger" strategy whereby *all* players cooperate under the threat of a contagious process of economy-wide defection.[3] We carry out an individual-level analysis that sheds light on empirically-relevant equilibrium selection criteria. We find that only one out of four individuals behaves in a manner consistent with the use of the grim trigger strategy. These findings challenge the descriptive power of theories based on the notion that everyone cooperates because of the threat of unforgiving, generalized punishment. The data suggest that subjects dealt with the heterogeneity in behavior by tolerating some defections in attempting to coordinate with a subset of participants in the economy. We observe that systematic defectors and systematic cooperators coexisted within most economies. Subjects tried out a variety of strategies but these attempts at coordinating failed to sustain high cooperation rates.

The paper proceeds as follows: Section 2 discusses related works; Section 3 presents the experimental design; Section 4 provides a theoretical analysis; Section 5 proposes an empirical procedure for classification of individual strategies and the main results are reported in Section 6; Section 7 presents results for a maximum likelihood analysis of strategies; Section 8 expands the analyses by considering an extended strategy set, and Section 9 concludes.

2. Related experimental literature

There are a few experimental studies of the strategies played by subjects in indefinitely repeated

[3] This implies that to punish a single defection, innocent cooperators must also be penalized. In this sense, our work bears similarities to experiments with public good games and common pool resources (e.g., Ostrom et al., 1992, Fischbacher et al., 2001), where cooperation is not individually rational in the stage game but can be sustained in the indefinitely repeated game if the grim strategy is adopted.

games. With one exception, discussed below, all studies we are aware of refer to two-person economies. The key differences in our set-up are the following: first, economies include four persons; second, subjects do not interact as partners; third, we consider long-duration economies. We now present an overview of the papers most related to our study.

Engle-Warnick, McCausland, and Miller (2004) retrieve subjects' strategies from experimental data on an indefinitely repeated prisoner's dilemma. They model behavior using finite automata, as we do, while implementing a different experimental design and empirical technique. Their subjects interacted in a supergame as partners for 5 periods in expectation, while our subjects interacted as strangers for 20 periods in expectation. Unlike our study, they employ Bayesian methods and numerical techniques to estimate the distribution of strategies represented by automata with errors on actions.

Engle-Warnick and Slonim (2006) study an infinitely repeated trust game with 5 periods of average duration. They empirically identify the strategies employed by subjects by formalizing strategies using the notion of finite automata (Rubinstein, 1986), as we do. They consider a large number of automata, which exclude the possibility of errors in the implementation of strategies. They find that the vast majority of data can be explained using only a small number of strategies. In particular they find support for grim trigger play between partners.

Aoyagi and Fréchette (2009) study an indefinitely repeated prisoners' dilemma with imperfectly observable actions, and 10 periods of average duration. They consider a family of threshold strategies where transitions between cooperative and punishment state depend on four free parameters. They find support for the use of forgiving strategies, rather than grim trigger.

Dal Bó and Fréchette (2011) study an indefinitely repeated prisoner dilemma with an expected duration of two or four periods. They estimate subjects' strategies fitting the data via a maximum likelihood approach, to a set of six possible strategies. They look at the behaviour of experienced subjects. They find support for "tit for tat" and "always defect", but unlike Engle-Warnick and Slonim (2006), they do not find evidence for grim trigger strategies in their economies of two players and of very short duration.

Fudenberg et al. (forthcoming) consider a design adapted from Dal Bó and Fréchette (2011) where actions are observed with noise and there are longer sequences of play. They fit the data using a set of twenty automata with up to four states following a maximum likelihood approach. The most common strategies are "tit for two tats" and strategies that trigger permanent punishment after the second or the third consecutive defection.

Stahl (2009) studies the impact of reputation mechanisms in finitely repeated prisoner's dilemmas with random matching. In the absence of a reputation mechanism, subjects did not sustain cooperation in economies of 22-24 members. With a color-coded reputation mechanism in place, cooperation was sustained through the use of strategies based on individual reputation, which is in line with the evidence in Camera and Casari (2009) for indefinitely repeated games in smaller economies. In contrast, the present study identifies strategies in economies where reputation mechanisms are ruled out by design and it focuses on individual behavior, which is not studied in Camera and Casari (2009).

3. Experimental design

The experiment is based on the same design of the Private Monitoring treatment in Camera and Casari (2009), which is suitable to study strategy selection in an indefinitely repeated prisoner dilemma where reputation formation is impossible. An economy comprises four persons who interact *privately* and *anonymously*. The interaction is private because subjects observe only outcomes in their pair but not in the rest of the economy (private monitoring). Subjects are not in a stable partnership but are randomly matched in pairs after every encounter. The interaction is anonymous because subjects cannot observe identities.

The underlying game is the prisoners' dilemma described in Table 1. In the experiment, subjects could choose between *Y*, for cooperation, and *Z* for defection. Because of the empirical difficulty in supporting high levels of cooperation in economies of strangers, the parameters of the experiment are calibrated to promote some cooperative choices. This is necessary to study and draw conclusions about the type of strategies people adopt to support cooperative outcomes.[4]

[Table 1 approximately here]

A supergame (or *cycle*, as it was called in the experiment) consists of an indefinite interaction among subjects achieved by a random continuation rule; see Roth and Murninghan (1978). The interaction is of finite but uncertain duration, because in each period a cycle continues with a constant probability δ=0.95. For a risk-neutral subject δ represents the discount factor. In each period the cycle is expected to continue for 19 additional periods. To implement this random stopping rule, at the end of each period the program drew a random integer between 1 and 100, using a uniform distribution. The cycle continued with a draw of 95 or below. All session participants observed the same random draw, which means that cycles for all economies terminated simultaneously.

Each experimental session involved twenty subjects and five cycles. We built twenty-five economies in each session by creating five groups of four subjects in each of the five cycles. Matching across cycles followed a perfect stranger protocol: in each cycle each economy included only subjects who had neither been part of the same economy in previous cycles nor

[4] We selected this parameterization as it scores high on the indexes proposed by Rapoport and Chammah (1965) and Roth and Murnighan (1978) that correlate with the level of cooperation in the indefinitely repeated prisoners' dilemma in a partner protocol.

were part of the same economy in future cycles. Subjects did not know how groups were created but were informed that no two participants ever interacted together for more than one cycle. This matching protocol across supergames reduces the possibility of contagion effects, as opposed to a stranger protocol. In short, it is as if each subject had five distinct "lives" in a session.

Participants in an economy interacted in pairs according to the following matching protocol within a supergame. At the beginning of each period of a cycle, the economy was randomly divided into two pairs. There are three ways to pair the four subjects and each one was equally likely. So, a subject had one third probability of meeting any other subject in each period of a cycle. For the whole duration of a cycle a subject interacted exclusively with the members of her economy. In each economy, subjects interacted locally in the sense that they could only observe outcomes in their pair. In addition, they could neither observe identities of opponents, nor communicate with each other, nor observe histories of others. As a consequence, subjects did not share a common history. With this private monitoring design, the efficient outcome can be supported as an equilibrium.

Studying strategy adoption within this experimental design offers several advantages in comparison to designs based on two-person economies. First, anonymity implies that strategies based on reputation cannot be employed. In empirical studies, these strategies show a strong drawing power, although they are not theoretically essential to sustain cooperation.[5] Hence, with anonymity subjects are forced to consider other strategies. Second, in economies with more than two participants coordination on strategies and outcomes is more interesting and challenging. For

[5] Our design adopts private monitoring and random matching among subjects as in the Private Monitoring treatment of Camera and Casari (2009), where subjects' behavior substantially differed from behavior in a non-anonymous public monitoring treatment. With non-anonymous public monitoring subjects mostly defected with opponents who previously defected with her. Stahl (2009) reports that cooperation could be sustained in a finitely repeated prisoner's dilemma game only when a color-coded reputation mechanism was introduced.

instance it is possible to study if and how a subset of subjects can separately coordinate on a strategy. Third, subjects are exposed to a variety of behaviors, which facilitates the empirical identification of strategies.

The experiment involved three distinct groups of subjects: 40 undergraduate students from various disciplines at Purdue University, 20 full-time MBA students in the Krannert School of Management, and 40 clerical workers employed as staff throughout Purdue University. Both MBAs and undergraduates have a strong international component. The clerical workers are mostly long-time state residents, who exhibit a wide variation in age and educational backgrounds. Having multiple subject pools is methodologically appealing because it enhances the external validity of our results.[6]

All 100 subjects were recruited through e-mail and in-class-announcements. The sessions were run in the Vernon Smith Experimental Economics Lab. No eye contact was possible among subjects. Instructions were read aloud with copies on all desks. A copy of the instructions is in the Appendix. Average earnings were $18 excluding show-up fees. A session lasted on average 79 periods for a running time of about 2 hours, including instruction reading and a quiz. Each session had 20 participants and 5 cycles.[7]

4. Theoretical predictions

The theoretical predictions are based on the works in Kandori (1992) and Ellison (1994), under the assumption of identical players, who are self-regarding and risk-neutral. Here we concisely

[6] Differences across subject pools are studied in the companion paper Bigoni, Camera and Casari (2012).

[7] Sessions took place on the following dates: 21.4.05 (71), 7.9.05 (104), 29.11.05 (80), 06.12.05 (50), 07.02.07 (91). The total number of periods for the session is in parenthesis. Show-up fees are as follows: undergraduates received $5; clerical workers received $10; MBAs received $20. Data of the first two sessions are also analyzed in Camera and Casari (2009). The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007).

present the relevant theoretical predictions; for additional details see Camera and Casari (2009). The stage game is the prisoner dilemma in Table 1. Players simultaneously and independently select an action from the set $\{Y,Z\}$. Total surplus in the economy is maximized when everyone cooperates, i.e., when all players choose Y. Thus, we refer to the outcome where every player in the economy selects Y as the *efficient* or fully cooperative outcome. If both pairs in the economy select $\{Z,Z\}$, then we say that the outcome is *inefficient*. There exists a unique Nash equilibrium where both agents defect and earn 10 points.

Under private monitoring, indefinite repetition of the stage game with randomly selected opponents can expand the set of equilibrium outcomes. Following the work in Kandori (1992) and Ellison (1994), we present sufficient conditions so that the equilibrium set includes the efficient outcome, which is achieved when everyone cooperates in every match and all periods.

The inefficient outcome can be supported as a sequential equilibrium using the strategy "always defect." Since repeated play does not decrease the set of equilibrium payoffs, Z is always a best response to play of Z by any randomly chosen opponent. In this case the payoff in the indefinitely repeated game is the present discounted value of the minmax payoff, $z/(1-\delta)$. If δ is sufficiently high, then the efficient outcome can be sustained as a sequential equilibrium, by threatening to trigger a contagious process of defection, leading to minmax forever. In an economy with full cooperation, every player receives payoff $y/(1-\delta)$. Hence, the main theoretical consideration is the following:

Let $\delta^* \in (0,1)$ be the unique value of δ that satisfies

$$\delta^2(h-z) + \delta(2h-y-z) - 3(h-y) = 0 .$$

If $\delta \geq \delta^*$, then the efficient outcome is a sequential equilibrium. In the experiment, the efficient outcome can be sustained as an equilibrium, because δ=0.95 and δ^*=0.443.

We now provide intuition for the above statement. Conjecture that players behave according to actions prescribed by a social norm. A social norm is a rule of behavior that identifies desirable play and a sanction to be selected if a departure from the desirable action is observed. We identify the desirable action by Y and the sanction by Z. Thus, every player must cooperate as long as she has never played Z or has seen anyone select Z. However, if the player observes Z, then she must select Z forever after. This is known as a *grim trigger strategy*.

Given this social norm, in equilibrium everyone cooperates so the payoff to everyone is the present discounted value of y forever: $y/(1-\delta)$. A complication arises when a player might consider defecting, however, as defection always grants a higher payoff in the stage game. To deter players from behaving opportunistically, the social norm employs the threat of contagious process of defection leading to minmax forever. Notice that a player deviates in several instances—first, in equilibrium, if she has not observed play of Z in the past but chooses Z currently, and second, off-equilibrium, if she has observed play of Z in the past but plays Y currently. Cooperating when no defection has been observed is optimal only if the agent is sufficiently patient. The future reward from cooperating today must be greater than the extra utility generated by defecting today (unimprovability criterion). Instead, if a defection occurs and everyone follows the social norm, then everyone ends up defecting since the initial defection will spread by contagion. Given that our experimental economies have only four players, contagion can occur very quickly.

For a strategy to be an equilibrium strategy, cooperating after observing a defection should also be suboptimal. Choosing Y can delay the contagion but cannot stop it. To see why, suppose a player observes Z. If she meets a cooperator in the next period, then choosing Y produces a current loss to the player because she earns y (instead of h). If she meets a deviator, choosing Y

also causes a current loss because she earns *l* rather than *z*. Hence, the player must be sufficiently impatient to prefer play of *Z* to *Y*. The smaller are *l* and *y*, the greater is the incentive to play *Z*. Our parameterization ensures this incentive exists for all $\delta \in (0,1)$ so it is optimal to play *Z* after observing (or selecting) *Z*.[8]

Two remarks are in order. First, due to private monitoring, T-periods punishment strategies *cannot* support the efficient outcome as an equilibrium. Suppose a pair of agents starts to punish for T periods, following a defection in the pair. Due to random encounters, this initial defection will spread at random throughout the economy. Hence, over time different agents in the economy will be at different stages of their T-periods punishment strategy. Hence, agents cannot simultaneously revert to cooperation after T periods have elapsed from the initial defection.[9] Second, cooperation is risk-dominant in our design, in the following sense. Consider two strategies, "always defect" against "grim trigger." Grim trigger is risk-dominant if a player is at least indifferent to selecting it, given that everyone else is believed to select each strategy with equal probability. Indifference requires $\delta = 0.763$.[10]

5. Estimation procedure for individual strategies

To empirically identify in the experimental data the strategies employed by each subject, we formalize strategies using the concept of finite automata. As a robustness check, we consider both deterministic automata as well as automata with a random element.

An automaton is a convenient way to represent the process by which a player implements a

[8] With our parameterization, the upper bound to δ is 1.125 and the lower bound is 0.443.
[9] In every period of the session, all participants observed the same random integer 1,…,100, which could have served as a public randomization device. Hence, the efficient outcome could be sustained also by strategies based on contagious punishment that exploit the availability of a public randomization device; see Ellison (1994).
[10] Details on derivations are available upon request. See Blonski and Spagnolo (2001) for an application to infinitely repeated games among partners. Blonski, Ockenfels, and Spagnolo (2011) and Dal Bó and Fréchette (2010) present experimental evidence on how risk-dominance impact partners' play in indefinitely repeated games.

rule of behavior in a repeated game (Rubinstein, 1986). The automaton is described by (i) a set of actions, (ii) a set of states for the player, (iii) an outcome function that specifies the action to be taken given the player's state, and (iv) a transition function that specifies the next state that the player will reach, given his current state and the actions taken by the opponent.

Automata with sufficiently many states can describe *any* type of behavior observed in the experiment. We consider only two-state automata. There are various reasons for doing so. This class of automata is small—there are only 2^5=32 two-state automata—and yet it allows to represent most common strategies in the literature, such as "tit-for-tat," "grim trigger," "always defect," and "always cooperate."[11] Clearly, not all of these automata describe equilibrium strategies. Moreover, two-state automata describe strategies that are relatively simple, hence likely to be devised and used by experimental subjects. As an example Figure 1a illustrates "tit for tat" and "grim trigger." Actions are either C=Cooperate or D=Defect; a circle corresponds to a state for the player, where the initial state is a bold circle; the outcome function is the identity function, i.e. the unique action prescribed is written inside each circle; the solid arrows represent transitions between states, which depend on the opponent's action reported next to each arrow.

As seen above, an automaton defines a deterministic action plan, which provides a rigid rule to capture subject's behavior. When fitting the data, we relax the rigidity of the rules of behavior by introducing a random element in the automata. This can accommodate subjects who make some mistakes in implementing a plan or who pursue some intentional experimentation within their strategy. The estimation procedure allows for random transitions, i.e., the possibility to reach the

[11] There are 2 initial states identified by the action prescribed in that state, C and D, and 2 subsequent states, C or D, that are reached depending on the 4 possible outcomes of the match (each player has two actions). See Table 2. We use automata to represent subjects' play because it is an empirically helpful technique to classify subjects' behavior. Clearly, such a representation does not restrict in any way subjects' freedom of choice during the experiment. Automata are simply a convenient tool to characterize subjects' observed patterns of choice. Expanding the set of automata considered, for instance by including three- or four-state automata, would simply increase the number of classified individuals and would not alter the equilibrium set.

incorrect state with some probability $p \geq 0$.[12] We estimate strategy fitting from a range of values for p from 0 through 0.40. This allows us to assess how the explanatory power of a given automaton varies as we increase the probability of errors. With two states, departures from a plan take one of two forms: the subject may either fail to switch state (say, keeps playing C instead of switching to D) or may incorrectly switch state (say, plays D instead of keep playing C). The dashed lines in Figure 1b represent such incorrect or accidental transitions for the case of grim trigger and tit-for-tat. Randomness on transitions is different from randomness on outcome functions, as in Engle-Warnick et al. (2004).

[Figure 1 approximately here]

We group the 32 strategies considered into six *strategy sets* (Table 2). The *initial action* is C for four strategy sets and is D for two sets. An additional distinction is whether play is *unconditional* or *conditional* on the observed outcome. Unconditional strategies prescribe only one action unless mistakes are made. Such strategies comprise the classes of automata called *systematically cooperate* and *systematically defect*, which include as a special case "always cooperate" and "always defect" (see the note to Table 2 for more details). Conditional strategies starting with C are divided into three groups: grim trigger, a set of *forgiving* strategies, and a set of *unconventional* strategies. Forgiving strategies prescribe a switch to playing D only if an opponent chooses D, but allow a switch back to C (e.g., "tit for tat"). Unconventional strategies, instead, may prescribe D even if no defection has been observed.

[Table 2 approximately here]

The strategy-fitting procedure is a mapping from the experimental data into the strategy sets

[12] Note that we are interested in social norms and especially the grim trigger strategy, because it sustains cooperative equilibria. If everyone plays grim trigger, one mistake moves the economy to permanent punishment. Errors in transitions allow us to study cases where players may reconsider the wisdom of carrying out extreme punishments or wish to give a second chance to cooperation.

of Table 2. The unit of observation is the sequence of all choices of a subject in a cycle, i.e., the behaviour of an individual or, simply, an *individual*. We may also refer to such a sequence as one *observation*. Hence, one subject contributes five observations in the dataset. For every individual, we first select the strategy that best describes ("fits") her sequence of actions among the thirty-two strategies available. Then, we check whether the description of behavior provided by this best-fitting strategy is sufficiently accurate. If it is so, then we classify the individual by that strategy; otherwise, we say that the individual is unclassified by that strategy. Note that one individual could be classified by more than one strategy. Those who cannot be classified by *any* strategy are denoted unclassified individuals.

We say that strategy q "fits" an observation (i.e., an individual) if it can generate an action sequence *consistent* with the behavior of the subject in the cycle. The definition of consistency allows for some experimentation or occasional mistakes. More precisely, let $x_{q,t}=1$ if a subject's action in period t of a cycle corresponds to the outcome generated by a correct implementation of strategy q, and let $X_q(T)=\sum_{\tau=1}^{T} x_{q,\tau}/T$ denote the *consistency score* of that strategy, in a cycle of duration T. The score ranges from zero (no action taken is consistent with strategy q) to one (correct implementation of q).[13] To account for the possibility that subjects may occasionally depart from the chosen plan of action, we presume a probability p of an incorrect transition exists that is (i) identical across subjects, (ii) constant across periods and cycles, and (iii) independent of the strategy considered. Under these conditions, the number n of a subject's actions that are inconsistent with a strategy q in a cycle of duration T is distributed according to a binomial with

[13] For example let q be "grim trigger" and suppose a subject observes D only in period 1 of a four-period cycle. The sequence CDDD generates the score X_q=4/4=1. With random transitions, however, the sequence CDCC would generate X_q=3/4 because only the action in period 3 is inconsistent with grim trigger. An incorrect transition occurs in period 2 (from state D to C, whereas D should be an absorbing state), but the action in period 4 is consistent with being (incorrectly) in state C.

parameters p and T-1. The expected number of inconsistent actions increases with T and decreases with p so that if p and T are sufficiently small the expected number of inconsistent action is lower than one. Hence, the average length of a cycle is a crucial design parameter.

Fixing p, we say that strategy *q fits an observation* or, equivalently, that one individual is *classified* according to strategy q, if the following three conditions are satisfied. First, q correctly predicts the initial action, $x_{q,1} = 1$. This is because errors in transition can occur only across periods; hence, an error can be made only after period 1. Second, q must have the largest consistency score among all strategies considered, $X_q(T) \geq X_{q'}(T)$ for all $q' \neq q$. Finally, if n actions are inconsistent with q, then the probability of such a realization must be within chance, given p and T. As a statistical test, strategy q does not fit the observation if the observation lays in the 10% right tail of the distribution of errors, i.e., the strategy does not fit the observation if the probability of observing n or more inconsistent actions is smaller than 10%. To fix ideas, suppose p=0.05. According to our criterion, not even one inconsistent action is admissible in cycles lasting less than four periods. In a cycle lasting 20 periods, instead, we expect one incorrect transition and admit at most two incorrect transitions; this means that, for example, a "grim trigger" player who has started punishing has the chance to move back to a cooperative state and to retrace his steps back to full defection, and yet to be classified as "grim trigger," according to this third condition. If one or more of the above conditions is not met, then the observation is "unclassified."

Definition: *The total fit N(q) of a strategy q is the number of observations that q fits. The total fit N(Q) of a strategy set Q is the number of observations that can be explained by at least one strategy* $q \in Q$.

Both the total fit of a single strategy and the total fit of a strategy set are useful measures because individuals could sometimes be classified by more than one strategy. For this reason, the total fit of a strategy q provides an upper bound for the number of individuals that employ strategy q.[14] . This problem of classification overlap is more relevant for subjects who did not experience a sufficient variety of actions in a cycle or played a short cycle. For instance, in a two-period cycle "grim trigger" and "tit-for-tat" identically fit the observation C in period 1 and D in period 2 when the initial opponent plays D, and also fit the observation CC when the initial opponent plays C. When a subject observes the same action (e.g., C) in every meeting, we cannot infer what the subject would have done if D was observed. As a consequence, if one sums up the total fit of single strategies in a set Q, $\sum_{q\in Q} N(q)$, then the figure may be greater than the total fit $N(Q)$ of the set Q, and can even exceed the total number of observations.

To obtain a tighter classification of individuals, the strategy-fitting procedure is refined as follows. First, strategy sets are constructed in order to include closely-related strategies. It is thus possible that a single individual is classified by two strategies if they are in the same set, but less likely if they belong to different sets. Second, the strategy fitting procedure has been run separately for individuals who observed heterogeneous actions, i.e. for the subsample of the data where opponents played both C and D. Tracing the response to out of equilibrium behavior improves the chances for unique identification. Third, the estimation on this subsample is also carried out for deterministic strategies. This further reduces classification overlap; for instance, an individual cannot be identified both by "grim trigger" and by strategies in the set "systematically cooperate".[15]

[14] This measure is context-independent, i.e., it is invariant to the number and type of strategies considered.
[15] More specifically, we find the following among all classified individuals: first, the average individual is classified by 4.51 strategies (1827/405, Table 2) and by 1.70 strategy sets (688/405, Table 3). Second, focusing on strategy

6. Results

There are seven main results. Result 1 concerns an analysis of strategy play at the aggregate level. Results 2-7 are about individuals, i.e., strategies employed by single subjects in each cycle.

Result 1. *Consider subjects' behavior at the aggregate level. In period 1 subjects exhibited a high cooperation rate. If in the cycle subjects observed a defection, then they persistently lowered her cooperation rate.*

This finding is broadly consistent with the theories in Kandori (1992) and Ellison (1994) regarding the existence of a rich equilibrium set, including full cooperation, under private monitoring.

Choices in the first period of each economy help us determine whether some equilibrium (among the many possible) had a particularly strong drawing power. Average cooperation level in period 1 was 67.2%, and in all periods it was 53.8%. Hence, we can rule out that subjects attempted to coordinate on defection (see Table 4 for cooperation rates disaggregated by cycle and for period 1 in each cycle). What behavior can explain such patterns of cooperation? Due to private monitoring, cooperation cannot be supported through T-period trigger strategies. In contrast, grim trigger can theoretically sustain an equilibrium with 100% cooperation. To investigate whether the data are consistent with such strategies, we ran a probit regression that explains subjects' choice to cooperate (1) or not (0) using two groups of regressors. In this regression, the unit of observation is a subject's choice in a period. We introduce dummy variables that control for fixed effect (cycles, periods within the cycle, subject), as well as for the

sets, the average individual is classified by 1.43 strategy sets when her opponents play both C and D (454/317, Table 3). Third, this figure further reduces to 1.28 with deterministic automata (234/182, Table 3, p=0).

duration of the previous cycle. To trace the response of subjects in the periods following an observed defection we include a “grim trigger” regressor that has value 1 in all periods following an observed defection (0 otherwise). We also include five “lag” regressors that have value 1 only in one period following an observed defection (0 otherwise). More specifically, the “lag *n*” regressor takes value 1 after one, two, three, four or five periods after the observed defection (*n*=1,2,3,4,5, respectively) and 0 in all other periods. If subjects switched from a cooperative to a punishment mode after seeing a defection, then the estimated coefficient of at least one of the six strategy regressors should be negative. For example, if subjects punished for just two periods after a defection, then the sum of the estimated coefficients of grim trigger and “lag *n*” regressors should be negative for the first two periods after a defection (0 afterwards).

The key result from this analysis is: the defection of an opponent triggered a persistent decrease in cooperation with very little reversion to a cooperative mode. Figure 2 provides supporting evidence; it illustrates the marginal effect of experiencing a defection on the frequency of cooperation in the following periods.[16] The marginal effect curves are L-shaped, i.e., after an initial drop, the curves look generally flat, and no recovery to pre-defection cooperation levels after five periods can be detected.[17] Instead, if subjects adopted a strategy that allows for reversion to full cooperation, then curves should be U-shaped.[18]

[Figure 2 approximately here]

Cooperation was the focal point of period 1 play for subjects. When first confronted with a defection in the match, a substantial fraction of initial cooperators responded with an immediate,

---

[16] The representation for “any more than five” period lags is based on the marginal effect of the grim trigger regressor only. The representation for period lags 1 though 5 is based on the sum of the marginal effects of the grim trigger regressor and the “lag *n*” regressors with the appropriate lag.
[17] The marginal effects of “lag 1” and “lag 5” regressors are not significantly different (p-value: 0.1144).
[18] Additional details on supporting evidence, including regression results, are in the supplementary appendix.

downward and persistent shift in the frequency of cooperation. Similar analyses are carried out by Camera and Casari (2009) both for anonymous and non-anonymous economies. Depending on the anonymity level, such Probit regression analysis revealed substantial differences in the empirical choice of strategies. The present study enhances the validity of the earlier results on anonymous economies, by using a larger and more diverse subject pool.

The above analysis is compatible with a fraction of subjects acting as if playing "Grim trigger" and others playing "Always defect" or "Always cooperate." However, this is not the only pattern of strategies that could possibly generate this result. Therefore, to expand on this initial assessment we carried out a statistical analysis of strategies adopted at the individual level, proceeding as follows. First, we empirically identify strategies and determine the number of observations that can be explained by those strategies. Second, we empirically characterize the strategies most commonly used. Third, we analyze the dynamics of behavior by subject to understand whether subjects learn to coordinate on certain outcomes and cooperative strategies.

As discussed in the previous Section, the unit of observation is the sequence of all choices of a subject in a cycle, i.e., the *individual*. Since there are 100 subjects and five cycles, there are 500 individuals. As an initial step, we wish to determine (i) whether the 32 simple strategies considered classify a high or small number of individuals and (ii) which strategies are most successful in doing so. A central result is that, given the identification technique proposed in Section 5, a high number of individuals can be classified.

Result 2. *Consider the behavior of single individuals. When allowing for limited randomness in behavior (p=5%), thirty-two simple strategies classify 81% of individuals.*

The empirical findings are reported starting with Figure 3, illustrating the percentage of classified individuals as we vary the probability p of incorrect transition (i.e., the experimentation rate) from 0 to 0.40. Varying the probability *p* serves as a robustness check. Figure 3 shows the marginal gain in total fit as one changes the probability of incorrect transition. Fully deterministic automata (*p*=0) classify more than half of the individuals. The total fit is 53.0% of individuals. If we increase the probability of incorrect transitions to *p*=0.05, then the total fit of the entire strategy set improves substantially, reaching 81.0%. The fit then slowly tapers out. With *p*=0.30 we classify almost 100% of individuals. Therefore, in the analysis that follows, we will report results for *p*=0.05, unless otherwise stated, and will include detailed results in Tables 3-4.

[Figure 3 approximately here]

No single group of strategies classifies a majority of individuals. "Systematically cooperate," which is the most relevant group, classifies 26.8% of individuals when *p*=0 and 42.4% when *p*=0.05 (Table 3). Considering *p*=0.05 (Table 2), the best-fitting single strategy is one of those in the "systematically cooperate" class and it classifies 37.6% of individuals. When taking just two strategies into account (11100 and 00000), the total fit is 59.8%.[19] When taking into account differences in cycle duration, these figures are in line with the results reported in other studies. In an indefinitely repeated trust game, Engle-Warnick and Slonim (2006) achieve a total fit of 89.6% when they employ 32 strategies. When taking just two strategies into account they fit 66.8%. However, the average length of a cycle in Engle-Warnick and Slonim (2006) was 5.1 periods, considerably shorter than in our experiment, which matters for comparison purposes because statistically a strategy has more difficulty in fitting behavior emerging from longer cycles. Longer cycles allow a better identification of strategies, so as cycle duration increases a larger strategy set is needed to fit a given number of observations. To see why, consider that a set

[19] 00000 corresponds to unconditional defection; 11100 describes the behavior of someone who always cooperates but may switch to permanent defection by mistake.

of just two strategies such as "grim trigger" and "always defect" fits 100% of observations of one-period cycles. This explains why it is more difficult to classify individuals who played longer cycles; unclassified individuals played cycles lasting 25.8 periods on average, as opposed to 13.6 periods for classified individuals. The key issue is thus to determine under what dimensions unclassified and classified individuals differ.[20]

[Table 3 approximately here]

Result 3. *Classified individuals exhibited higher average payoffs than unclassified individuals.*

Support for Result 3 is in Table 3. Mean profits are significantly greater for classified than unclassified individuals (18.7 vs. 15.2; p-value is 0.061 when controlling for cycle length and 0.014 without it; N_1=96, N_2=406). This suggests, although it does not prove, that the two-state automata considered include the best-performing strategies.

As it may be expected, classified individuals exhibited lower volatility of play than unclassified individuals. Volatility of play is defined as the *frequency of switch* between cooperation and defection choices. The average switch frequencies for classified individuals are significantly lower than for unclassified individuals: 9.3% vs. 34.3% (p-value of 0.005 in both cases; N_1=94, N_2=406).[21] Interestingly, both classified and unclassified individuals faced a volatile

[20] We generated data on individual play through simulations, to determine whether p increases the fit by simply capturing random play. Each simulated player was assigned action C or D at random in each period, drawing from a fixed probability distribution, which corresponded to the empirically observed distribution of actions in the experiment. When running the same strategy-fitting procedure on this simulated dataset, we classify 25% of individuals when p=0 and 41% when p=0.05. These figures are directly comparable with those from the experimental data in Table 3 (53% and 81%, respectively). Adding a probability p of incorrect transition does increase the number of individuals classified, but to a much lesser extent than in the experimental data. Hence, p does not simply capture random play in the experimental data.

[21] The p-values reported in this Section are obtained from regressions results as explained below. We did not rely on statistical tests because observations are generally not independent. Statistics are computed aggregating observations by subject and cycle, unless otherwise noted. Thus we always have 500 observations in total. Strictly speaking, these observations are independent only if we focus on the first cycle. In this Section comparisons are carried out through

environment. Hence, the difference in switch frequency is not a mere consequence of being exposed to opponents with more erratic behavior.

The Z-tree software recorded the number of seconds a subject employed to make each choice. The *decision time* is the number of seconds elapsed between the appearance of the input screen and the confirmation of the choice. Decision time is an additional descriptive variable for subjects' strategies for which there is a growing interest in experimental economics as well as psychology (e.g., Chabris at al. 2008, or see Kosinski, 2006 for psychology). In particular, the literature has suggested that decision time is related to the difficulty of the task, learning, and impulsive or deliberate nature of the decision being made (Rubinstein, 2007). The median decision time for choosing between C and D is more than 35% longer for unclassified than classified subjects (Table 3: 4.26 vs. 3.09 seconds). However, this difference is not significant.

One could think of two alternative interpretations of Result 3. On the one hand, unclassified individuals may be more sophisticated than classified individuals, and so they adopt strategies that are more complex than the ones that can be identified by two-state automata. This greater complexity requires higher cognitive effort, thus longer decision time, and include more frequent action switches due to richer contingencies. On the other hand, unclassified individuals may simply be undecided on what behavior to adopt, and so experiment more within their strategy. The difference in profits for classified and unclassified subjects emerging from Table 3 suggests that experimentation is a likely explanation.[22]

---

regressions where the dependent variable is alternatively (i) average frequency of switch, (ii) average profit, and (iii) mean decision time per observation. The independent variable is a dummy taking value 1 if the observation is classified by any of the 32 strategies considered (zero otherwise). The regression includes fixed effects at the subject level, and errors are computed clustering at the session level. Full regression results are available upon request.

[22] Notice however that this does not imply that every set of classified strategies does better than any set of unclassified strategies. For example, in stationary environments (last two panels of Table 3), individuals following "systematically cooperate" earn less than unclassified individuals because defecting always increases payoffs.

Having described a central difference between classified and unclassified individuals, we now turn to examining what strategies characterize the behavior of classified individuals. We are especially interested in the grim trigger strategy, as it has a prominent role in the way folk theorems define the equilibrium set and the efficiency frontier. Such a strategy supports cooperation by prescribing the harshest possible penalty through decentralized, contagious punishment. Hence, it may appeal to subjects interested in sustaining cooperation in four-person economies where individual reputation cannot be developed.

Result 4. *The grim trigger strategy classifies at most one individual out of four, even when allowing for limited randomness in behavior.*

At most 26.8% of individuals' behavior is consistent with adoption of the grim trigger strategy. This percentage falls to 18.4% when considering deterministic automata. Recall that we use the word "individual" to denote the sequence of all choices of a subject in a single cycle. This classification does not require a subject to follow the same strategy consistently across all cycles. Support for Result 4 comes from Table 3. There is a discrepancy between the results from the Probit analysis (Result 1) and fitting automata on individuals (Result 4).[23] Our probit analysis overestimates the behavior compatible with grim trigger in comparison with the classification of individuals based on automata. To reconcile this apparent discrepancy, we note that a strong aggregate response to an observed defection may result from use of strategies that prescribe punishment forms other than grim trigger. From Table 3, one can see that 33.6% of individuals employ conditional punishment strategies that are unlike grim trigger. Adoption of such strategies can generate the observed aggregate pattern of response to a defection.

[23] As pointed out by a referee, the disparity between aggregate and individual behavior is what one would expect in economies where some subjects playing "Grim Trigger" coexisted with subjects playing "Always defect" and "Always cooperate."

Result 5. *There was heterogeneity in individual behavior and no single strategy can classify the majority of individuals.*

Table 3 displays a summary of results from the empirical identification of strategies. The data suggest the use of heterogeneous strategies. The most common behavior was consistent with "systematically cooperate." The three other largest clusters of classified individuals acted as if having adopted a strategy from classes of strategies, which we denoted "systematically defect," "forgiving." and "grim trigger". Interestingly, individuals adopting unconditional strategies greatly outnumbered those using conditional ones, which marks a difference from other experimental studies about cooperative tasks (Fischbacher et al., 2001) in which a majority of subjects adopt strategies of conditional cooperation. We address this discrepancy in Section 8, where we consider also conditional strategies with longer memory. Moreover, twice as many individuals selected a strategy with cooperation as the initial action, as opposed to an initial defection. In sum, the data suggest the existence of heterogeneity in the strategies followed by individuals and a "preference" for strategies that, roughly speaking, are more cooperative.

As a robustness check, Table 3 reports the strategy-fitting procedure run on three disjoint subsamples: observations in which opponents (i) cooperated as well as defected, (ii) always cooperated, (iii) always defected. Not surprisingly, it is easier to classify observations in a stationary environment because subjects tend to adopt more stable behavior, which is consistent with a wide set of the strategies studied. The percentage of classified individuals grows from 78% (317 out of 406) in subsample (i) to 94% (88 out of 94) in subsamples (ii) and (iii). However, subsample (i) is clearly the most useful for the purpose of identifying strategies, and Result 5 is robust when we only consider subsample (i). Among classified individuals, the grim trigger strategy was only the third most common strategy (22.7%). Instead, behavior consistent with "systematically cooperate" had the highest total fit and classified 45.8% of individuals. Conversely, "systematically defect" classified 33.4% of the individuals.

To sum up, the strategy fitting analysis uncovered significant heterogeneity in individual behavior. Only a minority of individuals acted as if using grim trigger (see also Offerman et al., 2001), while a significant fraction of individuals exhibited unconditional behavior, i.e., played an action that was fixed and independent of the opponents' actions. The uncovered heterogeneity in individual behavior sheds light on the observed patterns of cooperation in society. When people follow a social norm that involves threats of unforgiving economy-wide punishment, it may take just one individual that follows a different strategy to completely unravel cooperation. This suggests that, for a given degree of heterogeneity in the population, full cooperation is harder to sustain in larger economies. Studying four-person economies is an initial step in this direction.

Up to this point the analyses were static. Further information can be gathered by extending the analysis to study dynamic patterns of choices.

Result 6. *Individual behavior changed with experience: 81% of subjects changed strategy from cycle to cycle. Yet, experience did not lead to the general adoption of any specific strategy.*

Recall that each subject in the experiment generates five observations on strategies, i.e., five individuals, one per cycle. If a given strategy q fits all five observations generated by a subject, then we say that strategy q classifies that subject. When we follow each subject across cycles, the data yields a very strong result. Only 19 out of 100 subjects can be classified according to the same strategy in all cycles. Of these, 11 and 7 can be classified as playing "systematically cooperate" and "systematically defect," while only 1 subject adopted a steady behavior consistent with grim trigger. This suggests that most subjects experimented with various strategies across cycles perhaps in an effort to search for a strategy that is a “best response” to play experienced in earlier cycles. In principle, the possibility to experiment with strategies across cycles could improve the chances to reach full cooperation in later cycles. However the data show this was not the case, despite the fact that the economies had only four subjects. Local interactions and anonymity proved to be frictions sufficient to put full cooperation out of reach.

The presence of a learning pattern is confirmed by the analyses of decision times. Decision times display two major patterns (see Table 4). The median decision time is much longer in cycle 1 than other cycles, which suggest learning takes place (9.50 seconds in cycle 1 vs. 2.2 seconds in cycle 5). Also, within each cycle the mean decision time is much longer in period 1 than other periods (13.53 vs. 3.18 seconds), which suggests that the initial decision in a cycle is the most difficult to make. Both patterns emerge when subjects choose either C or D in period 1, which suggests that subjects choose a strategy in the first period of a cycle, thus they need to spend more time thinking. A longer decision time in period 1 of later cycles may reflect experimentation with strategies across cycles.

[Table 4 approximately here]

The data suggest that subjects dealt with the heterogeneity in behavior by tolerating some defections in attempting to coordinate with a subset of participants in the economy. Unlike in two-person economies, in four-person economies a coalition of subjects can profitably coordinate on cooperation. To fix ideas, given the parameterization chosen, a subject can earn more than the minmax payoff even if two persons in the economy always defect. The key requirement is that the remaining subject must cooperate sufficiently often. If two subjects always defect, then a subject who always cooperates earns more than the minmax payoff as long as the third subject cooperates at least 75% of the times. This suggests that a stable subset of systematic cooperators could emerge even if there are systematic defectors. The empirical relevance of these behavioral considerations is well illustrated in the result that follows.

Result 7. *Systematic defectors and systematic cooperators coexisted within most economies.*

To provide evidence for Result 7, we categorize each of the 125 experimental economies depending on the classification of individuals within each economy. Individuals classified as systematic cooperators coexisted with systematic defectors in more than half of the economies.[24] In addition, we categorize the 212 individuals classified by the "systematic cooperation" class according to their presence within each economy. Only 41 individuals were the sole systematic cooperator in the economy, while 56 individuals were found in economies where everybody systematically cooperated. The remaining 115 individuals systematically cooperated in economies where someone, though not everyone, sometimes defected.[25] In other words, the data show that oftentimes subjects unconditionally cooperated even in economies where defectors were present, which supports the view that subgroups successfully coordinated on cooperation. As earlier noted, disciplining a lone, anonymous defector by punishing future random opponents impairs the possibility of coordinating on cooperation with the others. This provides a behavioral justification for why grim trigger is not the strategy of choice in our experimental economies. Indeed, Results 7 shows that persistent opportunistic behavior goes often unpunished. Clearly, there may be other reasons for the observed behavior, such as other regarding preferences. Other-

[24] In only 3 economies we could not classify individuals as either systematic cooperators or systematic defectors; in 64 economies both classes of strategies were observed; in 39 economies there were systematic cooperators but no systematic defectors; and in 19 economies the reverse was true. The average length of cycles in each of these four categories of economies was, respectively, 32.7, 11.8, 18.2, and 22.4 periods. This evidence also suggests, although it does not prove, that strategies based on the public randomization device were uncommon.

[25] Subjects who followed "systematically cooperate" faced environments characterized by different degrees of cooperation: 48 subjects faced 100% cooperation; 61 faced a cooperation rate between 67% and 99%; 64 subjects faced a cooperation rate between 33% and 66%, and 39 faced less than 33% cooperation. This means that the expected payoff for a subject following "systematically cooperate" is at least 15.3 points, which is higher than the minmax payoff of 10.

regarding preferences may support or hinder the use of "systematically cooperate" strategies depending on what motivates subjects. On the one hand, altruistic motives and positive reciprocity may prevent subjects from punishing after observing a defection because punishment destroys surplus and harms cooperators and defectors alike. On the other hand, positional motives reinforce the urge to punish after a defection in order to prevent others from getting ahead in terms of relative share of income.

7. Strategy identification through maximum likelihood

Here, we estimate the importance of each candidate strategy with a maximum likelihood approach, as in Dal Bó and Fréchette (2011) and Fudenberg et al. (forthcoming). This methodology enhances comparability with related studies and allows us to study the case when automata can make errors in implementing actions, as opposed to errors in transitions.

The estimation comprises a set of twenty-six strategies.[26] The estimation employs data from all cycles and presumes that subjects (i) have a given probability of choosing one of the 26 strategies, (ii) may change strategies from cycle to cycle and (iii) may make errors on actions, i.e., with some probability may choose an action that is not recommended by the strategy.[27]

The maximum likelihood estimation differs from the one presented in Section 5 because (a) it relies on the assumption that subjects may make errors in actions (not in transitions) and (b) it estimates the prevalence of errors in implementing actions (the parameter γ), rather than

[26] There are 26 strategies now because if there are no errors in transitions, then every strategy in the class "Systematically cooperate" (Systematically defect") coincides with the one-state automaton "Always cooperate" (Always defect). All two-state automata in the class "Systematically cooperate" ("Systematically defect") prescribe the same behavior as the strategy "Always cooperate" ("Always defect").

[27] The estimation was executed adapting the code included in the supplementary material of Dal Bó and Fréchette (2011), where the reader can also find the details of the estimation procedure. However, unlike in Dal Bó and Fréchette (2011) and Fudenberg et al. (forthcoming), our estimation procedure assumes that subjects may change strategies from cycle to cycle; the code has been adapted accordingly.

imposing a maximal number of errors (the parameter p).[28] One can illustrate the difference between errors in actions and in transitions using as an example the “Grim trigger” automaton. If everyone plays grim trigger, then a single wrong choice moves the economy to permanent punishment, which is an absorbing state because a mistake is never forgotten. Instead, mistakes in transition allow for a sort of experimentation within the strategy, as it is still possible to change state and revert to a cooperative mode, hence we may observe “fresh starts” or alternating spells of cooperation and punishment. The maximum likelihood estimates the prevalence of errors in actions (the parameter γ) and “classifies” every individual. Instead, fitting automata to observations, as done in Section 6, implies that a fraction of observations may end up “unclassified.” Because this fraction depends on the exogenously-specified maximal number of errors in transitions (the parameter p), the strategy-fitting procedure helps us assessing the sensitivity of results to increasing margins of error (Figure 3).

[Table 5 approximately here]

Table 5 reports the estimates of the population proportions for each of the 26 strategies. These maximum likelihood estimates confirm the main results on strategy identification reported in Section 6. The four most likely strategies are “Always cooperate”, “Grim trigger”, “Tit-for-Tat” and “Always defect” and, as noted in Result 5, there is significant heterogeneity in individual behavior with no strategy clearly prevailing. In line with Result 4, grim trigger can account only for a minority of observations, in the order of one out of five. Table 5 provides additional support for the finding that a significant fraction of individuals (more than 50%) exhibited unconditional behavior, i.e., chose an action that was fixed and independent of the actions of previous opponents. In short, the results reported in Section 6 are robust to the use of a maximum

[28] The assumption that subjects can make mistakes in implementing actions also underlies the approach adopted by Engle-Warnick, McCausland, and Miller (2004).

likelihood approach and to consideration of errors in actions. The strategy-identification approach proposed in Section 5 allows us to perform additional empirical analyses of the behavior of individuals, such as differentiating classified and unclassified individuals. It also allows us to assess how the explanatory power of a given strategy varies as we increase the probability of errors.

The finding that a majority of subjects followed strategies compatible with unconditional behavior departs from findings in other experimental studies about cooperation (e.g., Fischbacher et al., 2001), and so it should be further investigated. Indeed, it is possible that some observations could result from subjects using complex conditional strategies that cannot be represented by two-state automata. More concretely, consider an individual who starts defecting after suffering several defections in the previous four rounds. This individual could be classified as "systematically cooperate" when allowing for some errors in transitions, although she may in fact punish as part of a strategy conditional on the number of defections suffered. We investigate this possibility in Section 8, where we repeat the strategy-fitting procedures adopted in Sections 6 and 7 for an expanded set of strategies.

8. Robustness check on strategy classification

We expand the strategy set to also include 11 conditional cooperative strategies with longer memory. In particular, we follow Fudenberg et al. (forthcoming) and include "lenient" variants of grim trigger and of "forgiving" strategies. These strategies prescribe initial cooperation and a switch to defection after the player has met $k \geq 1$ defecting opponents. We consider variants in which the punishment can be either permanent (as in Grim trigger) or temporary (as in TFT); the triggering event can be a cumulative, a consecutive or a proportional count of defections (see notes to Tables 6-7).

[Table 6 approximately here]

Table 6 reports the results for the classification approach proposed in Section 5. When considering the full sample, the conditional strategies with longer memory classify only 26 observations that were not already classified under any of the one- and two-state deterministic automata. Hence, the fraction of classified individuals marginally increase from 53% to 58.2% (p=0). In addition, no single strategy with longer memory that we considered has a total fit as high as "always cooperate" (which is 134 in the full sample and 70 in the sample restricted to opponents playing both C and D). To further assess the power of strategies with longer memory, Table 6 reports the number of individuals classified by each new strategy *in addition* to those jointly classified by "always cooperate," "grim trigger," and "tit-for-tat" (last column). One can see a substantial overlap between the individuals classified by the above 3 strategies and the 11 conditional strategies with longer memory.

Note also that most individuals classified by "always cooperate" are classified *also* by one or more conditional strategies with longer memory (126 out of 134). This is not surprising: any observation consisting of an uninterrupted sequence of cooperative actions is compatible with "always cooperate" and with cooperative strategies in which punishment is triggered by a condition that never took place. Grim2 is largely responsible for the increase in the total number of classified observations.

[Table 7 approximately here]

Table 7 reports the results obtained using the maximum likelihood approach. When considering the full sample, the parameter γ marginally decreases from 0.54 to 0.48. We find that "Always cooperate" and "Grim trigger" lose some share, which is captured by lenient longer-memory variants of conditional cooperative strategies (i.e., Grim and Tit-for-tat). Yet, we also find that the two unconditional strategies “Always cooperate” and “Always defect” are still the ones that capture the behavior of the greatest shares of individuals.

To sum up, based on the strategy classification method, one cannot strictly rule out that individuals classified as “Always cooperate” are following conditional strategies with longer memory. In Table 6, the longer-memory strategies considered classify most of the observations also classified by “Always cooperate.” However, based on the maximum likelihood estimates, we find that the unconditional strategy “Always cooperate” is still the one that captures the greatest share of cooperative strategies (Table 7).

9. Final Remarks

This experimental study offers novel insights about subjects’ strategies in decentralized trading environments where mutual gains from cooperation coexist with incentives to behave opportunistically. We studied economies where subjects faced an indefinite sequence of prisoner’s dilemmas played in pairs with changing opponents. Because the interaction was anonymous, subjects could not build a reputation. Moreover, each economy comprised four subjects, which made coordination harder to achieve than in two-person economies that have been the focus of previous experiments with indefinite interaction.

We empirically study equilibrium and strategy selection in supergames. In our setup, a social norm based on the threat of contagious punishment can support full cooperation. The analysis accounts for equilibrium strategies—such as grim trigger and unconditional defection—as well as for non-equilibrium strategies—such as tit-for-tat and unconditional cooperation. The experimental design helps the empirical identification of *individual* strategies thanks to substantially longer sequences of play than previous work, a design based on four-person economies, and a diverse subject pool (college students, MBA students, and white-collar workers).

An assessment based on a standard Probit regression analysis suggests that there is a strong initial attempt to coordinate on cooperation and the first experienced defection triggered a permanent downward shift in cooperation levels. More in-depth strategy-identification analyses reveal that grim trigger is not the prevalent norm of behavior. Subjects did not follow social norms that rely on contagious punishment schemes. Furthermore, we found substantial heterogeneity in the strategies used, which persisted with experience. These findings are robust to the adoption of three alternative strategy-identification techniques: fitting deterministic automata; fitting automata that can transition from state to state *stochastically*; identifying strategies through a maximum likelihood approach.

Subjects tried to reach a cooperative outcome but did so without being able to coordinate their strategy choices, independently searching for suitable strategies. In such an environment, unilaterally adopting grim trigger does not make full cooperation more likely. In fact, if a subgroup of subjects wants to coordinate on cooperation, then playing grim trigger may jeopardize coordination attempts and simply drag the economy towards full defection. This may explain why grim trigger was uncommon in the data and why systematic defection did not crowd out systematic cooperation.

Grim trigger seems to be rare also in field settings with repeated social dilemmas. Ostrom (2010) surveyed field studies of community management of fisheries and water resources and did not find a single case where harvesters used the grim trigger strategy. These considerations point to a weak predictive power of theories based on homogeneous agents who adopt strategies of uncompromising, contagious punishment. This suggests care must be taken in drawing immediate conclusions from applications based on folk theorems. For example, theories that trace the efficiency frontier by presuming everyone follows a norm of unforgiving, universal punishment, have low descriptive power vis-à-vis our experiment. Adoption of a social norm based on grim trigger did not emerge in our economies where reputation-based strategies were unavailable. The possibility to resort to norms of decentralized, contagious punishment did not stave-off opportunistic behavior. On the contrary, most subjects were willing to forgive a defection, to different degrees. Some reacted to a defection with a temporary punishment, while others systematically cooperated even in the presence of relentless defectors.

These findings suggest that, on the one hand further investigations should be conducted with larger economies, to determine whether the findings with four-agent economies hold when the economy's size increases. On the other hand, a theoretical challenge remains, which is to increase the descriptive power of folk theorem-type results in interactions among strangers.

Acknowledgements

We thank three anonymous referees for suggesting additional analyses that have helped us to improve the paper. We also thank, Jim Engle-Warnick, Tim Cason, Guillaume Fréchette, Hans Theo Normann, Davide Raggi and seminar participants at the ESA in Innsbruck and Chicago, IMEBE in Bilbao, the Einaudi Institute for Economics and Finance (EIEF), NYU and the universities of Bologna, Frankfurt, Illinois, Fudan and Siena for comments and suggestions. Research support was partially provided by grants from Purdue's CIBER, the EIEF and NSF Grant CCF-1101627. An early version of this paper was completed while G. Camera was visiting the University of Siena thanks to a Fulbright grant.

References

Aoyagi, M., and Fréchette, G., 2009. Collusion as Public Monitoring Becomes Noisy: Experimental Evidence. Journal of Economic Theory 144 (3), 1135-1165.

Bigoni, M., Camera, G., and Casari, M., 2012. Strategies of cooperation and punishment among students and clerical workers. Unpublished manuscript, Purdue University.

Binmore, K., 2005. Commentary to "Economic man" in Cross-cultural perspective: Behavioral experiments in 15 small-scale societies. In Joseph Henrich et al. (Eds.), Behavioral and Brain Sciences, 28, 817-818.

Blonski, M. and Giancarlo, S., 2001. Prisoners' Other Dilemma. SSE/EFI Working Paper 437.

Blonski, M., Ockenfels, P. and Spagnolo, G., 2011. Equilibrium selection in the repeated prisoner's dilemma: axiomatic approach and experimental evidence. American Economic Journal : Microeconomics, 3(3), 164–92.

Camera, G., and Casari, M., 2009. Cooperation among strangers under the shadow of the future. American Economic Review, 99(3), 979–1005.

Chabris, C. F., Laibson, D., Morris, C., Schuldt, J., and Taubinsky, D., 2008. Measuring Intertemporal Preferences Using Response Times. NBER Working Paper #14353.

Dal Bó, P., 2005. Cooperation under the Shadow of the Future: Experimental Evidence from Infinitely Repeated Games. American Economic Review, 95(5): 1591-1604.

Dal Bó, P., and Fréchette, G., 2011. The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence. American Economic Review.

Diamond, P. A., 1982. Aggregate Demand Management in Search Equilibrium. Journal of Political Economy, 90(5): 881-894.

Duffy, J., and Ochs, J., 2009. Cooperative Behavior and the Frequency of Social Interaction. Games and Economic Behavior, 66, 785-812.

Duffy, J., Xie, H. and Lee, Y., Forthcoming. Social Norms, Information and Trust among Strangers: Theory and Evidence. Economic Theory.

Ellison, G., 1994. Cooperation in the Prisoner's Dilemma with Anonymous Random Matching. Review of Economic Studies, 61: 567-88.

Engle-Warnick, J., McCausland, W., and Miller. J., 2004. The ghost in the machine: inferring machine-based strategies from observed behavior. Working paper 2004-11, University of Montreal.

Engle-Warnick, J., and Slonim, R.L., 2006. Inferring Repeated-Game Strategies from Actions: Evidence from Trust Game Experiments. Economic Theory, 28: 603–632.

Fischbacher, U., 2007. z-Tree: Zurich Toolbox for Ready-made Economic Experiments. Experimental Economics 10(2), 171-178.

Fischbacher, U., Gaechter, S., and Fehr, E., 2001. Are people conditionally cooperative? Evidence from a public goods experiment. Economics Letters 71, 397–404

Friedman, J. W., 1971. Non-Cooperative Equilibrium for Supergames. Review of Economic Studies, 38(113): 1-12.

Fudenberg, D., Rand, D. G. and Dreber, A., Forthcoming. Turning the Other Cheek: Leniency and Forgiveness in an Uncertain World. American Economic Review.

Kandori, M., 1992. Social Norms and Community Enforcement. Review of Economic Studies, 59: 63-80.

Kosinski, R. J., 2006. A Literature Review on Reaction Time. Unpublished manuscript, Clemson University.

Mortensen, D. T., and Pissarides, C. A., 1994. Job Creation and Job Destruction in the Theory of Unemployment. Review of Economic Studies, 61(3): 397–415.

Offerman, T., Potters, J. and Verbon, H., 2001. Cooperation in an Overlapping Generations Experiment. Games and Economic Behavior 36, 264-275.

Ostrom, E., 2010. Beyond Markets and States: Polycentric Governance of Complex Economic Systems. American Economic Review, 100, 641–672.

Ostrom, E., Walker, J., and Gardner, R., 1992. Covenants with and without a Sword: Self-Governance Is Possible. American Political Science Review, 86(2), 404–17.

Palfrey, T. R., and Rosenthal, H., 1994. Repeated Play, Cooperation and Coordination: An Experimental Study. Review of Economic Studies, 61: 545-565.

Rapoport, A., and Chammah, A., 1965. Prisoner's dilemma: A study in conflict and cooperation. Ann Arbor: The University of Michigan Press.

Roth, A. E., and Murnighan, K., 1978. Equilibrium Behavior and Repeated Play of The Prisoner's Dilemma. Journal of Mathematical Psychology, 17: 189-98.

Rubinstein, A., 1986. Finite automata play the repeated prisoner's dilemma. Journal of Economic Theory 39, 83–96.

Rubinstein, A., 2007. Instinctive and Cognitive Reasoning: a Study on Response Times. The Economic Journal, 177, 1245-1259.

Stahl, D. O., 2009. An Experimental Test of the Efficacy of Simple Reputation Mechanisms to Solve Social Dilemmas, unpublished manuscript, University of Texas at Austin.

Tables and Figures

(A) Notation in the theoretical analysis

Player 1/ Player 2	*Cooperate (Y)*	*Defect (Z)*
Cooperate (Y)	*y,y*	*l,h*
Defect (Z)	*h,l*	*z,z*

(B) Parameterization of the experiment

Player 1/ Player 2	*Cooperate (Y)*	*Defect (Z)*
Cooperate (Y)	25, 25	5, 30
Defect (Z)	30, 5	10, 10

Table 1: The stage game

Strategies starting with C=cooperate			Strategies starting with D=defect		
Strategy	**Strategy Set**	***N***	**Strategy**	**Strategy Set**	***N***
11111	*Systematically cooperate*	168	00000	*Systematically defect*	111
11110		178	01000		109
11101		166	00100		94
11100		188	01100		98
11000	*Grim trigger*	134	00001	*Unconventional*	*26*
11010	*Forgiving*	113	00010		11
11011		103	00011		5
11001		107	01001		25
10111	*Unconventional*	10	01010		13
10110		13	01011		7
10101		13	00101		21
10100		21	00110		9
10011		2	00111		3
10010		9	01101		22
10001		7	01110		11
10000		24	01111		6

Table 2: Classification of strategies

Notes: In this table *p*=0.05. *N* is the number of observations classified by a strategy. Each of the 32 strategies is coded as a five-element vector. Each element corresponds to a state, i.e., an action to be taken, with C =1 and D=0. The first element is the initial state. The remaining four elements identify the state reached following current play (equivalently, the action to be implemented in the next round). Denote c and d the actions of the opponent. The second element in the vector identifies the state reached if (C,c) is played. The remaining elements identify the states reached given play (C,d), (D,c) and (D,d), respectively. For instance the automaton 11010 represents "tit-for-tat." It starts with C, prescribes play D in two instances, if (C,d) or (D,d) are the outcomes (third and fifth element in the sequence), and prescribes play C if (C,c) or (D,c) are the outcomes. The first four automata in each column are called "systematically cooperate" and "systematically defect" because they prescribe the automaton should remain always in the initial state (cooperate or defect) unless a random shock generates a transition to an incorrect state. For instance, with 11110 the agent starts in state C and remains in C; state D can be reached only by mistake, in which case the player remains in D only if her opponent plays d (last element of the vector). Clearly the automaton 11111 is unconditional cooperation (always cooperate), i.e., does not allow for mistakes or experimentation. The same holds for unconditional defection, 00000 (always defect).

32 automata	Random transitions (p=0.05)			Deterministic (p=0.00)	
	N	median response time	average profit	*N*	average profit
All Observations	500	3.42	18.06	500	18.06
C in period 1	336	3.51	17.68	336	17.68
D in period 1	164	2.93	18.85	164	18.85
<u>Classified</u>	405	3.09	18.74	265	19.82
C in period 1	272	3.38	18.28	173	19.45
-systematically cooperate	212	3.00	18.60	134	19.92
-forgiving	130	4.33	20.52	90	22.44
-grim trigger	134	3.45	20.79	92	22.18
-unconventional	44	4.35	14.40	28	15.20
D in period 1	133	2.56	19.67	92	20.50
-systematically defect	120	2.49	19.51	86	20.33
-unconventional	48	5.09	21.87	28	24.16
<u>Unclassified</u>	95	4.26	15.18	235	16.08
C in period 1	64	4.32	15.11	163	15.80
D in period 1	31	4.20	15.33	72	16.74
Opponents play both C & D	406	3.50	17.12	406	17.12
<u>Classified</u>	317	3.41	17.65	182	18.13
C in period 1	201	3.84	16.90	105	17.26
-systematically cooperate	145	3.41	16.69	70	16.71
-forgiving	68	8.00	17.16	28	18.53
-grim trigger	72	4.85	17.86	30	18.02
-unconventional	32	3.45	16.63	19	18.26
D in period 1	116	2.54	18.95	77	19.31
-systematically defect	106	2.46	18.89	74	19.31
-unconventional	31	5.54	20.39	13	21.35
<u>Unclassified</u>	89	4.06	15.24	224	16.31
C in period 1	62	4.21	15.07	158	15.94
D in period 1	27	3.44	15.63	66	17.17
Opponents always play C	75	2.2	25.83	75	25.83
<u>Classified</u>	73	2.14	25.79	73	25.79
C in period 1	60	2.00	25.06	60	25.06
-systematically cooperate	59	2.00	25.00	59	25.00
-forgiving	59	2.00	25.00	59	25.00
-grim trigger	59	2.00	25.00	59	25.00
-unconventional	1	6.00	28.33	1	28.33
D in period 1	13	2.86	29.16	13	29.16
-systematically defect	10	5.01	30.00	10	30.00
- unconventional	13	2.86	29.16	13	29.16
<u>Unclassified</u>	2	4.94	27.23	2	27.23
C in period 1	1	4.55	26.13	1	26.13
D in period 1	1	5.33	28.33	1	28.33
Opponents always play D	19	6.84	7.52	19	7.52
<u>Classified</u>	15	6.84	7.44	10	7.03
C in period 1	11	7.60	6.63	8	6.29
-systematically cooperate	8	8.05	5.96	5	5.00
-forgiving	3	4.00	8.44	3	8.44
-grim trigger	3	4.00	8.44	3	8.44
-unconventional	11	7.60	6.63	8	6.29
D in period 1	4	2.37	9.64	2	10.00
-systematically defect	4	2.37	9.64	2	10.00
- unconventional	4	2.37	9.64	2	10.00
<u>Unclassified</u>	4	6.56	7.83	9	8.06
C in period 1	1	8.00	6.67	4	7.33
D in period 1	3	5.43	8.22	5	8.65

Table 3: Analysis of individual strategies

Notes: The unit of observation is the sequence of all choices of a subject in a cycle, i.e., an individual. When no confusion arises we refer to such a sequence as one *observation*. There are 500 observations. *N* refers to the number of observations classified by the corresponding strategy. An observation is classified according to strategy set Q, if at least one strategy $q \in Q$ fits, i.e.: (i) the initial action is correctly predicted by *q;* (ii) *q* has the largest consistency score (see explanation in text) among all strategies in Q; and (iii) when we allow for random transitions, the probability of observing *n* or more inconsistent actions is smaller than 10% given the experimentation parameter *p*=0.05. Otherwise, the observation is "Unclassified." Clearly, if we do not allow for random transitions, i.e. *p=0*, then item (iii) is modified as follows: the probability of observing any inconsistent action must be zero.

	Cycle					
	1	2	3	4	5	Total
(1 obs. = a single choice in a period)						
Cooperation in all periods (in %)	53.9	54.3	48.3	57.6	54.6	53.8
Cooperation in period 1 (in %)	74.0	64.0	65.0	68.0	65.0	67.2
Coordination on cooperation (in %)	33.3	31.0	30.6	40.3	34.5	33.9
Average profit per period (in points)	18.09	18.15	17.24	18.64	18.19	18.06
Median decision time (in seconds)	9.50	3.96	2.37	2.00	2.20	3.42
Switch frequency (in %)	33.2	25.3	25.6	23.9	32.8	28.2
All observations *(1 obs = an individual)*						500
Classified observations (in %)	76.0	83.0	69.0	89.0	88.0	81.0
of which: classified by grim trigger	23.0	18.0	25.0	31.0	37.0	26.8
Subsample: opponents play both C and D (N)						406
Classified observations (in %)	75.3	81.9	61.7	87.3	85.3	78.1
of which: classified by grim trigger	18.8	14.9	13.6	18.3	24.0	17.7

Table 4: Summary statistics by cycle

Notes: "Cooperation" reports the percentage of C choices; "Coordination on cooperation" reports the percentage of subject pairs were both chose C: it indicates the mean of cc where cc is 1 for a C choice when the opponent chose C and is 0 for a C choice when the opponent did D and for a D choice.

Strategies starting with C=cooperate				Strategies starting with D=defect			
Strategy	**Strategy Set**	**Coeff.**	**S.E.**	**Strategy**	**Strategy Set**	**Coeff.**	**S.E.**
11111	*Systematically cooperate*	0.336	0.043***	00000	*Systematically defect*	0.282	0.056***
11000	*Grim trigger*	0.190	0.043***	00001	*Unconventional*	0.000	0.002
11010	*Forgiving*	0.158	0.041***	00010		0.000	0.000
11011		0.000	n.a.	00011		0.000	0.000
11001		0.000	0.002	01001		0.005	0.005
10111	*Unconventional*	0.000	0.000	01010		0.000	0.013
10110		0.002	0.004	01011		0.000	0.000
10101		0.000	0.000	00101		0.003	0.004
10100		0.015	0.017	00110		0.000	0.004
10011		0.000	0.000	00111		0.000	0.001
10010		0.000	0.000	01101		0.009	0.010
10001		0.000	0.006	01110		0.000	0.000
10000		0.000	0.002	01111		0.000	0.000
	γ	**0.542**	**0.064*** **				

Table 5: Maximum likelihood estimates of individual strategies

Notes: p-values are calculated using bootstrapped standard errors. The coefficient for strategy 11011 is not estimated directly, but it is implied by the fact that the proportions must sum to one. γ is an endogenous parameter of the estimation that measures the probability of errors. A larger γ denotes a higher probability of errors. * Significant at the 10% level, ** significant at 5% level, *** significant at 1% level.

	N	Observations not classified by either "always cooperate," "grim trigger," or "tit-for-tat"
All Observations	500	336
Classified by at least one of the 26 strategies with short memory.	*265*	*101*
Classified by at least one of the 37 strategies	291	127
Grim 2 A	107	18
Grim 3 A	102	4
Grim 2 B	112	14
Grim 3 B	123	4
Grim 33%	103	1
Grim 67%	123	9
Grim 100%	118	0
2TFT	87	3
3TFT	91	4
TF2T	109	11
TF3T	120	1
Unclassified	209	0
Opponents play both C and D	406	309
Classified by at least one of the 26 strategies with short memory.	*182*	*85*
Classified by at least one of the 37 strategies	207	110
Grim 2 A	46	17
Grim 3 A	39	4
Grim 2 B	51	13
Grim 3 B	60	4
Grim 33%	41	1
Grim 67%	61	9
Grim 100%	56	0
2TFT	26	2
3TFT	30	3
TF2T	48	10
TF3T	57	1
Unclassified	199	0

Table 6: Conditional strategies with longer memory

Notes: Strategy Grim2 (Grim3) triggers to permanent punishment when two (three) past opponents defected. Strategies Grim A consider the total number of past defections (in the cycle), while Grim B considers only the number of consecutive defections. Strategy Grim33% (Grim67%, Grim100%) triggers the punishment phase when the frequency of defections reaches 33% (67%, 100%). Strategy TF2T (TF3T) prescribes C unless each of the last two (three) opponents played D. Strategy 2TFT (3TFT) prescribes C unless a defection was suffered in either of the last 2 (3) rounds. Automata are deterministic, i.e., p=0. The automaton "always cooperate" is 11111, "grim trigger" is 11000, and "tit-for-tat" is 11010.

Strategies starting with C=cooperate				Strategies starting with D=defect			
Strategy	**Strategy Set**	**Coeff.**	**S.E.**	**Strategy**	**Strategy Set**	**Coeff.**	**S.E.**
11111	*Systematically cooperate*	0.138	0.040***	00000	*Systematically defect*	0.282	0.064***
11000	*Grim trigger*	0.050	0.025**	00001	*Unconventional*	0.000	0.000
11010	*Forgiving*	0.067	0.022*	00010		0.000	0.000
11011		0.000	0.001	00011		0.000	0.000
11001		0.002	0.004	01001		0.005	0.006
10111	*Unconventional*	0.000	0.000	01010		0.001	0.014
10110		0.002	0.004	01011		0.000	0.000
10101		0.000	0.002	00101		0.004	0.005
10100		0.004	0.008	00110		0.000	0.000
10011		0.000	0.000	00111		0.000	0.000
10010		0.000	0.000	01101		0.007	0.011
10001		0.000	0.006	01110		0.000	0.000
10000		0.000	0.001	01111		0.000	0.000
Grim 2 A	*Variants of Grim trigger*	0.038	0.021*	2TFT	*Variants of Tit-for-tat*	0.012	0.012
Grim 3 A		0.038	0.035	3TFT		0.025	0.024
Grim 2 B		0.022	0.020	TF2T		0.072	0.025***
Grim 3 B		0.086	0.025***	TF3T		0.061	n.a.**
Grim 33%		0.000	0.011				
Grim 67%		0.085	0.052				
Grim 100%		0.000	0.014				
	γ	**0.486**	**0.064*** **				

Table 7: Maximum likelihood estimates of conditional strategies with longer memory

Notes: p-values are calculated using bootstrapped standard errors. The coefficient for strategy TF3T is not estimated directly, but it is implied by the fact that the proportions must sum to one. γ is an endogenous parameter of the estimation that measures the probability of errors. A larger γ denotes a higher probability of errors. * Significant at the 10% level, ** significant at 5% level, *** significant at 1% level. For the definitions of the variants of Grim trigger and Tit-for-tat see the notes to Table 6.

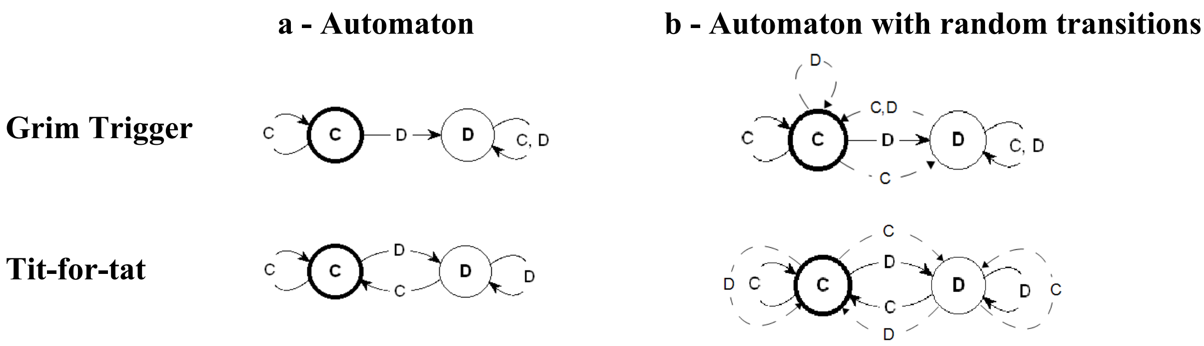

Figure 1: Strategy representation using automata (C=cooperate, D=Defect)

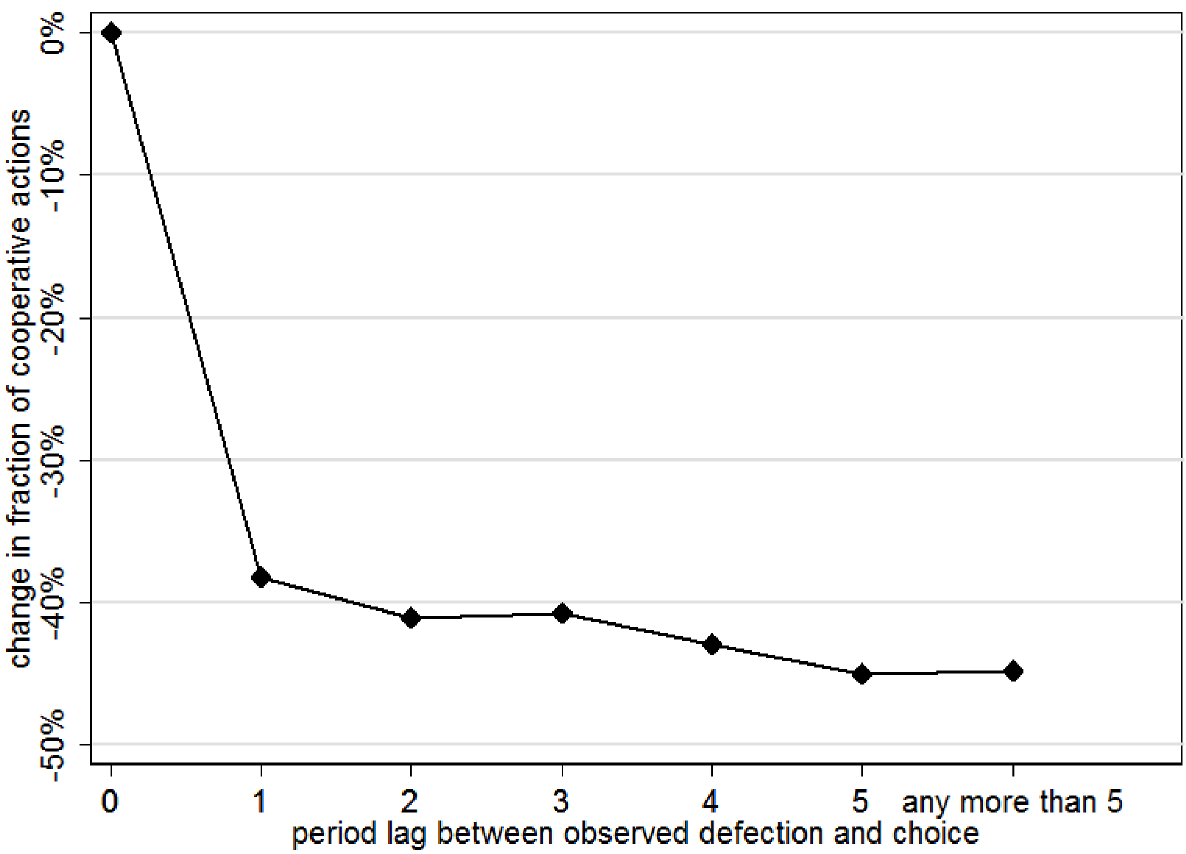

Figure 2: Aggregate response to an observed defection

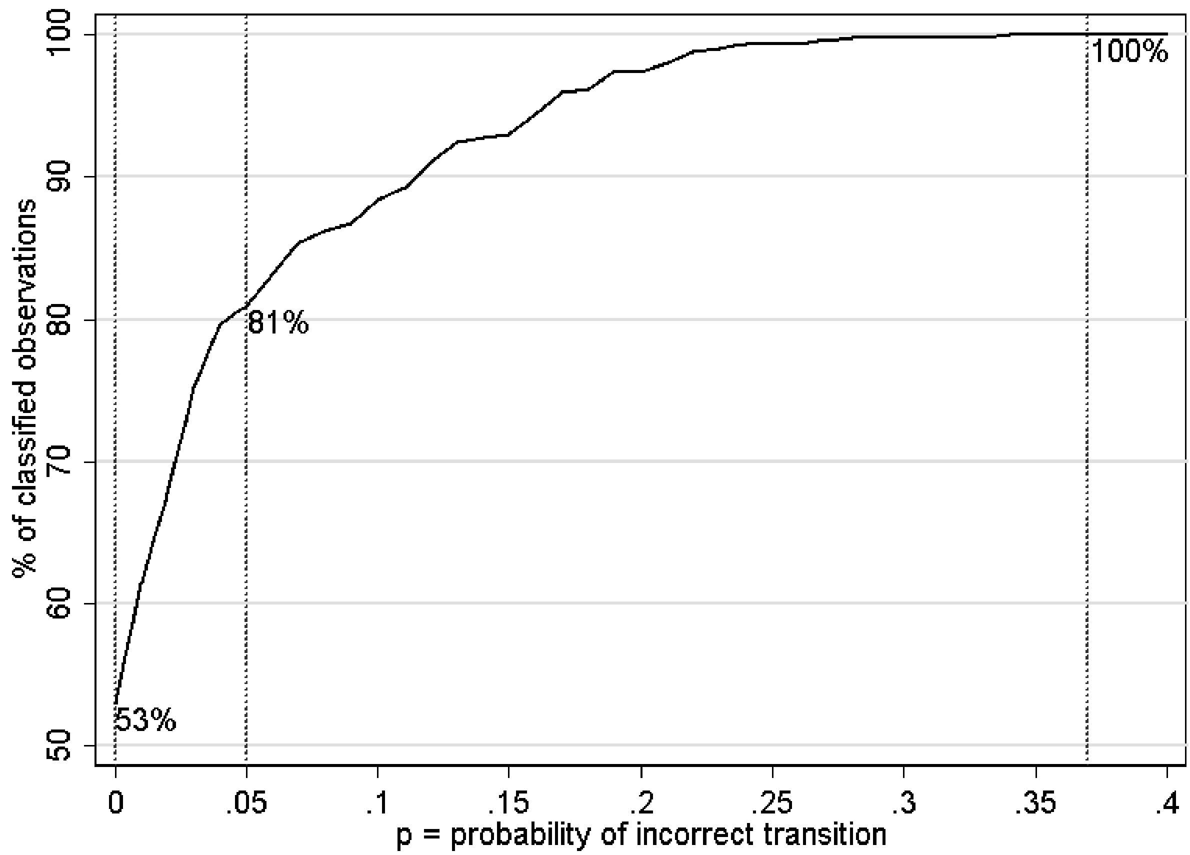

Figure 3: Percentage of classified observations