The background of the entire page is a grayscale photograph of a modern university building. The building features a prominent glass facade with vertical window lines. In the foreground, there is a paved courtyard with several trees and concrete benches. The overall scene is bright and clear.

KRANNERT SCHOOL OF MANAGEMENT

Purdue University
West Lafayette, Indiana

Cooperative Strategies in Groups of Strangers: An Experiment

By

Gabriele Camera
Marco Casari
Maria Bigoni

Paper No. 1237
Date: June, 2010

Institute for Research in the
Behavioral, Economic, and
Management Sciences

Cooperative strategies in groups of strangers: an experiment

Gabriele Camera, Marco Casari, and Maria Bigoni*

15 June 2010

* Camera: Purdue University; Casari: University of Bologna, Piazza Scaravilli 2, 40126 Bologna, Italy, Phone: +39-051-209-8662, Fax: +39-051-209-8493, marco.casari@unibo.it; Bigoni: University of Bologna, Piazza Scaravilli 2, 40126 Bologna, Italy, Phone: +39-051-2098890, Fax: +39-051-0544522, maria.bigoni@unibo.it. We thank Jim Engle-Warnick, Tim Cason, Guillaume Fréchet, Hans Theo Normann, and seminar participants at the ESA in Innsbruck, IMEBE in Bilbao, University of Frankfurt, University of Bologna, University of Siena, and NYU for comments on earlier versions of the paper. This paper was completed while G. Camera was visiting the University of Siena as a Fulbright scholar. Financial support for the experiments was partially provided by Purdue's CIBER and by the Einaudi Institute for Economics and Finance.

Abstract

We study cooperation in four-person economies of indefinite duration. Subjects interact anonymously playing a prisoner's dilemma. We identify and characterize the strategies employed at the aggregate and at the individual level. We find that (i) grim trigger well describes aggregate play, but not individual play; (ii) individual behavior is persistently heterogeneous; (iii) coordination on cooperative strategies does not improve with experience; (iv) systematic defection does not crowd-out systematic cooperation.

Keywords: repeated games, equilibrium selection, prisoners' dilemma, random matching.

JEL codes: C90, C70, D80

I. Introduction

Fostering cooperation in society can be problematic when individual reputation is not easily established. 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 theorems suggest that, over the long haul, none of these frictions present a fundamental obstacle to cooperation.¹ Groups of self-regarding individuals can overcome the short-run temptation to cheat others by threatening permanent defection through a decentralized punishment scheme that spreads by contagion. The open question is how, in practice, groups of individuals reach cooperation when theoretically feasible and what strategies they adopt to sustain it.

To address the above question, we gathered data from an experiment where four subjects interacted over a long horizon. 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). Our design makes it possible to empirically identify and characterize strategies employed by subjects. Little is known about how individuals play indefinitely repeated games. The existing evidence on individual strategies is limited 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, 2009). The present study advances the understanding of how individuals play indefinitely repeated games by studying individual behavior in four-person economies of substantially longer duration than in the literature, and with a varied subject pool (undergraduates, MBA students, and white-collar workers).

¹ The foundation for this affirmation traces back to the folk theorems in Friedman (1971) and the random-matching extensions in Kandori (1992) and Ellison (1994).

Folk theorems show that long-run interaction can sustain a multiplicity of equilibrium outcomes, but do not offer much guidance regarding which equilibrium will be selected. Applications of theories of infinitely repeated games assume, often implicitly, that agents are homogenous and will select the most efficient among the available equilibria. We report that efficiency is rarely achieved in our experimental economies, which suggests that efficiency may not be the key equilibrium selection criterion. The individual-level analysis in this study sheds light on empirically-relevant equilibrium selection criteria.

Folk theorems trace the efficiency frontier through a “grim trigger” strategy whereby all players cooperate under the threat of a contagious process of economy-wide defection. At the aggregate level the data look consistent with the notion of grim trigger play. However, when we dig deeper and analyze data at the individual level, this is no longer true. Only one out of four subjects behaves in a manner consistent with the use of the grim trigger strategy.

The data suggest that different subjects adopted different strategies, and no single strategy was prevalent. Subjects tried out a variety of strategies but this process failed to yield full cooperation, and did not improve coordination on cooperative strategies. Moreover, a substantial fraction of participants did not use conditional strategies: they either systematically cooperated or systematically defected, independently of the actions taken by their randomly encountered opponents. In short, theories based on the notion that cooperation emerges because individuals adopt an identical strategy based on the threat of unforgiving, generalized punishment have limited descriptive power. Our results challenge theory to provide more descriptive guidance.

The set-up we have adopted to study individual strategies, is novel relative to the existing experimental literature. An economy comprises four persons who interact *locally* and *anonymously*. The interaction is local because subjects observe only outcomes in their pair but

not in the rest of the economy. 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. This makes individual reputation impossible to build and coordination harder to achieve than in two-person economies.

Our design with four-person economies is a convenient abstraction for representing a wide range of economies where reputation is hard to establish, it is difficult or costly to monitor the actions of all other members of society, to communicate intentions, and where institutions for enforcement have limitations. By experimentally controlling the informational flows and the matching process, our small laboratory economies capture essential features observed in larger economies, without the need to let hundreds of people interact together. An important segment of the economic literature adopts anonymous random matching economies as the platform for theoretical analysis. For example, in macroeconomics consider the matching models as in Diamond (1982) and in microeconomics consider the decentralized trading models as in Kandori (1992) and Milgrom, North, and Weingast (1990). Experiments with anonymous economies provide much needed empirical evidence to assess the validity of such theories.

Studying strategy adoption within this richer framework can offer novel insights about subjects' behavior in long-term interactions. Anonymity implies that strategies based on reputation, which have a strong drawing power, cannot be employed.² Hence, subjects are forced to consider other strategies. In economies with more than two persons, coordination on strategies and outcomes is more challenging. In addition, it is possible to study if and how a subset of subjects can separately coordinate on a strategy. Subjects are also exposed to a variety of

² In Camera and Casari (2009) subjects in a Non-anonymous public monitoring treatment adopt widely different strategies than in Private Monitoring treatment. In particular, the representative subject defected mostly only when interacting with those participants who previously defected with her.

behaviour, which facilitated the empirical identification of strategies. In this way, our work complements and innovates on previous studies on the identification of individual strategies.

The paper proceeds as follows: Section II discusses related works; Section III presents the experimental design; Section IV provides a theoretical analysis; Section V describes the empirical estimation procedure; results are reported in Section VI; and Section VII concludes.

II. Related experimental literature

There are a few experimental studies of individual strategies played in repeated games. They all 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. Though they model behavior using finite automata, as we do, there exist important differences both in the experimental design and in the empirical technique. Regarding the design, their subjects interacted in a supergame as partners for 5 periods in expectation, while our subjects interacted as strangers for 20 periods in expectation. As already mentioned, both features of our experiment facilitate the empirical identification of strategies. Regarding the underlying behavioral model, their automata choose action *stochastically* while transitions from state to state are *deterministic*. In other words, subjects can make mistakes in implementing actions. Instead, our automata choose actions *deterministically* and can transition from state to state *stochastically*. In other words, subjects experiment within the strategy, which can help moving an economy from a punishment mode to a cooperative mode. For example, if everyone follows a grim trigger strategy, an individual

mistake in implementing an action moves the economy to a permanent punishment mode. Instead, with experimentation within the strategy, it is possible to change state and revert to a cooperative mode. As a consequence, in Engle-Warnick et al (2004) the absorbing state of grim trigger is the punishment mode because a mistake is never forgotten. Our behavioral model instead allows for fresh starts and alternating sequences of cooperative mode and punishment mode. Moreover, they employ Bayesian methods and numerical techniques to estimate the distribution of strategies while we follow a straightforward maximum likelihood approach.

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.³

Dal Bó and Fréchette (2010) study an indefinitely repeated prisoner dilemma with an expected duration of two or four periods. They estimate individual 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-

³ Ule et al. (2009) also study a random matching setting, as we do. However, the game is finitely repeated, there is costly personal punishment, and the underlying game is a gift-giving game. Their focus is not on equilibrium strategies identification, but on the classification of subjects into types according to their behavior.

Warnick and Slonim (2006), they do not find evidence for grim trigger strategies in their economies of two players and of very short duration.

III. 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. 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.

[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 Murningham (1978). The interaction is of finite but uncertain duration, because in each period a cycle continues with a constant probability $\delta=0.95$. 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

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.

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.

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 \$26.22 for undergraduate subjects, \$40.15 for MBAs, and \$23.59 for clerical workers. 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.⁴

IV. 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 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

⁴ 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 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 $\delta=0.95$ and $\delta^*=0.443$.

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.

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 .

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.

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 $\delta = 0.763$.⁵

V. 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 rule of behavior in a repeated game (Rubinstein, 1986). The automaton is described by (i) a set of actions, (ii) a set of individual states, (iii) an outcome function that specifies the action to be taken given the individual state, and (iv) a transition function that specifies what individual state is reached next, given the current individual state and the actions of 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.”⁶ 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 an individual state, 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

⁵ 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 (2009) and Dal Bó and Fréchette (2010) present experimental evidence on how risk-dominance impact partners’ play in indefinitely repeated games.

⁶ 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.

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 an incorrect state with some probability $p \geq 0$. As a robustness check, we estimate strategy fitting from a range of values for p from 0 through 0.40. 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 grim trigger, a set of *forgiving* strategies, and a set of *opportunistic* 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”). Opportunistic 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 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*. 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).⁷ 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

⁷ 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.

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 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, $x_{q,1} = 1$. 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 $F(q)$ of a strategy q is the fraction of observations that q fits. The total fit $F(Q)$ of a set of strategies Q is the fraction of observations that can be explained by at least one strategy $q \in Q$.*

Both measures of total fit, of a single strategy and of a set, are useful. The total fit of a strategy q provides an upper bound for the fraction of subjects that employ strategy q .⁸ It is an upper bound, as subjects can be classified by more than one strategy because either they did not experience a sufficient variety of actions, or they played a short cycle. For instance, “grim trigger” and “tit-for-tat” identically fit the observation CD in a two-period cycle where the initial opponent plays D, and also the observation of a constant sequence C when the opponents always play C. As a consequence, the sum of total fit of all strategies in set Q , $\sum_{q \in Q} F(q)$, can exceed one, which it sometimes does in the analysis. The problem of overlapping strategies is particularly relevant when a subject observes the same action (e.g., C) in every meeting. In this case, we cannot infer what the subject would have done if D was observed, hence strategy identification is less meaningful. For this reason, the strategy-fitting procedure has been run separately for subjects who observed heterogeneous actions (Table 3).

VI. Results

There are seven main results. Result 1 concerns the strategy played by the average subject. Results 2-7 are about individuals, i.e., strategies employed by single subjects in each cycle.

Result 1. *Consider the behavior of the average subject. In period 1 she exhibited a high cooperation rate. If in the cycle she observed a defection, then she 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

⁸ This measure is context-independent, i.e., it is invariant to the number and type of strategies considered.

monitoring. It also confirms the empirical evidence presented in Camera and Casari (2009) and enhances the validity of those earlier results, by using a larger and more diverse subject pool.

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 (e.g., tit-for-tat). 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 the individual choice to cooperate (1) or not (0) using two groups of regressors. We introduce dummy variables that control for fixed effect (cycles, periods within the cycle, individuals), as well as for the duration of the previous cycle. To trace the response of the representative subject 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). For example, the “lag n ” regressor takes value 1 only in period $n=1, \dots, 5$ after the defection (0 otherwise). If the representative subject 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.⁹ 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. Instead, if the representative subject tended to revert to cooperation, then curves should be U-shaped.¹⁰

[Figure 2 approximately here]

The above analysis suggests that the representative subject acted as if playing a grim trigger strategy. Cooperation was the focal point of period 1 play for the representative subject. When first confronted with a defection in the match, the representative subject responded with an immediate, downward and persistent shift in the frequency of cooperation.

To expand on this initial assessment we carried out a disaggregated statistical analysis of subjects' strategies, proceeding as follows. First, we empirically identify strategies used by subjects and determine the fraction of observations that can be explained by those strategies. Second, we empirically characterize the strategies most commonly used. Third, we analyze the dynamics of individual behavior to understand whether they 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 fraction of individuals and (ii) which strategies are most

⁹ 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 through 5 is based on the sum of the marginal effects of the grim trigger regressor and the “lag n ” regressors with the appropriate lag.

¹⁰ Additional details on supporting evidence, including regression results, are in the supplementary appendix.

successful in doing so. A central result is that, given the identification technique proposed in the previous Section, a high fraction of individuals can be classified.

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

The empirical findings are reported starting with Figure 3, illustrating the fraction 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. Notice from Table 3 that no single group of strategies can classify more than 26.8% of individuals (“systematically cooperate”). 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. For instance, with $p=0.15$ more than 90% of individuals are classified. 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]

The best-fitting strategy is one of those in the “systematically cooperate” class and it classifies 37.6% of individuals.¹¹ When taking just two strategies into account, the total fit is 59.8%. 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

¹¹ The strategy coding is 11100. It is also the best-fitting strategy with $p=0.15$ with a fit of 40.6% of observations.

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 fraction of observations. To see why, consider that a set of just two strategies such as “grim trigger” and “always defect” fits 100% of observations of one-period cycles. This explains why in our experiment unclassified individuals played longer cycles (25.8 vs. 13.6 periods). The key issue is thus to determine under what dimensions unclassified and classified individuals differ.

[Table 3 approximately here]

Result 3. *Classified individuals exhibited less volatile play and higher average payoffs than unclassified individuals.*

Support for Result 3 is in Table 3 and Figure 4. 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; $N_1=96$, $N_2=406$).¹² Figure 4 plots each subject’s switch frequency and their opponents’ switch frequency. The circles’ size reflects the number of individuals associated to a specific switch frequency. Differences in switch frequencies of classified and unclassified individuals are not a simple effect of meeting different types of

¹² In this section statistics are computed aggregating observations by subject and cycle, unless otherwise noted. Thus we always have 500 observations in total. Statistical comparisons are done by means of a regression 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 (it takes value zero otherwise). The regression includes fixed effects at the subject level, and errors are computed clustering at the session level.

opponents. In addition, Table 3 reports that mean profits are significantly greater for classified than unclassified individuals (18.7 vs. 15.2; p-value is 0.014; $N_1=96$, $N_2=406$). This suggests that the two-state automata considered include the best-performing strategies.

[Figure 4 approximately here]

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 et 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.

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. Support for Result 4 comes from Table 3. There is a discrepancy between the behavior of the representative subject (Result 1) and the behavior at the individual level (Result 4).¹³ The experimental behavior recorded is compatible with the use of a grim trigger strategy at the level of the representative subject (Figure 1) but not when we consider individual behavior, because only a minority of disaggregated observations is compatible with the use of a grim trigger strategy. 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.

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

¹³ see Blanco et al., 2007 for a related result.

Table 3 displays a summary of results from the empirical identification of strategies. The data suggest that subjects acted as if following heterogeneous strategies. The four largest clusters of classified subjects acted as if having adopted a strategy from one of four main classes of strategies, which we denoted “systematically cooperate,” “systematically defect,” “forgiving,” and “grim trigger”. The most common behavior was consistent with “systematically cooperate,” though it leaves unexplained about half or more of the subjects. Interestingly, subjects adopting unconditional strategies greatly outnumbered those using conditional ones. Moreover, twice as many subjects 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 subjects 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, subjects’ behavior was more predictable in a stationary environment. The fraction of classified individuals grows from 78% (317 out of 406) in subsample (i) to 94% (88 out of 94) in subsamples (ii) and (iii). 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 subjects’ behavior, suggesting the existence of behavioral types. Only a minority of subjects acted as if using grim trigger (see also Offerman et al., 2001), while a significant fraction of participants

exhibited unconditional behavior, i.e., played an action that was fixed and independent of the opponents' actions. This suggests the existence of heterogeneous types among our subjects, which is in line with findings in the experimental literature on finitely repeated games (Houser et al., 2004), indefinitely repeated games (Engle-Warnick et al., 2004), and individual choice experiments (El-Gamal and Grether, 1995).

Up to this point we have analyzed data aggregated across cycles. Further information can be gathered by extending the analysis to study dynamic patterns of choices.

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

Recall that each participant 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 participant, then we say that strategy q classifies that participant. When we follow each participant across cycles, the data yields a very strong result. Only 19 out of 100 participants 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 participants 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 participants. 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 mean 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]

What can explain the persistent heterogeneity in strategy adoption found in the data? Unlike in two-person economies, in four-person economies a small group 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 participants always defect, then a subject who always cooperates earns more than the minmax payoff as long as the third participant 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.¹⁴ 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 even if not everybody else did the same in their economy.¹⁵ 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-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.

¹⁴ More precisely, 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.

¹⁵ More specific evidence comes from the following data. 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 who followed “systematically cooperate” is at least 15.3 points, which is higher than the minmax payoff of 10.

VII. Final Remarks

This study offers novel insights about subjects' behavior in decentralized trading environments where mutual gains from cooperation coexist with incentives to behave opportunistically. We designed experimental economies where four persons interacted locally and anonymously. Subjects faced an indefinite sequence of prisoner's dilemmas played in pairs but were not in a stable partnership as they were randomly rematched after every encounter. Because the interaction was anonymous, individual reputation was impossible to build and coordination was harder to achieve than in two-person economies, which have been the focus of previous experiments with indefinite interaction.

We empirically study equilibrium and strategy selection in supergames. The empirical analysis accounts for equilibrium strategies, such as grim trigger and unconditional defection, as well as non-equilibrium strategies, such as tit-for-tat and unconditional cooperation. The experiment facilitates 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).

Our conclusions about strategy selection crucially depend on whether the empirical analysis is conducted aggregating data from all subjects or at the individual level, disaggregating data subject by subject. At the aggregate level, results are compatible with the use of the grim trigger strategy. There is a strong initial attempt to coordinate on cooperation and defections triggered a permanent downward shift in cooperation levels. This finding is coherent with folk theorems. However, at the individual level, the conclusion is different: grim trigger is not the prevalent norm of behavior. In our experiment, subjects avoided the permanent, economy-wide punishment prescribed by grim trigger schemes.

Why were these schemes uncommon among strangers? Grim trigger is a theoretically appealing way to decentralize punishment because it prescribes permanent economy-wide defection, which is the harshest possible threat. However, the data show that there was substantial heterogeneity in strategy adoption, which persisted with experience. 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, 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 such 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.

These considerations point to a weak predictive power of theories based on homogeneous agents who adopt strategies of uncompromising, contagious punishment. They also suggest that empirical findings based on two-person experimental economies cannot be easily generalized to larger economies. In particular, the widespread adoption of grim trigger documented in two-person economies did not emerge in our four-person economies where reputation-based strategies were unavailable. In turn, 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. The possibility to resort to norms of decentralized, contagious punishment did not stave-off opportunistic behavior in our various subject populations. 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 further experimental and theoretical work lies ahead. On the one hand, it is necessary to empirically investigate whether the discontinuity in behavior observed when going from two-person to larger economies, is due to the anonymity of interaction, the lack of stable partnership, the size of the economy, or something else. On the other hand, a theoretical challenge remains, which is to increase the descriptive power of folk theorems.

References

- Aoyagi, Masaki, and Guillaume Fréchette.** 2009. "Collusion as Public Monitoring Becomes Noisy: Experimental Evidence." *Journal of Economic Theory* 144 (3), 1135-1165.
- Blanco, Mariana, Dirk Engelman, and Hans Theo Normann.** 2007. "A Within-Subject Analysis of Other-Regarding Preferences." Manuscript, Goethe University of Frankfurt.
- Blonski, Matthias and Giancarlo Spagnolo.** 2001. "Prisoners' Other Dilemma." SSE/EFI Working Paper 437.
- Blonski, Matthias, , Peter Ockenfels and Giancarlo Spagnolo.** 2009. "Equilibrium selection in the repeated prisoner's dilemma: axiomatic approach and experimental evidence." Unpublished manuscript, Goethe University Frankfurt.
- Camera, Gabriele, and Marco Casari.** 2009. "Cooperation among strangers under the shadow of the future," *American Economic Review*, 99(3), 979–1005.
- Chabris, Christopher F., David Laibson, Carrie Morris, Jonathon Schuldt, and Dmitry Taubinsky.** 2008. "Measuring Intertemporal Preferences Using Response Times", NBER Working Paper #14353.
- Dal Bó, Pedro.** 2005. "Cooperation under the Shadow of the Future: Experimental Evidence from Infinitely Repeated Games." *American Economic Review*, 95(5): 1591-1604.
- Dal Bó, Pedro, and Guillaume Fréchette.** Forthcoming. "The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence." *American Economic Review*.
- Diamond, Peter A.** 1982. "Aggregate Demand Management in Search Equilibrium." *Journal of Political Economy*, 90(5): 881-894.
- Duffy, John, and Jack Ochs.** 2009. "Cooperative Behavior and the Frequency of Social Interaction." *Games and Economic Behavior*, 66, 785-812.
- El-Gamal, Mahmoud and David Grether.** 1995. "Are People Bayesian? Uncovering Behavioral Strategies," *Journal of the American Statistical Association* 90(432), 1137-1145.
- Ellison, Glenn.** 1994. "Cooperation in the Prisoner's Dilemma with Anonymous Random Matching." *Review of Economic Studies*, 61: 567-88.
- Engle-Warnick, Jim, William McCausland and John Miller.** 2004. "The ghost in the machine: inferring machine-based strategies from observed behavior." Working paper 2004-11, University of Montreal.

- Engle-Warnick, Jim, and Robert L. Slonim.** 2006. "Inferring Repeated-Game Strategies from Actions: Evidence from Trust Game Experiments." *Economic Theory*, 28: 603–632.
- Friedman, James W.** 1971. "Non-Cooperative Equilibrium for Supergames." *Review of Economic Studies*, 38(113): 1-12.
- Houser Daniel, Michael Keane and Kevin McCabe,** 2004. "Behavior in a Dynamic Decision Problem: An Analysis of Experimental Evidence Using a Bayesian Type Classification Algorithm," *Econometrica*, Econometric Society, vol. 72(3), pages 781-822, 05
- Kandori, Michihiro.** 1992. "Social Norms and Community Enforcement." *Review of Economic Studies*, 59: 63-80.
- Kosinski, Robert J.** 2006. "A Literature Review on Reaction Time." Unpublished manuscript, Clemson University.
- Offerman, Theo, Jan Potters and Harrie Verbon.** 2001. "Cooperation in an Overlapping Generations Experiment", *Games and Economic Behavior* 36, 264-275.
- Palfrey, Tom R., and Howard Rosenthal.** 1994. "Repeated Play, Cooperation and Coordination: An Experimental Study." *Review of Economic Studies*, 61: 545-565.
- Roth, Alvin E., and Keith Murnighan.** 1978. "Equilibrium Behavior and Repeated Play of The Prisoner's Dilemma." *Journal of Mathematical Psychology*, 17: 189-98.
- Rubinstein, Ariel.**1986. "Finite automata play the repeated prisoner's dilemma." *Journal of Economic Theory* 39, 83–96.
- Rubinstein, Ariel.** 2007. "Instinctive and Cognitive Reasoning: a Study on Response Times." *The Economic Journal*, 177, 1245-1259.
- Ule, Aljaz, Arthur Schram, Arno Riedl and Timothy Cason.** (2009). "Indirect punishment and generosity towards strangers." *Science*, 326, 1701-1703.

Tables and Figures

(A) Notation in the theoretical analysis

Player 1/ Player 2	<i>Cooperate</i> (Y)	<i>Defect</i> (Z)
<i>Cooperate</i> (Y)	<i>y,y</i>	<i>l,h</i>
<i>Defect</i> (Z)	<i>h,l</i>	<i>z,z</i>

(B) Parameterization of the experiment

Player 1/ Player 2	<i>Cooperate</i> (Y)	<i>Defect</i> (Z)
<i>Cooperate</i> (Y)	25, 25	5, 30
<i>Defect</i> (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		168	00000		111
11110	<i>Systematically cooperate</i>	178	01000	<i>Systematically defect</i>	109
11101		166	00100		94
11100		188	01100		98
11000	<i>Grim trigger</i>	134	00001		26
11010		113	00010		11
11011	<i>Forgiving</i>	103	00011		5
11001		107	01001		25
10111		10	01010		13
10110		13	01011	<i>Opportunistic</i>	7
10101		13	00101		21
10100	<i>Opportunistic</i>	21	00110		9
10011		2	00111	3	
10010		9	01101	22	
10001		7	01110	11	
10000		24	01111		6

Table 2: Strategies and strategy sets

Notes: 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).

	p=0.05			p=0	
	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	18.6	134	19.92
-forgiving	130	4.33	20.52	90	22.44
-grim trigger	134	3.45	20.79	92	22.18
-opportunistic	44	4.35	14.4	28	15.2
D in period 1	133	2.56	19.67	92	20.5
-systematically defect	120	2.49	19.51	86	20.33
-opportunistic	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	15.8
D in period 1	31	4.2	15.33	72	16.74
Opponents play both C & D	406	3.5	17.12	406	17.12
<u>Classified</u>	317	3.41	17.65	182	18.13
C in period 1	201	3.84	16.9	105	17.26
-systematically cooperate	145	3.41	16.69	70	16.71
-forgiving	68	8	17.16	28	18.53
-grim trigger	72	4.85	17.86	30	18.02
-opportunistic	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
-opportunistic	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	25.06	60	25.06
-systematically cooperate	59	2	25	59	25
D in period 1	13	2.86	29.16	13	29.16
-systematically defect	10	5.01	30	10	30
<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	7.6	6.63	8	6.29
-systematically cooperate	8	8.05	5.96	5	5
D in period 1	4	2.37	9.64	2	10
-systematically defect	4	2.37	9.64	2	10
<u>Unclassified</u>	4	6.56	7.83	9	8.06
C in period 1	1	8	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., the subject's behavior. When no confusion arises we refer to such a sequence as one *observation*. There are 500 observations.

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					Total
	1	2	3	4	5	
All Observations						500
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
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						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

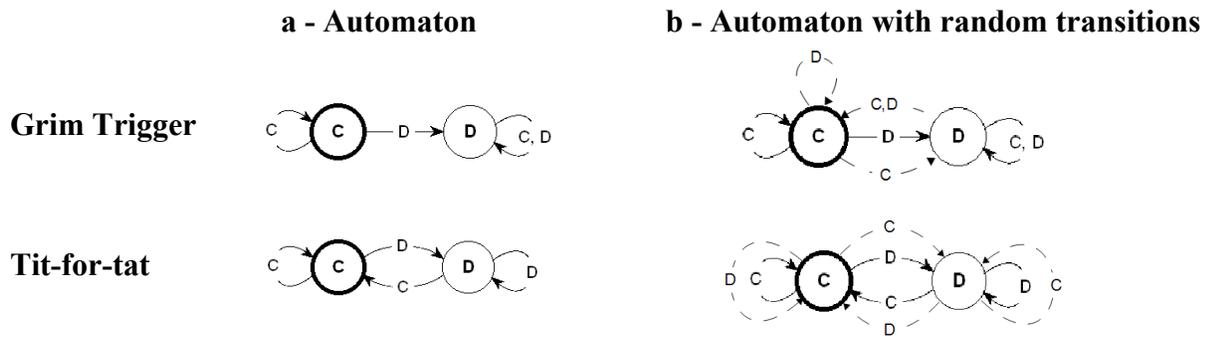


Figure 1: Strategy representation using automata (C=cooperate, D=Defect)

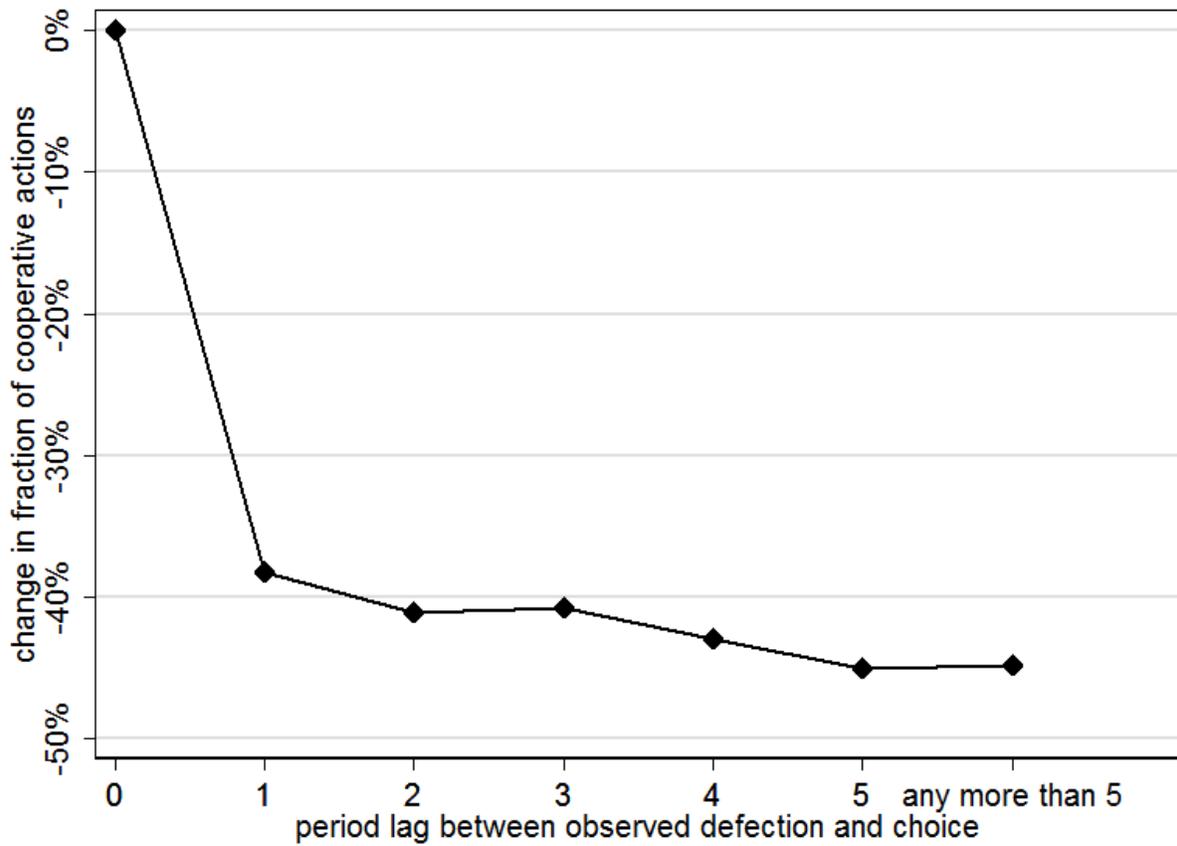


Figure 2: Aggregate response to an observed defection

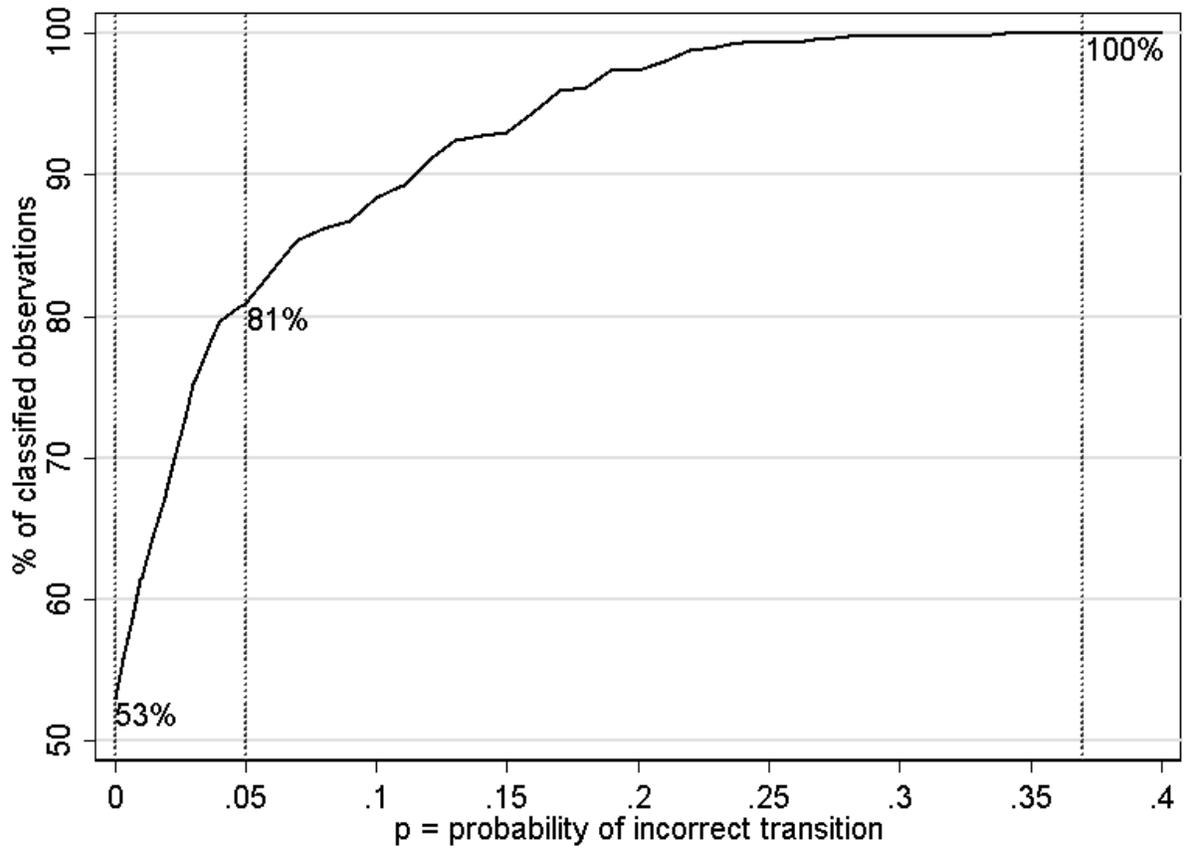


Figure 3: Fraction of classified observations

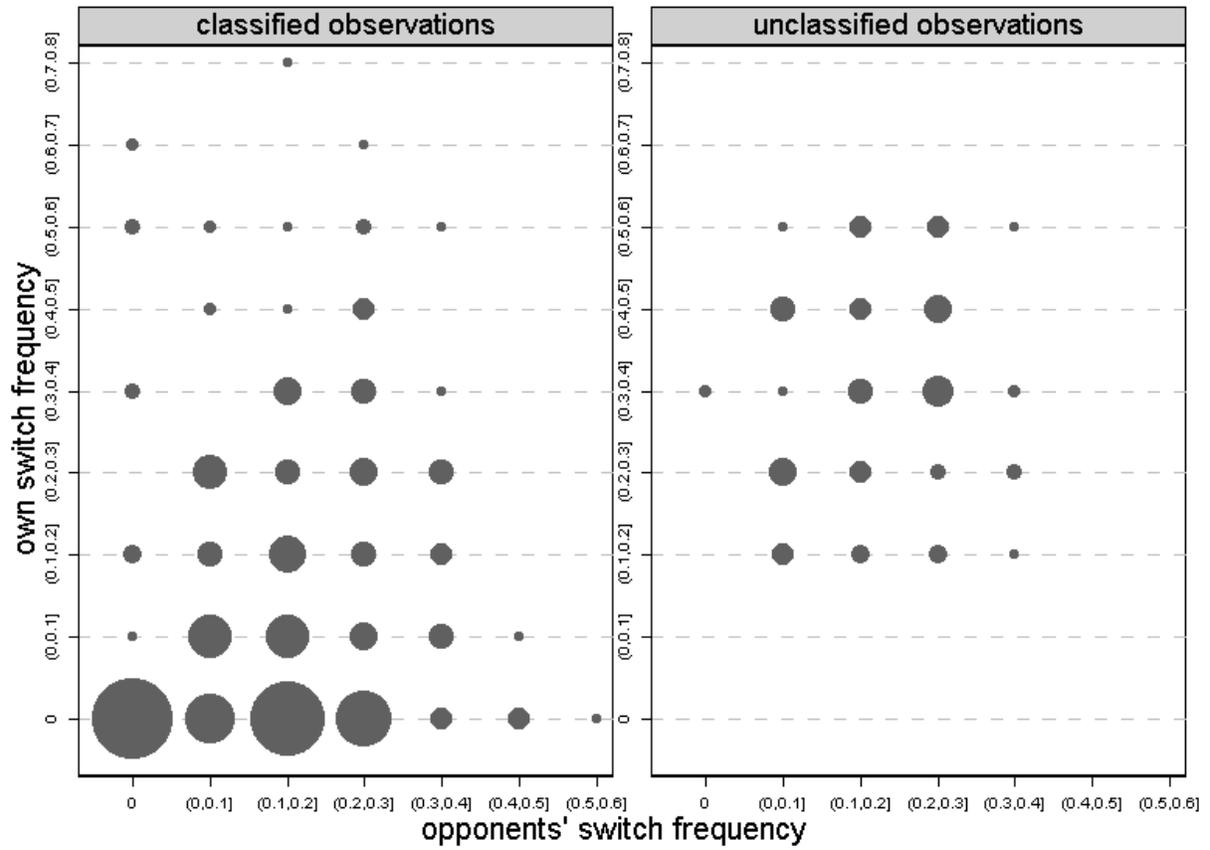


Figure 4: Volatility of play

Notes: 500 observations in total (405 classified, and 95 unclassified). Volatility of play at the individual level is defined as the *frequency of switch* between cooperation and defection choices. By definition, switch frequency is equal to zero for cycles lasting only one period.

Appendix

Appendix A

Table A1 reports the results from a probit regression that explains the individual choice to cooperate (1) or not (0) using two groups of regressors. First, we introduce several dummy variables that control for fixed effect (cycles, periods within the cycle, individuals), as well as for the duration of the previous cycle. Second, we include a set of regressors used to trace the response of the representative subject in the periods *following* an observed defection. For simplicity, we limit our focus to the five periods following an observed defection. This specification is more general than tracing behavior in periods 1-5 only, and it allows us to shed light on the type of strategy employed by the representative subject. Of course, there are several ways to choose regressors in order to trace strategies. Our specification has the advantage to detect whether subjects followed theoretically well-known strategies, such as grim trigger or tit-for-tat (Robert Axelrod, 1984). Indeed, we include a “grim trigger” regressor, which has a value of 1 in *all* periods following an observed defection and 0 otherwise. We also include five “lag n ” regressors, which have a value of 1 only in *one* period following an observed defection and 0 otherwise. For example, the “lag 1” regressor takes value 1 exclusively in the period after the defection (0 otherwise). The “lag 2” regressor takes value 1 exclusively in the second period following a defection (0 otherwise). And so on.

If the representative subject 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 following a defection, then the sum of the estimated coefficients of the grim trigger regressor and the “lag n ” regressors should be negative for the first and second period following a defection, and zero afterwards.

Figure 1 in the text illustrates the marginal effect on the frequency of cooperation in the periods that followed an observed defection.¹⁶ The focus on the five-period lags is for convenience in showing relevant patterns. 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 ” regressor with the appropriate lag. The L-shaped pattern of response to an observed defection suggests a persistent downward shift in cooperation levels immediately after a defection. The grim trigger coefficient estimate is significantly different than zero at a 1 percent level.¹⁷ While there is evidence that the representative subject employed a reactive strategy, not all observed actions fit this type of strategy.

¹⁶ Figure 1 is based on Table A1 using the coefficient estimates coding for reactive strategies. Zero-period lag is exogenously set at 0 percent. Marginal effects for the tit-for-tat regressors are computed for grim trigger regressor set at 1 (i.e. defection)

¹⁷ Table A1 reports that the actual length of the previous cycle influenced the propensity of participants to cooperate in period 1—the longer the previous cycle, the higher the cooperation level in the first period of the current cycle.

Table A1: Probit regression on individual choice to cooperate – marginal effects^(*)

<i>Dependent variable:</i>		
<i>I= cooperation, 0=defection</i>	All Periods	Period 1
<i>reactive strategies:</i>		
grim trigger	-0.448*** (0.051)	
lag 1	0.068*** (0.016)	
lag 2	0.039 (0.028)	
lag 3	0.042 (0.026)	
lag 4	0.020 (0.028)	
lag 5	-0.002 (0.032)	
<i>cycle dummies:</i>		
Cycle 2	0.104** (0.043)	-0.073* (0.039)
Cycle 3	0.044** (0.018)	-0.064** (0.032)
Cycle 4	0.108*** (0.042)	-0.101* (0.058)
Cycle 5	0.038 (0.046)	-0.080* (0.048)
duration of previous cycle	0.000 (0.001)	0.005** (0.002)
Observations	7440	500

^(*) Marginal effects are computed at the mean value of regressors. Robust standard errors for the marginal effects are in parentheses computed with a cluster on each session; * significant at 10%; ** significant at 5%; *** significant at 1%. For a continuous variable the marginal effect measures the change in the likelihood to cooperate for an infinitesimal change of the independent variable. For a dummy variable the marginal effect measures the change in the likelihood to cooperate for a discrete change of the dummy variable. First periods of each cycle are excluded (except in the last column). Individual fixed effects are included in columns “NP” and “PP”, and period fixed effects are included in all but the last column. These fixed effects are not reported in the table (individual dummies; period dummies: 3, 4, 5, 6-10, 11-20, 21-30, >30). Duration of previous cycle was set to 20 for cycle 1.