An Experimental Study of Imperfect Public Monitoring: Efficiency versus Renegotiation-Proofness*

Matthew Embrey
Maastricht University

Guillaume R. Fréchette
New York University

Ennio Stacchetti
New York University

October 28, 2013

Abstract

We study experimentally behavior in a repeated partnership game with public imperfect monitoring, and focus on whether subjects are affected by renegotiation concerns. The signal in our design is rather simple: it indicates only a success or a failure in each period. In some treatments, the equilibrium with the highest payoffs is renegotiation-proof, while in others it is not. Results indicate subjects' play is affected by the inclusion of a choice that permits some cooperation with more forgiving punishments, but that they do not play the renegotiation-proof equilibrium. However, when the renegotiation hypothesis predicts forgiving (short) punishments, subjects using cooperative strategies are indeed more likely to be forgiving. The experiment also reveals the use of strategies that have not been documented before, highlighting the importance of exploring different monitoring structures. Finally, our design includes communication, which we observe to be used to reduce strategic uncertainty.

*We wish to thank Emanuel Vespa and Sevgi Yuksel as well as seminar participants at the University of Pittsburgh, and conference participants at the European Economic Association, the Econometric Society (Oslo), and the ESA North American meetings for helpful comments and suggestions that improved the paper. We would further like to thank the National Science Foundation (grants SES-0924780 and SES-1225779), the CV Starr Center, and the Center for Experimental Social Science (CESS) for research support. We are responsible for all errors.

1 Introduction

In many economic interactions, individuals have the incentive to take actions that are privately profitable but collectively detrimental. This tension between group-interest and self-interest is a natural element of many environments – including problems of collective action and public goods, performance in teams and collusion in industrial organization – and often results in stark equilibrium predictions where agents end up in the least desirable outcome. The theory of repeated games, where the same interaction is played over a number of periods, provides an environment in which agents are able to circumvent this tension and sustain cooperation using inter-temporal incentives (Abreu et al. 1986, Fudenberg & Maskin 1986). Complementing this theoretical effort is a continuing empirical agenda investigating not only the extent to which subjects use such dynamic incentives to achieve greater cooperation, but also the strategies being used and the factors that might predict when the possibilities for cooperation will be exploited.

The experiment presented in this paper addresses these research questions by implementing in the laboratory a repeated partnership game. In the stage game, two agents choose an effort level. The probability that a joint project will be successful is determined by the sum of these efforts. Agents do not observe the effort choice of their partner. Instead, at the end of the period they discover whether or not the project was a success. Their payoffs, which are higher if the project is a success and decreasing in their own effort, are such that in a one-shot interaction only the lowest level of effort would be chosen by either party in equilibrium. More cooperative behavior in the repeated game can be sustained, using credible inter-temporal incentives, but requires phases of punishment as well as cooperation, with the former being triggered by failed projects.

Relatively little is known about how people behave in infinitely repeated games with imperfect monitoring. This paper adds in this dimension by introducing a monitoring technology that has not previously been studied. This monitoring structure, a binary success or failure signal, is one often considered in theoretical work and features naturally in the partnership game. Indeed the partnership game as a whole is new to the experimental literature, despite te fact it corresponds to many situations of interest. Finally, our experiment is the first to include communication in an infinitely repeated game with imperfect monitoring. The communication technology is highly structured, limiting the messages that can be sent. The aim is to ensure that communication is used for strategic considerations, to facilitate coordination and/or renegotiation.

The results suggest, despite the imperfect monitoring, a significant number of subjects act cooperatively, and do so by conditioning their behavior on the outcome of projects. Indeed, in the baseline games that have just two actions available, the efficiency of subjects' choices is consistent with the most cooperative super-game equilibrium. Our strategy analysis, however, reveals a degree of heterogeneity among subjects. While more use conditional strategies, a significant proportion repeatedly play the one-shot Nash action. The most commonly used conditional strategies have punishment phases that are either permanent reversion to Nash, monotone in the outcome, or belong to a family of counting strategies, which cooperate if there have been more successful outcomes than failures.[1]

The first two are natural analogues of the grim trigger and tit-for-tat, strategies commonly observed in prisoner's dilemma experiments with perfect monitoring. The counting strategies, which are not common in prior experimental studies, are reminiscent of the strategies predicted in favor exchange environments, such as Möbius (2001), Skrzypacz & Hopenhayn (2004), Hopenhayn & Hauser (2008). Overall, in our data, subjects tend to respond to the imperfect monitoring by adopting strategies that are more lenient. That is, they are inclined to let at least one failed project go by before entering a punishment phase. If they wish to implement a forgiving strategy in addition to leniency – that is, to transition out of a punishment phase at some point – they tend to do so by using one of the counting strategies.

The repeated partnership game implements several important features. First, the stage game has, in expectation, a prisoners' dilemma structure, which provides the tension between individual incentives and optimality for the group. Second, it implements an imperfect monitoring environment with a simple public signal structure. Such information environments are common in economic applications, such as oligopolies with demand shocks, principal-agent problems with unobservable effort, as well as partnerships and team interactions, as is more directly suggested by the game implemented in the experiment.

In repeated games, imperfect monitoring reduces but does not eliminate the scope for cooperation. Consequently, whether subjects can use dynamic incentives to support cooperation in such environments is of interest in itself. The imperfect monitoring game confronts the players directly with considerations that are not present in a perfect monitoring game. To sustain cooperation,

[1]Permanent reversion to Nash repeatedly plays the Nash action once the punishment state has been entered. A monotone punishment will transition out of the punishment state with sufficient success signals.

the players must be prepared to impose punishments when failures occur, an event that has positive probability even when both players cooperate. The severity of the punishments helps keep discipline among the players but it also decreases the total expected payoff of a cooperative strategy. Thus, while still being able to provide incentives for cooperation, the players would actually like to decrease the severity of punishments. As a result, these naturally occurring punishment phases provide a rich source of data on the means by which subjects support cooperative equilibria, as well as important learning mechanisms.

Furthermore, this setting highlights a potential inter-temporal consistency issue that has been raised in the theoretical literature on repeated-games: if agents are able to reach an initial implicit agreement, is it not also reasonable to presume that they can reach different implicit agreements at later stages? In particular, should they not be able to renegotiate away from a punishment phase, should it transpire that such a phase is encountered? If so, this calls into question the credibility of such punishment phases. Theories of renegotiation-proofness in repeated games attempt to resolve this inter-temporal inconsistency. As a treatment variable, our design introduces an alternative cooperation level, one that is less efficient but easier to support.[2] The manipulation of the available levels of cooperation permits a test of the extent to which such renegotiation-concerns might be a factor in determining cooperation rates.

To address the coordination problem inherent in situations of equilibrium selection, our design includes a structured pre-play communication mechanism before each stage of the repeated game. Pre-play communication has been found to facilitate the selection of more efficient equilibria in a number of, mostly static, experimental settings (for example, Cooper et al. 1992). By providing a means to resolve the coordination problem, the design gives cooperative equilibria their best chance of surfacing. The repeated communication stage also provides a means for renegotiation to emerge during a super-game, thus emphasizing the tension between efficiency and renegotiation-concerns in treatments where the latter are present.

Overall, efforts to cooperate appear to be aided by the communication mechanism in all our treatments. Most communication stages end in agreements that coordinate on the most efficient outcome. Even though they carry no commitment, most agreements are honored, especially those made in round

[2]Full cooperation requires the most severe form of punishment - permanent reversion to the one-shot Nash action - while a medium level of cooperation can be supported with a punishment phase consisting of just a single period of the Nash action.

1 or following a successful project. Should a pair not reach an agreement, the most likely choice in the subsequent stage-game is the one-shot Nash action. Despite correctly predicting a reduction in efficiency with the inclusion of the alternative cooperation level, we do not find strong support for the hypothesis that inefficient super-game behavior is the result of renegotiation concerns. In particular, subjects do not appear to switch to the renegotiation-proof strategy when predicted to do so, although we do find some evidence that they use forgiving strategies more often – that is, strategies less susceptible to being renegotiated during a punishment phase.

The paper is organized as follows. The next section provides a brief review of the related experimental literature. Section 3 sets up the repeated partnership game that is implemented in the laboratory and outlines the key predictions, while Section 4 provides details of the experimental procedures and Section 5 the results of the experimental sessions. Section 6 relates the results to prior perfect monitoring studies and other imperfect monitoring environments, and discusses the renegotiation-concerns hypothesis. A final section concludes.

2 Experimental Literature

Repeated games have received quite some attention in the experimental literature, especially with regard to the issue of equilibrium selection and coordination (see, for example, Fréchette & Yuksel 2013, and the cites therein). The results of earlier experiments in this literature suggest that subjects often fail to make the most of opportunities to cooperate.[3] However, more recent experiments (such as Dal Bó 2005, Duffy & Ochs 2009, Aoyagi & Fréchette 2009, Dal Bó & Fréchette 2011) find higher rates of cooperation. The key innovation in these later papers, and one that is included in the current design, is that subjects play the repeated interaction multiple times, as opposed to just once for money.[4] This allows subjects to learn to support coordination when it is incentive compatible.

Much of the experimental literature on repeated games has perfect monitoring. That is, subjects know the choices made in the previous period by the player they are matched with. Here, the information environment is one

[3]See Murnighan & Roth (1978), Murnighan & Roth (1983), Holt (1985), Feinberg & Husted (1993) and Palfrey & Rosenthal (1994).

[4]All the articles listed, in both the earlier and later experiments, share the same design feature with regard to how the infinite horizon is implemented in the lab: they all follow Roth & Murnighan (1978) in using random termination.

of imperfect public monitoring. Of the papers with imperfect monitoring, our experiment is closest in design to those of Aoyagi & Fréchette (2009) [AF] and Fudenburg et al. (2012) [FRD].[5]

AF implement a continuous, unidimensional public signal. Subjects' choices are implemented correctly but the signal includes a noise term, akin to a discrete Cournot quantity-choice game with stochastic demand. The stage payoffs, however, do not depend on the signal. Instead, subjects play the expected value game to circumvent the unbounded payoffs that are implied by the continuous signal.[6] They focus attention on strategies with two states in which transitions are governed by thresholds on the realization of the public signal. They find that subjects can cooperate even if monitoring is imperfect but fail to make the most of the opportunities to cooperate when cooperation is theoretically possible. They also find that subjects tend to use strategies that return to cooperation after a punishment phase.

In the repeated prisoners' dilemma game of FRD, the public signal is generated by implementing each player's choice with a small probability of a mistake. The signal is thus two-dimensional and binary. The payoffs in the stage game are determined by the outcome of the signals, that is the implemented choices. In their imperfect information environment, they find that strategies are more lenient and forgiving than in their perfect monitoring environment.

Notably, the signal in the FDR model discriminates among the players. For example, the signal (cooperate, defect) is far more likely to occur when player 1 cooperates and player 2 defects than when both players cooperate. Hence, to deter cheating when players are cooperating, player 2 should be punished after the signal (cooperate, defect). However, player 1 does not need to be punished. Thus, in their setting, it would be natural to consider asymmetric equilibria. By contrast, in our partnership model, the probability of failure is the same whether player 1 cheats or player 2 cheats. Hence, when they are cooperating, *both players* should be punished if a failure occurs. Thus, in our partnership environment, strongly symmetric strategies are more plausible.

The implications of renegotiation concerns for supporting cooperation and

[5]Work with imperfect monitoring also includes Holcomb & Nelson (1997), Cason & Khan (1999) and Feinberg & Snyder (2002). However, the nature of the monitoring environment is significantly different in these papers. Specifically Holcomb & Nelson (1997) and Feinberg & Snyder (2002) have private monitoring environments, whilst Cason & Khan (1999) have perfect but delayed monitoring. Rand et al. (2013) compare the case when there is perfect monitoring, despite noise in the implementation of choices, to their previous paper, FRD, where the noise in the implementation of choices generates imperfect monitoring.

[6]See Bigoni et al. (2011) for a recent experiment in a similar setting but with stage payoffs that do depend on the signal, although the objective there is to test a novel theoretical prediction regarding flexibility and cooperation.

selecting equilibria has been the subject of few experimental papers. Davis & Holt (1999), Andersson & Wengström (2012) and Cooper & Kuhn (2009) all investigate renegotiation in two-stage games, although the latter two do not repeat the same stage game. Instead, the second stage of their games are coordination games that have the interpretation of being the continuation game from infinitely repeating the first stage, if one restricts attention to subjects playing symmetric strategies. In this sense, their results speak to renegotiation in infinitely repeated games. As is the case in our design, Andersson & Wengström (2012) and Cooper & Kuhn (2009), who are directly interested in communication, both include intra-play communication. However, we do not use communication as a treatment variable. Instead a structured-communication format is chosen and maintained across treatments.

There are a number of differences between our design and the aforementioned experiments on renegotiation. First, the others implement a perfect monitoring environment. As argued earlier, we believe the imperfect monitoring environment provides a rich test bed for renegotiation concerns. Second, in our experiments subjects actually play the indefinite repeated game, rather than a two-stage game, bringing the experimental design closer to the theoretical setting.[7] More importantly, these two points together imply that subjects, should they follow the most cooperative equilibrium strategy, can expect to experience exactly the sort of paths of play that renegotiation concepts have focussed on. Finally, the experimental design manipulates game parameters, in particular the number of available actions, rather than communication, to systematically vary the possibilities for renegotiation concerns.

3 A Repeated Partnership Game

The theoretical framework for the experimental investigation is a repeated partnership game. In the stage game, two agents simultaneously make an action choice, which can be interpreted as an effort level. These actions have

[7] Fréchette & Yuksel (2013) provide evidence that subjects might not approach two-stage games as they would randomly terminated games, in that they do not necessarily condition behavior in the second stage game on the outcome of the first stage. However, these games do have a number of features that facilitate investigating directly the role of communication. For example, a super-game is much shorter on average, thus allowing for many re-matches of pairs even with communication protocols that might be time-consuming. Furthermore, there is a clear and even distinction between pre-play communication – i.e. before the first stage – and intra-play communication – i.e. before the second stage. Since we are not investigating directly the role of communication, such factors are not an advantage for our study.

an associated cost, borne by the individual agent. Once both agents have made their action choice, they observe the outcome of their joint project. This can be either a success or a failure (binary public signal). The probability of success depends on the choice of action made by both agents. The more costly actions increase the probability of success of the project. An agent's payoff from the interaction is composed of the payoff from the outcome of the project, $\overline{y}$ if a success and $\underline{y}$ if a failure, minus the cost of their action choice.

This stage game is repeated with an infinite horizon. Agents have risk neutral preferences and discount future payoffs by the common, constant rate δ. While agents observe the public signal, the outcome of the previous period's project, they do not observe the action choice made by their partner. Note that the stage game and the information environment of the repeated game are entirely symmetric.[8]

3.1 Main Parameters

Two sets of parameters, referred to as A and B, are considered. Throughout, a discount rate of $\delta = 0.8$ is used. Each set of parameters consists of three action choices and project payoffs. These actions are the *high* effort action, the *medium* effort action and the *Nash* action. The cost and probability of success for each of these actions, along with the payoffs if the project is a success or a failure, are given in Table 1 for both sets of parameters.

Note that the probability of success if both players choose the high action and if both players choose the Nash action are the same for both sets of parameters, at 0.9 and 0.1 respectively. The same is true for the cost of the Nash action, which is set to a value strictly larger than zero, and the gain in payoff from success compared to failure, which is set to 100. The payoff from failure is then set to ensure that the net payoff of choosing the high action when the project is a failure is still greater than zero.

Using these actions, two games are constructed for each set of parameters: the *2-action* game, in which only the high and Nash actions are available, and the *3-action* game, in which all three are available. The choice of action made by the two agents determine their *ex-ante* expected payoff, before learning the outcome of the project. Table 2 gives these payoffs for each combination of the two players' choices. Viewing the strategic interaction in this manner makes the underlying prisoners' dilemma structure transparent. The joint expected utility maximizing choice is for each player to take the high action. However,

[8]See, for example, Mailath & Samuelson (2006) chapters 7 through 11 for a detailed treatment of repeated games with an imperfect public monitoring environment.

Parameter set A: $\overline{y} = 166$ and $\underline{y} = 66$						
		Payoff if:		Probability of success if other chooses:		
Action	Cost	success	failure	High	Medium	Nash
High	65	101	1	0.9	0.6	0.56
Medium	41	125	25	0.6	0.6	0.32
Nash	10	156	56	0.56	0.32	0.1
Parameter set B: $\overline{y} = 163$ and $\underline{y} = 63$						
		Payoff if:		Probability of success if other chooses:		
Action	Cost	success	failure	High	Medium	Nash
High	62	101	1	0.9	0.7	0.61
Medium	50	113	13	0.7	0.7	0.34
Nash	10	153	53	0.61	0.34	0.1

Table 1: Action choices

Parameter set A						
2-action game			3-action game			
	Other's action:			Other's action:		
Action	High	Nash	Action	High	Medium	Nash
High	(91, 91)	(57, 112)	High	(91, 91)	(61, 85)	(57, 112)
			Medium	(85, 61)	(85, 85)	(57, 88)
Nash	(112, 57)	(66, 66)	Nash	(112, 57)	(88, 57)	(66, 66)
Parameter set B						
2-action game			3-action game			
	Other's action:			Other's action:		
Action	High	Nash	Action	High	Medium	Nash
High	(91, 91)	(62, 114)	High	(91, 91)	(71, 83)	(62, 114)
			Medium	(83, 71)	(83, 83)	(47, 87)
Nash	(114, 62)	(63, 63)	Nash	(114, 62)	(87, 47)	(63, 63)

Table 2: Ex-ante expected payoffs

each has an incentive in the one shot game to deviate to the Nash action and free-ride on the other's effort. In the one-shot game, both players choosing the Nash action is the equilibrium outcome in dominant strategies.

While repeated play of the one-shot Nash action is always an equilibrium of the infinitely repeated game, more cooperative equilibria can be supported by providing inter-temporal incentives. The cooperative equilibria of the repeated game are analyzed using the perfect public equilibrium concept. An equilibrium is a public perfect equilibrium if i) all strategies of players depend only on the public history (that is the outcomes of the previous periods), and ii) after any possible public history the equilibrium prescribes strategies for the players that form a Nash equilibrium of the repeated game starting from that period onwards.

Two further restrictions are added to this equilibrium concept. First, attention is restricted to the simplest class of strategies that can support cooperation. These are two state automata, in which players play a prescribed action in each state. The states are labeled reward and punishment, with the former having an expected value, given being in that state, at least as high as the latter. Transition between states is solely a function of the signal of the previous period. Second, given the symmetric nature of the game under consideration, attention is restricted to equilibrium strategies in which players play the same action after any history of signals (strongly symmetric strategies). An equilibrium that achieves the maximum expected value in the reward state, among all perfect public equilibria using strongly symmetric two-state automata, is referred to as the maximum perfect public equilibrium. This is denoted by just PPE since it is the focus of the equilibrium predictions without any concerns for renegotiation.

In all four games, the maximum expected payoff in the reward state can be achieved by supporting the high action with permanent Nash reversion. This is a strategy where, in the reward state both players fully cooperate and continue to do so until the first realization of a failure signal. Following the failure signal, the players move to the punishment state where they both play the Nash action forever. The automaton representation of the strategy is given in Figure 1(a) and is referred to as Grim HN. As will become clear from the analysis in the subsequent subsection, the severity of this Nash reversion punishment is in fact necessary. More forgiving strategies, such as leaving the punishment state following the realization of a high signal, are not incentive compatible.

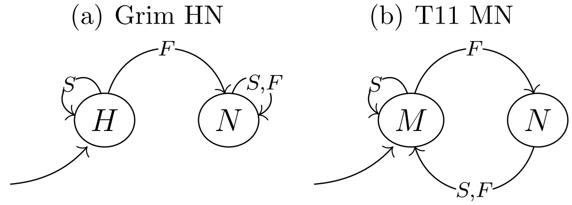


Figure 1: Automaton representation of predicted strategies

A key difference between the perfect and imperfect monitoring environments is the need in the latter to punish even in a cooperative equilibrium. That is, to maintain correct incentives for future cooperation, players should enter a punishment phase with strictly positive probability even when they both believe each other to have chosen the cooperative action. This feature of cooperative equilibria make theories of equilibrium selection based on the severity of the required punishment natural candidates for such imperfect monitoring environments. Furthermore, this setting highlights an inherent time-consistency issue: if players are able to reach an initial agreement, is it not also reasonable to presume that they would also agree to renegotiate away from severe punishment phases, should such phases be encountered at later stages.

For the renegotiation-proof prediction, we specialize the concept of Pearce (1987) to the current environment: a candidate equilibrium would survive potential renegotiation at later rounds if there is no other perfect public equilibrium, using strongly symmetric two-state automata, that has a larger expected value in the punishment state. Such an equilibrium of the repeated game is denoted by RE. The driving force of this approach to renegotiation is, as noted in Abreu et al. (1993), that players are not concerned with whether a punishment was needed to deter deviations in the past, but whether it is unavoidable to support any future equilibrium. In this sense, this renegotiation-proof concept selects an equilibrium that does not suffer from this inherent time-consistency issue.

The RE prediction differs from the PPE prediction in that, among the incentive compatible machines, it maximizes the punishment state value, rather than the reward state value. In the 2-action games, since all more forgiving punishment strategies cannot support the high action choice, and the only other choice is the Nash action, the PPE and RE concepts coincide. In the 3-action games, the medium action can be supported in the reward state by using a less severe punishment than Nash reversion. Consequently, the PPE and RE concepts will diverge for these games. The strategy predicted by the latter in this case plays the medium action in the reward state, and continues this partial cooperation until a failure signal is observed. Upon observing a failure, the strategy plays the Nash action for exactly one period before switching back to the reward state. The automaton representation of the strategy is given in Figure 1(b) and is referred to as T11 MN. [9]

Farrell & Maskin (1989) proposed an alternative renegotiation-proof refine-

---

[9] T11 because it is a Trigger strategy, it takes 1 bad signal to move to punishment, and the punishment phase lasts 1 period only. MN because it plays M in the reward state and N in the punishment state.

ment for repeated games (see also Bernheim & Ray 1989, Benoit & Krishna 1985). Within the family of finite automata, a subgame-perfect equilibrium is *weakly renegotiation-proof* if the expected values of any two of its states are not Pareto ranked. The idea here is that if the expected value of of one state is Pareto dominated by the value of another state, both players would have an incentive to renegotiate away from the state with the Pareto inferior payoff. In our partnership model, the only weakly renegotiation-proof equilibrium in strongly symmetric strategies is Nash forever, which supports no cooperation. As mentioned in the introduction, we find a significant number of subjects acting cooperatively in all treatments.

3.2 Incentive Compatibility of Predicted Strategies

Consider first the case of parameter set A and the PPE prediction. When the players follow the strategy Grim HN, each expects a total discounted value of $V_P = 66$ in the punishment state, where they both choose the Nash action in the current and all future periods.[10] The value in the reward state is then $V_R = 83.857$.[11] Once in the punishment state, the players play the repeated one-shot Nash strategy and clearly have no incentives to deviate.

In the reward state, it is necessary to check that player 1, for example, has no incentive to deviate. By the "one-shot deviation" principle, it is enough to check that player 1 does not profit should he deviate once and only once. If he chooses Nash instead of high and never deviates from Grim HN again, he gets 83.2, which is less than $V_R = 83.857$, the payoff when he does not deviate.[12] Note that the incentive constraint is relatively "tight" – the expected value of a deviation is just below the equilibrium payoff. Consequently, the two-state automata Grim HN delivers "almost" the weakest possible punishment that supports full cooperation in the reward state.

The incentive constraint at the reward state can be re-written as

$$(1 - 0.8)[112 - 91] \leq 0.8[0.34(V_R - V_P)]$$

to illustrate that the incentives to play high are weakened if the value of punishment V_P is increased. If the value of punishment were larger than 69.38, the high action would no longer be supported in equilibrium. Thus, it

[10]Note that the expected values reported in this analysis of incentive compatibility are all normalized by $(1 - \delta) = 0.2$, as is common in the analysis of repeated-game equilibria. Thus, if a player receives a sequence of payoffs $\{\pi_t\}$, then the total discounted (normalized) payoff is $(1 - \delta) \sum_t (\delta)^t \pi_t$.

[11]Explicitly, $V_R = (1 - 0.8) \times 91 + 0.8[0.9V_R + 0.1V_P]$

[12]Explicitly, $(1 - 0.8)112 + 0.8[0.56V_R + 0.44V_P] = 83.2$

is not possible to sustain full cooperation (play high) with a more forgiving strategy that can be represented by a two-state automata. A strategy that supports the high action with the weakest possible punishment can only be implemented with an eight-state automata.[13]

To complete the incentive compatibility check of the Grim HN machine, if player 1 chooses medium instead of high and never deviates again, he gets $78.371 < 83.857$.[14] Therefore, no deviation is profitable and Grim HN is a perfect public equilibrium. The analysis for the strategy T11 MN is similar, as is the analysis for parameter set B. Table 3 summarizes the relevant values for both strategies in parameter sets A and B.

		Parameter set	
Automaton	State	A	B
Grim HN	Reward	83.857	83.000
	Punishment	66.000	63.000
T11 MN	Reward	80.394	79.129
	Punishment	77.515	75.903

Table 3: Expected values of predicted automata

4 Experimental Design

The experiments were conducted at NYU's Center for Experimental Social Science using undergraduate students from all majors recruited via e-mail. Instructions were read aloud to students and they interacted solely through computer terminals.[15] The basic design of the experiment is as follows: Subjects are randomly matched into pairs for the length of a repeated game, referred to as a *match*. In each round of a match subjects play the same *stage game*, which is the repeated partnership game described in Section 3. The length of a match is randomly determined to replicate an infinite horizon with discounting environment. The probability that a match will continue for at

[13]In this strategy, a failure in the reward state is punished by seven periods of Nash reversion. For this machine, $V_R = 84.755$, and the expected value at the state when the punishment is just beginning is $V_P = 69.108$.

[14]Explicitly, $(1 - 0.8)85 + 0.8[0.6V_R + 0.4V_P] = 78.371$

[15]See Appendix A for sample instructions. The computer interface was implemented using zTree (Fischbacher 2007).

least one more round is 0.8.[16] After a match terminates, subjects are given detailed feedback: for each round, they are shown their choice, the choice of their partner and the realization of the public signal. Pairs are then randomly rematched for the next match. A session consists of 10 such matches.

4.1 Communication

In addition to the above, matched partners were given the opportunity to engage in a structured communication stage before making their action choices. Although the theory does not indicate a direct role for communication in order for renegotiation concerns to bite, introspection suggests that having the opportunity to communicate in future periods may be behaviorally important for such concerns to emerge. Indeed the theory of repeated games in general is ambivalent about the role communication plays in super-game equilibria, irrespective of renegotiation. Although the theory does not point to an explicit role, communication is often cited as a facilitating factor for cooperation, especially for collusion in antitrust applications. Consequently, communication is included and allows these two countervailing forces to come into play: on the one hand the role communication can play in coordinating on more cooperative equilibria, and on the other, the role it can play in bringing renegotiation concerns to the fore.

The structured communication stage was implemented as follows: One of the two players in a pair was chosen at random to send the first message. They then took turns to exchange messages. The subjects were given the following list of messages to send:

- "No message";
- "I propose that you choose X and I choose Y", where X and Y were picked from the list of available action choices;
- "I propose that you choose X and I choose Y. And if the outcome is high in the next round, you choose X1 and I choose Y1. And if the outcome is low in the next round, you choose X2 and I choose Y2", where X, X1, X2 and Y, Y1, Y2 were picked form the available action choices;
- "Agree to the proposal."[17]

[16] This choice trades off a number of competing forces: A larger continuation probability leads to longer matches, providing more potential reward and punishment phases within the same match. However, longer matches reduces the number of re-matches that can feasibly be done within a single experimental session, thus reducing experience.

[17] Note that this option was only available if a message containing a proposal was sent beforehand.

Communication ended in one of two ways. Either, once both players had the opportunity to send at least one message, the first "No message" or "I accept the proposal" ended the exchange. Alternatively, after both players had the opportunity to send two messages, the player who sent the first message was asked to respond to the other player's last message with either "No message" or "I agree to the proposal". This then ended the exchange. Consequently, both players had the opportunity to send at least one message, while there were at most two rounds of communicating back and forth. Furthermore, at all times during the message exchange, subjects had the option to send no message.

This sequential message structure was chosen to avoid potential coordination problems that could arise if messages were sent simultaneously, in particular, the issue of how subjects would interpret simultaneously sending contradicting messages. There are two reasons for choosing a structured message protocol instead of a free form chat protocol. First, the structure gives the experimenter greater control over what subjects are communicating over. The selected message list focusses on what actions subjects are going to choose and how to support such choices with future choices. Second, the structured message format provides a ready-coded data set for analyzing what messages were sent and the subsequent actions taken. Finally, it should be emphasized that the messages carry no commitment technology. Once the message exchange has been completed, subjects are free to make what ever choice they feel is in their best interests.

4.2 Actions in the Stage Game

To systematically vary the possibilities for renegotiation concerns, two stage games are implemented, differing only in the number of actions available. The control treatment is the 2-action game, where only the high and Nash actions are available. Here the PPE and RE concepts coincide, which is to support high cooperation using the most severe form of punishment, namely permanent Nash reversion. This treatment provides a benchmark for the ability of subjects to coordinate on an equilibrium without the impact of renegotiation concerns.

The treatment variation is the inclusion of the medium action, which generates the 3-action game. This medium cooperation level, while maintaining the underlying prisoners dilemma structure,[18] gives subjects the option of a level of cooperation that is easier to support, in the sense that it does not need as

[18] In particular the binding incentive constraints are deviations to the Nash action rather than another cooperative action that is less costly (for example deviating from the high action to the medium action).

severe a punishment. Under the RE prediction, full cooperation is no longer sustainable since the Nash-reversion punishment is no longer renegotiation-proof. Consequently, the PPE prediction, which is the same as in the 2-action game, and the RE predictions diverge. The latter prediction is that subjects will cooperate on the medium cooperation level, using a forgiving-Nash punishment to support it. Thus, the comparison between the 2-action and 3-action game provides the test for the impact of renegotiation concerns on the ability of subjects to cooperate.

4.3 Parameter Sets

As detailed in Section 3, two sets of parameters are considered, each with an associated 2-action and 3-action game. Implementing two sets of parameters provides a more robust investigation since the observations are not entirely dependent on one choice of parameters. Furthermore, there are two elements to providing dynamic incentives in any cooperative equilibrium of this environment. The first is to incentivize paying the private cost of higher effort. The second is to detect deviations from cooperation. The two parameter sets differ in the relative importance of these two elements in supporting the high effort choice: the private cost is higher in parameter set A than B, yet the probability of success if the other player deviates is lower - it is easier to detect deviation. As a consequence, the sucker payoff of the 2-action game in the latter parameter set is almost as high as the Nash payoff.[19] Thus, should fear of mis-coordination be a primary concern for supporting cooperation, parameter set B gives cooperation its best chance, since the difference between the sucker payoff and the Nash payoff is minimal.

A further difference between these parameter sets is the value of the sucker payment for different levels of cooperation. In parameter set A, this is fixed at 57 for all levels of cooperation; in parameter set B it is actually lower for the middle level of cooperation (values of 62 and 47 for the high and medium cooperation levels respectively). In fixing the sucker payments, the first set of parameters ensures that a factor, which could be behaviorally relevant for supporting cooperation, is kept constant across the cooperation levels.[20] In

[19] The sucker payment refers to the ex-ante expected payoff for a player that attempts to support a cooperative effort level, whether that be high or medium, but faces a partner that chooses the Nash action. This is common terminology from the prisoners' dilemma game, which is adapted here to fit the context of the repeated partnership game.

[20] This factor is referred to as behaviorally relevant, since the sucker payoff does not enter the equilibrium analysis outlined in Section 3. This is because only the "temptation" payoff (i.e. what a player receives when they deviate from a cooperative action pair) is relevant

this sense, the first set of parameters can be considered a weaker test of the potential impact of including less efficient, but also less punishing to support, cooperation level.

5 Results

For each treatment, three sessions were conducted, indexed one to three. For each index, the random round-match composition was drawn once so that, for a given session number, the round-match composition is the same across all treatments.[21] On average a session lasted between an hour and half and two hours; subjects earned on average between \$35 and \$43 for their participation. Each subject faced only one set of parameters (between-subjects design). Each session had 12 to 16 subjects.[22] The results are presented in four sections dealing with cooperation rates in general, the role of communication, the evolution of cooperation, and the estimation of strategies, respectively.[23]

5.1 Cooperation in the Repeated Partnership Game

One measure of the extent to which subjects are cooperating is expected payoff efficiency - the expected joint-payoff of a match, given the choices made by subjects, minus the expected joint-payoff from repeated one-shot Nash play divided by the difference in expected joint-payoffs between repeated full cooperation and repeated one-shot Nash. Figure 2 reports the average of this efficiency measure by treatment, along with the predictions from the theory. The left-hand panel reports data and predictions from all rounds, which corre-

for checking incentives to deviate from a cooperative equilibrium.

[21]See Table 17 of Appendix C.2 for a breakdown of the round-match composition by session.

[22]See Table 18 of Appendix C.2 for a summary of the treatments. One further treatment was run using the B set of parameters. In this benchmark game, a low action was included with the medium and Nash actions. The low action is a redundant action that is neither more efficient nor less punishing to support. The aim was to further test the renegotiation-proof hypothesis should behaviour in the 3-action game be consistent with it – in particular, should the medium action prove to be a popular cooperative action. However, as will be detailed in what follows, this was not the case. Consequently, these sessions are not reported in the main text. Results are available upon request.

[23]Throughout the results section, tables report data for the last five matches, unless otherwise explicitly stated. The first half of the matches are excluded to allow for adaptation that is commonly observed at the very beginning of an experiment. All reported regressions and statistical tests use cluster-robust standard errors, corrected for arbitrary correlation at the session level. See Fréchette (2012) for a discussion of session-effects.

sponds to the initial/reward-state prediction of the theory. Subjects are able to earn significantly more than that expected from always playing the one-shot Nash equilibrium, which has a zero efficiency rate. Efficiency rates are consistently below 100%, which would result if subjects always played high, a strategy that is not incentive-compatible. Over all rounds, the observed efficiency rates are very similar to those predicted by the RE concept. In particular, there is a drop in efficiency when the medium action is introduced, as predicted by the RE.

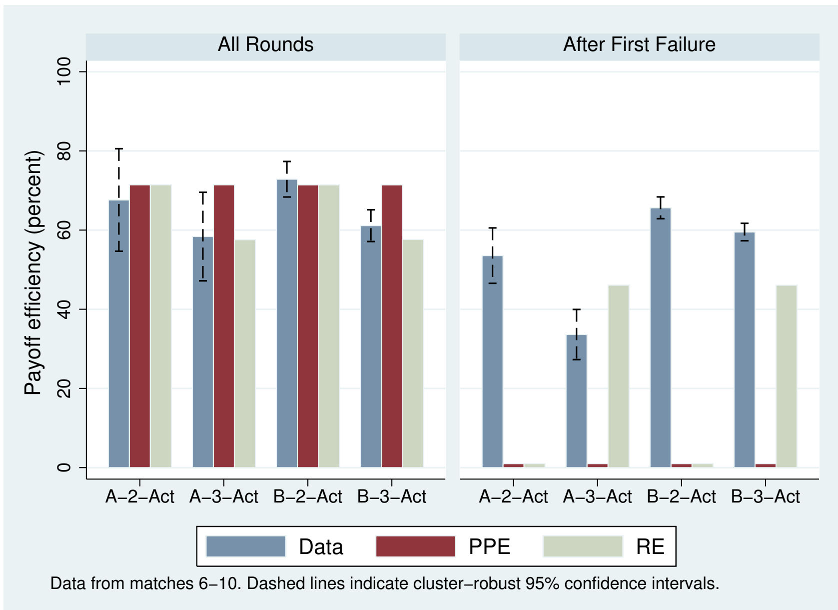


Figure 2: Expected payoff efficiency: data versus predictions

The right-hand panel from Figure 2 reports efficiency rates averaged over rounds that follow the first failure within a match. This sub-sample corresponds to the punishment-state prediction from the theory. As is clear from the figure, both the PPE and RE concepts fail to explain the observed pattern of behaviour for this sub-sample. First, subjects achieve significantly higher expected efficiency rates in the 2-action games than the zero predicted by both theories. Furthermore, efficiency is actually lower in the 3-action games than the 2-action games when the RE concept predicts that efficiency rates should increase in the punishment states of the 3-action games.

Figure 3 gives the frequency of each action choice across the four treatments, separated into choices in the first round of a match and choices in the later rounds of a match. The majority of round-one choices are cooperative (either medium or high) in every treatment, although more than a third of them are non-cooperative, suggesting a degree of heterogeneity in the way subjects approach this environment. If subjects played the payoff-maximizing equilibrium, then all round-one choices would be high. At the other extreme, the equilibrium with the lowest payoff would have all round-one choices being Nash. The renegotiation hypothesis would instead predict high round-one choices in the 2-action games and medium round-one choices in the 3-action games.[24] None of these hypotheses fully explain round-one choices. Although there is a decrease in the use of the high action when comparing the 2-action and 3-action games, the majority of round-one behavior is consistent with either the payoff maximizing equilibrium or the repeated-Nash equilibrium. The increase in Nash action choices generally observed after the first round is consistent with the provision of dynamic incentives, which requires punishment phases following failures.

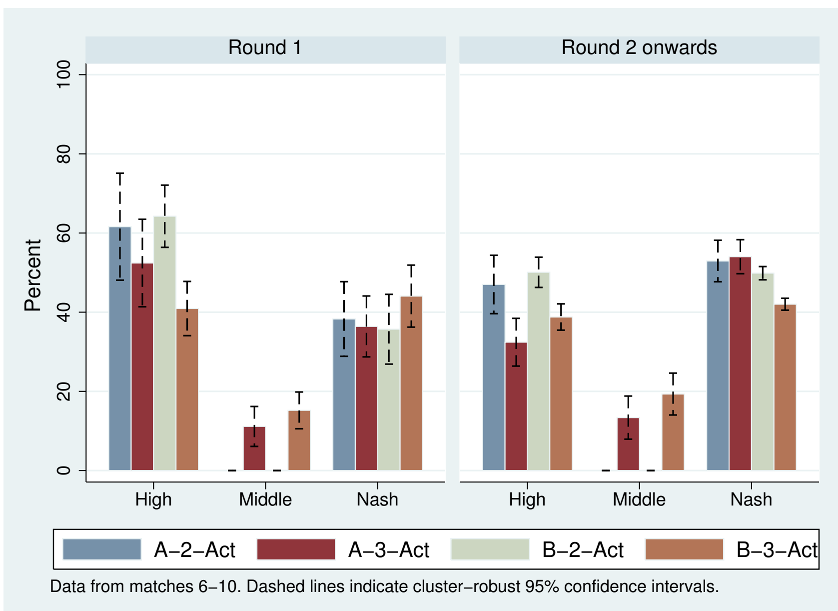


Figure 3: Summary of choice frequencies

[24]In all treatments, the simplest strategies that support some cooperation involve the use of only 2 choices. This is the case in the majority of matches, but for 14% of the subject-matches where all three choices are used.

5.2 Messages

Subjects are given the opportunity to exchange messages before every action choice. Table 4 summarizes the outcomes from the communication stage, which could end in no message (none), an agreement of the do-now format (1-part) or an agreement including a specification of what to do in the subsequent round conditional on the outcome (2-part). Even after a failure, the majority of communication ends in an agreement of some variety. These agreements are usually of the do-now format, especially in the 3-action games, and tend to be cooperative. In particular, agreements are overwhelmingly symmetric and involve making the highest effort choice, irrespective of whether the agreement was made in round 1, or in a later round following a failure or a success (see Table 19 in Appendix C for further details). Communication stages ending in no message are most prevalent in rounds following a failure.

Message sent in/after	Message type			Agreement followed if			Choice if no agreement		
	None	1-Part	2-Part	High	Med.	Nash	High	Med.	Nash
				A-2-Action					
Round 1	19	53	27	74		89	26		74
Failure	39	42	19	38		93	21		79
Success	17	57	26	77		100	43		57
				A-3-Action					
Round 1	15	71	15	68	54	89	25	22	53
Failure	34	58	7	31	47	89	10	13	77
Success	12	81	7	66	62	94	26	21	52
				B-2-Action					
Round 1	8	65	28	70		100	31		69
Failure	27	53	19	44		96	24		76
Success	16	67	17	70		88	44		56
				B-3-Action					
Round 1	19	62	18	57	81	91	15	17	67
Failure	29	63	9	48	74	96	16	9	75
Success	14	81	5	67	92	100	30	13	57

Table 4: Further details of messages (in %) in the last five matches

Even though the communication stage carries no commitment technology, messages appear to be informative about intended action choices. Agreements to play cooperative actions are generally honored when they are made either at the beginning of a match or following a successful outcome. This is no longer the case following a failure. Communication stages that end without an agreement are likely to be followed by a Nash action choice.

In every round, each pair of subjects had at least two communication stages, and no more than five, to reach an agreement or end communication with a no-message. Figure 4(a) shows how many stages on average subjects took according to whether they were in round 1, a round following a failure or one following a success. Both during the first five matches and the last five matches, subjects required little more than the minimum two stages on average in which to conclude their communication, and they became shorter with experience. However, irrespective of experience, communication is longer following a failure than a success.

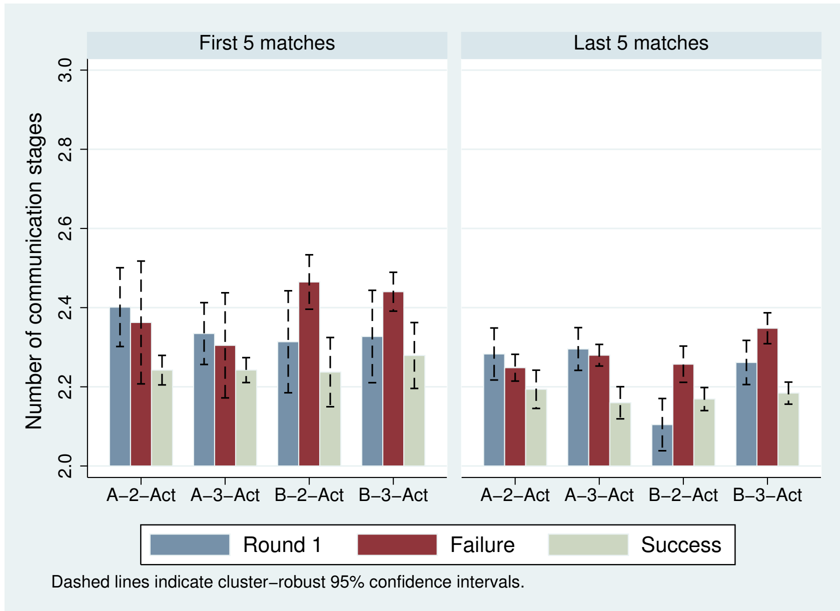
Although subjects in the renegotiation-concerns hypothesis, through a process of introspection, would never actually renegotiate, the hypothesis does identify histories from the PPE equilibrium that should lead to renegotiation efforts. These histories include those where the reward path of the PPE equilibrium had been followed until the previous round, at which point the outcome turned out to be a failure. Figure 4(b) restricts attention to observations consistent with a subject following the reward path of the PPE equilibrium prediction prior to the outcome of the last round.[25] As can be seen from the left-hand panel of Figure 4(b), there is some evidence that such histories – those following a failure – identify the rounds in which communication took longer during the first five matches.[26] However, by the second half of the experiment, the length of communication, as a function of the outcome in the previous round, is similar irrespective of whether the prior behaviour was consistent with the PPE reward path or not. Furthermore, in these later matches, the content of agreements is also similar (see Tables 19 and 21 in Appendix C).

5.3 Evolution of Cooperation

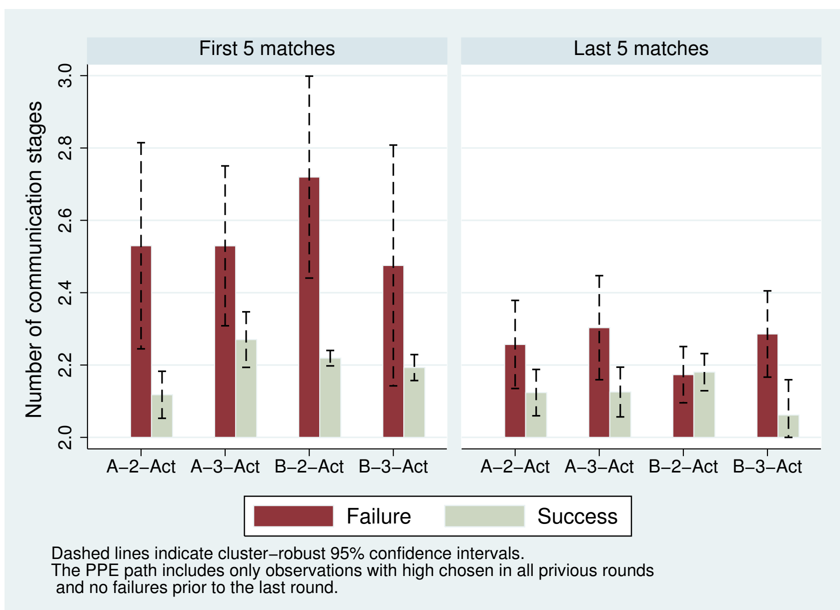
Figure 5 shows the evolution of high choices in the 2-action games, and the evolution of high and medium choices in the 3-action games. The top panel graphs the percentage of round 1 choices for each action across matches, while the bottom panel graphs the percentage of choices across rounds, aggregating

[25]That is, observations in which the subject has chosen high in every prior round and the outcome was a success in every prior round except possibly the last round.

[26]Note that there is less data for PPE reward path histories with a prior failure outcome than with a prior success outcome. See Table 20 in Appendix C.



(a) All histories



(b) PPE reward path histories

Figure 4: Number of communication stages

data from matches 6-10.[27] As can be seen in Figure 5(a), the high action is chosen in round 1 more than 60% of the time for the majority of matches of the 2-action games, with the rate dropping just under this mark during the later matches of parameter set A, while holding above this mark under parameter set B. In the 3-action games, the high action is chosen less often in round 1. While this difference persists under parameter set B, it is minimal by the later matches of parameter set A.[28] Under both parameter sets, the rate of medium choice is rarely above 20%, especially after the first match.

The evolution of cooperative initial choices is investigated in more detail using data from the 2-action games. Table 5 reports the results of a correlated random-effects probit regression of the probability of choosing high in round 1, run separately for parameter set A and parameter set B.[29] The first regressions, specifications 1, consider only match variables and choice histories. These are analogous variables to those that have proved important for the evolution of cooperation in prior studies of the (perfect monitoring) prisoner's dilemma.[30] As is the case in these prior studies, the other subject cooperating in the first round of the previous match has a positive effect on the probability of cooperating in the current match. However, the trend over matches is either very slightly negative or not significant at all. The overall impact of the length of the previous match is statistically significant in most cases,[31] while the initial choice has a positive impact on cooperation, but is only marginally significant for parameter set B.

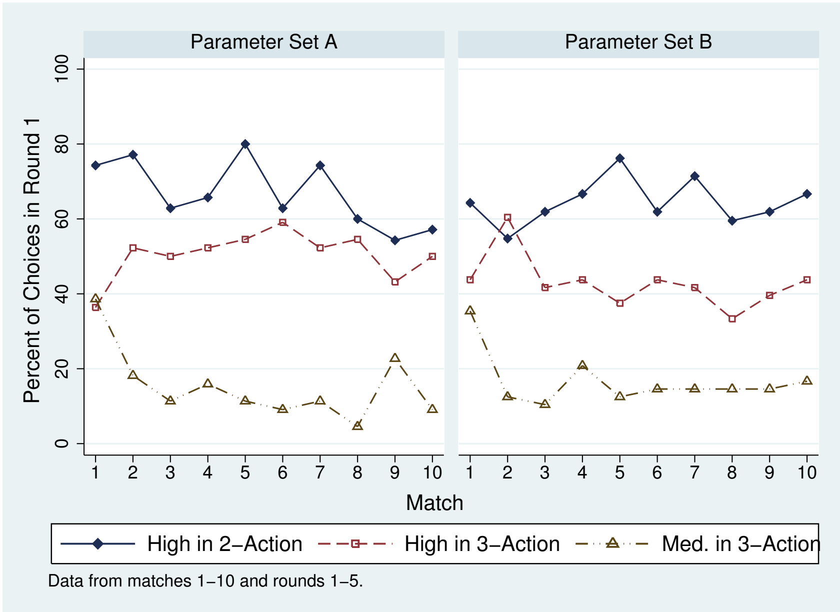
[27]For this analysis, only data from rounds 1-5 is used. Round five is the latest round that can be used while still having data from at least one of the sessions for every match; see Table 17 in Appendix C.

[28]Significance, or lack thereof, of the results illustrated in Figure 5(a), as well as those illustrated in Figure 5(b) that are discussed in the subsequent paragraph, is established using a regression analysis of the probability of choosing the high action. See Table 22 of Appendix C for the results. A similar pattern of behavior is observed for the evolution of expected efficiency, both across matches and across rounds, as reported for the proportion of high effort choices. See Figure 8 and Table 23 of Appendix C.

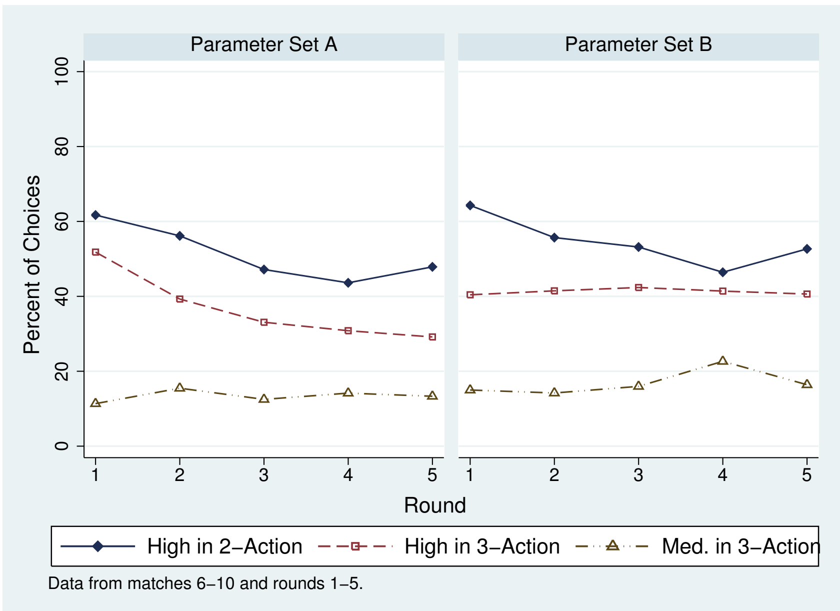
[29]We estimate a correlated random-effects model to deal with the initial conditions problem. We assume that the mean of the subject specific random effects is proportional to the choice in round 1 of match 1. See Heckman (1981) or Chamberlain (1982) for the static case. See Wooldridge (2002) for a clear exposition of the initial conditions problem and methods to address it.

[30]See for example Dal Bó & Fréchette (2011, 2013*a*). The reported analysis is restricted to the 2-action games since this will provide a better comparison with these prior experiments. A similar pattern emerges in the 3-action game; see Table 24 of Appendix C for details.

[31]Although difficult to read from the table, the variables length and $(\text{length})^2$ are jointly significant for all regressions except specification (2) for parameter set B. The p-values of the test are 0.00, 0.00, 0.00 and 0.21, respectively. In most cases, the previous match would need to be at least the expected length for the joint effect to be positive.



(a) Across matches



(b) Within a match

Figure 5: Evolution of high and medium choices

	Parameter set A				Parameter set B			
	(1)		(2)		(1)		(2)	
Match	-0.01^{**}	(0.007)	-0.01	(0.006)	0.01	(0.019)	-0.01^{***}	(0.002)
Length M-1	-0.05	(0.055)	-0.04	(0.041)	-0.03	(0.017)	0.00	(0.016)
(Length M-1)^2	0.00	(0.004)	0.00	(0.003)	0.00^{***}	(0.000)	0.00	(0.002)
Other coop. M-1	0.16^{***}	(0.061)	0.12^{*}	(0.071)	0.10^{***}	(0.014)	0.11^{***}	(0.041)
Coop. M=1	0.34^{**}	(0.159)	0.28	(0.176)	0.20^{*}	(0.118)	0.12	(0.144)
Coop. agree.			0.31^{***}	(0.080)			0.40^{***}	(0.062)
Cheated on M-1			-0.08	(0.070)			-0.19^{***}	(0.073)

M-1 stands for prior match; M=1 stands for first match. Specification 2 includes communication variables as well as the match and outcome variables included in specification 1.
Table reports average marginal effects. Clustered standard errors in parentheses.
***1%, **5%, *10% significance.

Table 5: Correlated random-effects probit regression of the probability of choosing high in round 1 of the 2-action games.

The second regressions, specification 2, add outcomes from the communication technology. Reaching a cooperative agreement at the beginning of a match significantly increases the chance of subsequently choosing high. Observing that your partner from the previous match cheated on an agreement at some point during that match has a significant negative impact on the probability of making a cooperative initial choice in parameter set B, but not in parameter set A.

Behavior within a match, as shown in Figure 5(b), is generally characterized by a gradual reduction in the choice of high effort across rounds; the exception being the 3-action game under parameter set B. Choice of medium effort in the 3-action games, on the other hand, does not vary much over rounds. Such falling rates of high effort is consistent with the provision of dynamic incentives, a key feature of all equilibria that support some form of cooperation. Figure 6 investigates the evolution of the subjects' reaction to the signal directly for the case of the 2-action game.[32] In both parameter sets, the percent of high choices is much lower following a failure than either in round 1 or in rounds that followed a success. Furthermore, this rate declines over matches, especially under parameter set B. By the last match, less than 40% of effort choices are high following a failure, whereas both in the first round, and following a success, rates are around the 60% mark.[33]

[32]See Table 25 of Appendix C for details of the number of observations, and distinct subjects, for each of these signal histories by match.

[33]A similar, if more noisy, pattern is observed in the 3-action games under parameter set A. For parameter set B, while the after-a-failure rate also ends below 40%, the rates of

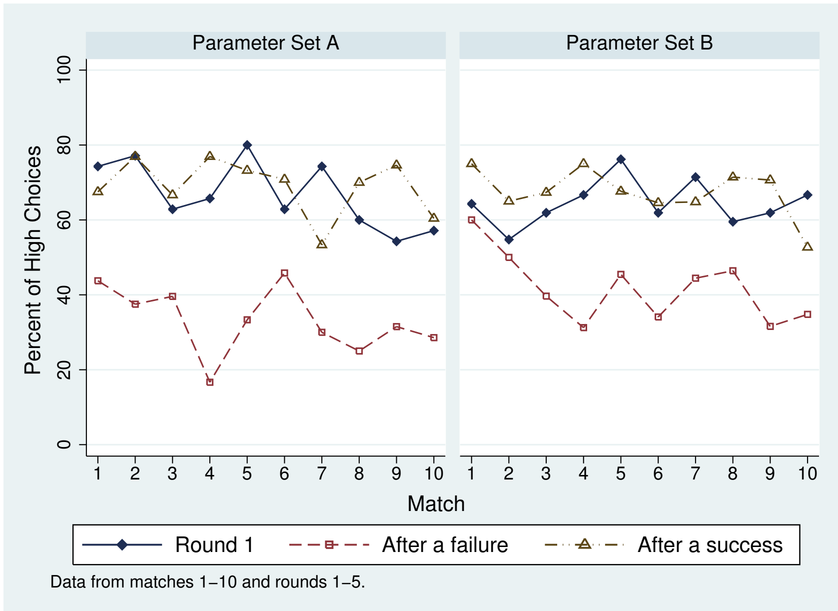


Figure 6: Evolution of high choices, either in round 1 or following a failure or success, in the 2-action games

In the 3-action game, the renegotiation-concerns hypothesis requires that subjects would incorporate the possibility to renegotiate a punishment phase and, as a result, coordinate on an equilibrium with an initial choice of medium effort. However, it is probably more reasonable to expect that subjects might have to experience such renegotiation before it would affect their initial choice. If this were indeed the case, then one would expect the probability of choosing medium in round 1 to be increasing, at least weakly, with the number of prior experiences of renegotiation histories. Figure 7 graphs the probability of choosing the high, medium and Nash effort levels in the first round as a function of the number of prior renegotiation histories in the 3-action games.[34] As is clear from the graph, there is no evidence of such experiences leading to an increase in initial choices of medium. Indeed, the evidence suggests quite the opposite, with the rate of high effort choices increasing, and the rate of both medium and Nash effort choices decreasing.[35]

---

choosing high in the first round or after a success are lower. See Figure 9 of Appendix C.

[34]See Table 26 of Appendix C for details of the number of observations, and distinct subjects, for each of these prior renegotiation-path experiences histories.

[35]A similar pattern for the high and Nash effort choices is observed for the 2-action games. See Figure 10 of Appendix C.

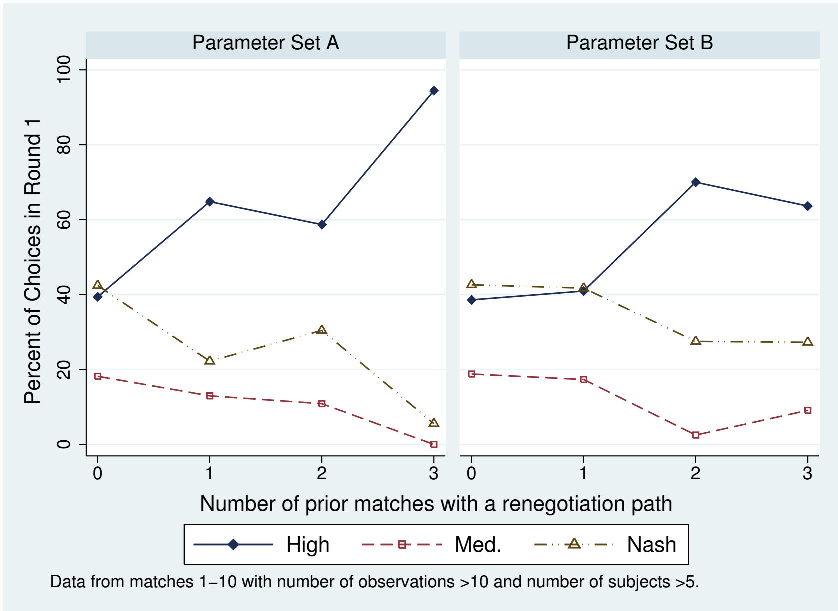


Figure 7: Effect of prior of renegotiation-path experiences on choices in the 3-action games

5.4 Strategy Estimation

The Strategy Frequency Estimation Method (SFEM) is used to investigate the strategies employed by subjects in our experiments.[36] Intuitively, the approach involves estimating how "close" the choices of a subject in a match are to the choices a given strategy would prescribes, then uses a mixture model to evaluate the frequency of each of the strategies considered. A more detailed description of the estimation procedure and the log-likelihood function can be found in Appendix B.1, but the likelihood has the following form:

$$\sum_I \ln\left(\sum_K \phi^k prob_i(s^k(\gamma))\right)$$

where s^k indicates strategy k (from the set of strategies considered K), i indicates a subject (from the set I), and $prob_i(s^k(\gamma))$ gives the probability that the choice i made are generated from strategy k and is a function of parameter γ ,which captures how closely the choices correspond to the strategies. The other parameters, ϕ^k, indicate the fraction of each strategy k.

[36] This method was introduced in Dal Bó & Fréchette (2011), and has also been used in Dreber et al. (2011), Fréchette & Yuksel (2013), Camera et al. (2012), Fudenburg et al. (2012), Dal Bó & Fréchette (2013*a*) and Vespa (2013).

As shown in Dal Bó & Fréchette (2013*a*), including the relevant strategies is important and we use the results of prior experiments to inform our choice. FRD is the closest guide since their implementation includes imperfect public monitoring with a binary, although multi-dimensional, signal. A version of each of the 11 strategies that are present in their environment with statistically significant frequencies is adapted for choices of high and Nash: All High, All Nash, Grim, Mono, Mono21, Mono12, Mono22, Mono31, Grim2, Grim3.[37] Also included are a win-stay-lose-shift (WSLS) strategy, a T11 strategy – which punishes after a single failure but only for 1 period – and a family of strategies we refer to as Sum and SumN, with N equal 2, 3 or 4. The Sum strategy cooperates if there were at least as many success as failures in the match so far and defects otherwise. The SumN strategies are similar but only look back N rounds.[38] Suspicious versions, meaning they start with the Nash effort, of some of the more popular strategies above are also included. These are indicated by adding an S at the beginning, as in Smono11 for instance.

For each of the strategies mentioned above, an analogous strategy that uses medium and Nash actions, instead of high and Nash, is included. A subset of analogous strategies that use high and medium actions are also added. In addition, some strategies that play all three action choices are included, as well as other strategies that do not condition on the signal. In the case of strategies that use all three choices, theory provides little guidance and thus most of these strategies were constructed on the basis that they corresponded to the most commonly observed transitions across choices. However, some were included because they seemed intuitively appealing. There is a trade-off between including too many strategies, making it difficult to distinguish anything, and excluding an important strategy. We first included any strategy that seemed plausible or that was observed in a prior experiment. We estimated this model, which allows for 23 strategies in the treatments with two choices and 78 strategies in the treatments with 3 choices. Any strategy that has a frequency of 1% or below is dropped and the model re-estimated.[39] The complete results can be found in Appendix B.3. A summary of the statistically significant results is given in Table 6, and a further cut of the results, which

[37] "Version" since in our case the signal is unidimensional and thus the strategies are not exactly the same. For instance their tit-for-tat is closest to, but does not match, our monotone strategy, which cooperates if the project was a success in the last round and defects if it was a failure.

[38] We included the sum family of strategies after noticing that, for aggregate level data, there is a monotonic relationship between the running summation of high signals and the probability of choosing the high effort choice. See Figure 11 of Appendix C.

[39] Each strategy used in the second-stage estimation is described in detail in Appendix B.2.

groups strategies according to certain properties, is given in Table 7.[40] The complete model includes 21 strategies in the treatments with two choices and 29 strategies in the treatments with three choices.

The first point to note is that the estimates of β are relatively high and better than a random baseline.[41] In treatments A-2 and B-2 these estimates are slightly lower than in previous papers studying the prisoner's dilemma, whether with perfect or imperfect public monitoring, such as Dal Bó & Fréchette (2011), FRD, and Dal Bó & Fréchette (2013*a*). On the other hand, the βs in the 3-action treatments are much higher than the $\frac{1}{3}$ random-choice benchmark, and in the case of A-3, it is barely below the 2-action estimate.

The frequency with which subjects play All Nash is around 27-29% for all treatments except the A-2-Action treatment, which is much lower at only 6%, where most of choices that initially defect are captured by the Ssum2 HN strategy with 28%. The frequency of All High varies a great deal across treatments, going from being insignificant in the A-2-Action and B-3-Action treatments to a substantial 17% in the B-2-Action treatment. Of the strategies that can support some cooperation – that is excluding All Nash – the vast majority of strategies are conditional. That is, most strategies that sometimes cooperate condition the cooperation decision on outcomes. This is a minimal requirement to find results (beside subjects always playing the one-shot stage Nash) in line with any rational theory in such an environment.

Strategies that support cooperation also vary across treatments. Most notably, both Grim HN and Mono HN are statistically significant for multiple treatments. However, of these two, Mono HN is the most popular strategy overall. On the other hand, the grim family of strategies as a whole totals 21% of the data, averaging across treatments (going from 13% to 30% depending on the treatment). This total is more than the total for the monotone family, which account for 13% of the data (between 5% and 19%). Also very popular are strategies of the Sum family. They average 17% across treatments and are substantially more popular in treatments with two choices than with three choices.

In the 3-action treatments, about 15% of strategies involve the medium action. Only 2 of these are ever individually statistically significant – All M

[40]Grouping strategies by families can be useful as correlation can be high between strategies of the same family and thus it can be challenging to identify as statistically significant specific strategies. See Table 15 in Appendix B.2 for details of which strategies are included in each grouping.

[41]The estimates of β give an indication of the quality of the fit, something difficult to read from γ. Random choice would imply a β of $\frac{1}{2}$ in the 2-action treatments and $\frac{1}{3}$ in the 3-action treatments. See Appendix B.1 for further details.

Strategy	A-2-Action		A-3-Action		B-2-Action		B-3-Action	
	Treatment							
All H	0.056	(0.101)	0.158***	(0.054)	0.171**	(0.083)	0.000	(0.019)
All N	0.059	(0.098)	0.278***	(0.079)	0.285***	(0.088)	0.274***	(0.072)
Grim HN	0.005	(0.050)	0.083**	(0.033)	0.110**	(0.049)	0.000	(0.055)
Mono HN	0.168**	(0.077)	0.145***	(0.049)	0.048	(0.063)	0.098**	(0.044)
T11 HN	0.006	(0.033)	0.000	(0.073)	0.021	(0.042)	0.101**	(0.048)
Grim2 HN	0.005	(0.012)	0.000	(0.059)	0.020	(0.055)	0.077***	(0.028)
Grim3 HN	0.356***	(0.120)	0.000	(0.026)	0.000	(0.005)	0.131***	(0.042)
Sum2 HN	0.005	(0.075)	0.000	(0.031)	0.104*	(0.063)	0.000	(0.007)
Smono HN	0.004	(0.021)	0.017**	(0.008)	0.000	(0.006)	0.000	(0.050)
Ssum2 HN	0.277***	(0.091)	0.000	(0.012)	0.000	(0.033)	0.032	(0.031)
HNNN	0.000	(0.051)	0.026	(0.057)	0.080*	(0.046)	0.022	(0.043)
All M			0.046	(0.049)			0.083**	(0.041)
333 H			0.000	(0.050)			0.059***	(0.023)
γ	0.752	(0.090)	0.563	(0.122)	0.655	(0.088)	0.775	(0.119)
β	0.791		0.747		0.821		0.645	

Bootstrapped standard errors in parenthesis. ***1%, **5%, *10% significance.
Only include strategies that are statistically significant in at least one treatment.
Complete results in Table 16

Table 6: Strategy Frequency Estimation Summary (last five matches).

and 333 H. Going from two to three choices increases forgiveness, as a proportion of cooperative strategies, for both parameter sets. However, for both parameter sets there is no evidence for the use of the T11 MN machine in the 3-action games, which is predicted by the renegotiations-concerns hypothesis. Although, the 333 H strategy is very close to T11 MN, while T11 HN is statistically significant in one treatment, and that is one of the three choices treatments.

6 Discussion

To put the results in perspective, we compare behavior from our 2-action games with previously studied prisoner's dilemma games, both in perfect and imperfect monitoring environments. This discussion also includes the role of

	Treatment			
	A-2-Action	A-3-Action	B-2-Action	B-3-Action
Some Key "Families"				
(S)GrimX	0.370***	0.130	0.129*	0.253***
(S)MonoXY	0.188	0.162***	0.048	0.141*
(S)SumX	0.293***	0.154	0.215*	0.045
1 round punishment	0.006	0.000	0.021	0.160***
Starts with...				
H	0.656	0.570	0.715	0.544
N	0.344	0.345	0.285	0.351
Not lowest effort	0.656	0.609	0.715	0.566
Leniency and forgiveness				
Lenient	0.664	0.154	0.235	0.377
Forgiving	0.495	0.363	0.284	0.412
Lenient / cooperative	0.589	0.253	0.328	0.609
Forgiving / cooperative	0.320	0.523	0.396	0.593
States				
1 or 2	0.331	0.846	0.765	0.623
3 or more	0.669	0.154	0.235	0.377
Conditional				
Yes	0.861	0.466	0.413	0.621
Yes / (1-All N)	0.915	0.645	0.577	0.855
M supported by N	0.000	0.039	0.000	0.080

For key "families", ***1%, **5%, *10% significance indicate the results of an F-test that the sum of the frequencies equal zero.

Cooperative strategies here refer to strategies that do not start with Nash effort.

Table 7: Proportion of estimated strategies with certain properties.

communication. A separate subsection takes up the renegotiation-concerns hypothesis. These latter two elements are not generally considered in the previous studies that we compare.

To facilitate these comparisons, the (ex-ante expected) payoffs in the stage game can be normalised so that the payoff to both players choosing the cooperative action is one, while the payoff to both players choosing the non-cooperative action is zero. This normalization results in the following stage game

	H	N
H	$1, 1$	$-l, 1+g$
N	$1+g, -l$	$0, 0$

where g is the one-shot gain from defecting, compared to the cooperative outcome, and l is the one-shot loss from being defected on, compared to the non-cooperative outcome. If g and l affect behavior, one would intuitively think that an increase in g would generate more defection, since defection becomes more tempting, while an increase in l would result in less cooperation, since it exposes the subject to greater losses. This intuition is confirmed empirically (see Dal Bó & Fréchette 2011). Using the ex-ante expected payoffs of our 2-action games gives a (g, l) combination of $(0.84, 0.36)$ in parameter set A and $(0.82, 0.036)$ in parameter set B.

Prior Prisoner's Dilemma Studies

Some of the factors that have been previously found to affect the evolution of a subject's initial choices are the number of matches the subject has experienced, the length of the matches previously experienced, and whether the subject's partner in the previous match was cooperative (see, for instance, Dal Bó & Fréchette 2011). These factors are not always important in an individual study, but when the data is combined across studies, they are found to matter. Experience, in terms of the number of matches played, can also influence the amount of conditioning – the difference in cooperation rates following a good versus a bad outcome. In the case of this experiment, the one factor that robustly has an effect is the choice of the other subject in the previous match: more cooperative choices increase effort. With the inclusion of communication outcomes, however, we find that reaching a cooperative agreement has the greatest impact on the evolution of cooperative behavior.

	Treatment			
	A-2-Action	A-3-Action	B-2-Action	B-3-Action
Agree on high	74	67	69	57
No agreement	26	25	31	15
Overall average	62	52	64	40

Table 8: Percentage of High in Round 1.

The game parameters, g, l, and δ, also have an impact on cooperation via the size of the basin of attraction of always defecting versus grim (or equivalently TFT with perfect monitoring; see Dal Bó & Fréchette 2011, Rand

& Nowak 2013). In particular, whether grim is a best response to a 50-50 chance that the opponent will always defect or play grim (risk-dominance) seems to be important predictor of cooperation in experiments with perfect monitoring. Under imperfect monitoring the picture is not as clear, and FRD find that risk-dominance does not seem relevant in their setting.

In this study, cooperation is not risk-dominant, as defined above, due to design constraints. Nonetheless, we observe non-negligible levels of full cooperation. One possible explanation is the role of communication in reducing strategic risk. In the one-shot stag-hunt game, Cooper et al. (1992) show that structured two-way communication – where, as in our experiment, subjects could indicate their intended choice or stay silent – increases the frequency of choices of the payoff-dominant equilibrium, as opposed to the risk-dominant equilibrium, compared to a baseline game with no cooperation. Since our focus is not on communication per se, our design is not meant to isolate the effect of communication. However, as can be seen in Table 8, our data suggest that subjects do rely on communication to help with coordination. The first row gives the probability of a choice of high if both subjects agreed on playing high in round 1. In all treatments, this frequency is higher than the overall probability of high, which is given in the third row – cross treatments, the difference is 12.23 percentage points. Even more striking, is how much lower the probability of selecting high is if there is no agreement.[42] This evidence is, to our knowledge, the first of communication being used to reduce strategic uncertainty in infinitely repeated games.

A clear result to come out of FRD is that when monitoring is imperfect (and public), strategies are more lenient and more forgiving than under perfect monitoring. Dal Bó & Fréchette (2013*b*) report that in six studies with perfect monitoring, six strategies (Always Cooperate, Always Defect, Grim, TFT, WSLS, and STFT) account for over 50% of strategy choices, and that three of those (Always Defect, Grim, and TFT) account for the majority of strategies in five of those six studies. On the other hand, in the five treatments of FRD with imperfect monitoring, only in one of those treatments is that the case. Similarly, in this study, the strategies All N, Grim HN, and Mono HN represents a majority of the strategies in only one of the four treatments; in two of the four if we consider the six strategies All H, All N, Grim HN, Mono HN, WSLS HN, and Smono HN. Much of the strategy choices are of the lenient Grim types, the forgiving Mono type, or the lenient and forgiving Sum kind. In that sense, these results echo the observation of FRD.

[42]Note that this does not mean there were no messages. In fact, in the last 5 matches, which is the data considered in this table, there are no matches where there was not at least one message sent in round 1.

The Sum strategies are new to the experimental literature on prisoner's dilemma-esque environments. Given the coarseness of the signal, they are a natural heuristic with which to implement a lenient and forgiving strategy in the repeated-partnership game. Indeed, our results suggest that subjects are much more inclined to use these than lenient derivatives of the monotone family. Although not common in prior experimental studies, the Sum strategy resembles the strategies predicted in the literature on favor exchanges (see Möbius 2001, Skrzypacz & Hopenhayn 2004, Hopenhayn & Hauser 2008).[43]

In summary, the results from imperfect public monitoring experiments as a whole suggest that the observation of FRD that under imperfect monitoring strategies tend to be more lenient and more forgiving is a general phenomenon not confined to the particular game and monitoring structure they considered. However, the type of strategies used and whether both leniency and forgiveness increase does depend on the game and monitoring structure. In environments with discrete signals, a certain measure of leniency is observed via strategies that do not immediately trigger punishments. When the public signal is continuous (as in AF), leniency is instead observed through the use of lower than optimal thresholds to trigger punishments. Finally, unlike in the environment of FRD, the majority of strategies observed use at most two states,[44] but just as in their case, there is no indication that a significant number of subjects play win-stay-lose-shift strategies, a useful and popular strategy in the theory of repeated games (see, for example, Imhof et al. 2007).

[43]In particular, it is possible to re-interpret the Sum family of strategies as a form of chips mechanism (Skrzypacz & Hopenhayn 2004), adapted for the repeated partnership game: Once cooperation has been established, that is the running sum is strictly positive, choosing the high action after a failure signal can be viewed as a 'favor'. In the basic Sum strategy, the more consecutive successes, the greater the entitlement of the other player for favors. In the SumN versions, this entitlement is capped at N, and a degree of forgivingness is introduced since a long sequence of failures can be made up with a smaller number of consecutive successes – a feature of the optimal public perfect equilibrium of Hopenhayn & Hauser (2008). It should be noted that there are a number of important differences between the repeated partnership game and the cited favor exchange environments: First, while it is clear to both players when there is the opportunity to grant a favor, it is not verifiable that the opportunity was taken and the favor granted; second, since the opportunities to grant favors are based on the symmetric history of the public signal, when there is such an opportunity it is available to both players, rather than just one at a time. For an experimental implementation looking to investigate directly the favor exchange environment, see Roy (2012).

[44]We find at least 60% of strategies in three of our four treatments use at most two states. AF cannot reject the restriction that the best-fitting strategy has only two states for any treatment where cooperation can be supported in equilibrium. Dal Bó & Fréchette (2013*a*), who elicit strategies in a perfect monitoring environment, find 85% of strategies have at most two states.

Renegotiation Concerns

The imperfect monitoring environment highlights a potential inconsistency in cooperative equilibria of repeated games: should subjects reach a punishment phase – which is both necessary and happens with positive probability in equilibrium – why would they not renegotiate away from it and start afresh? In the 2-action games, such efforts would be redundant, since there is no credible alternative to switch to other than defecting forever, which is exactly the path being renegotiated. In the 3-action games, the introduction of the medium choice provides a tension between supporting the efficient level of effort or a lower level, which can be supported by a more credible punishment strategy.

This renegotiation-concerns hypothesis appears to correctly predict the change in expected payoff efficiency, as well as a reduction in the use of high effort, when moving from the 2-action to the 3-action game. However, it is clear that, with no more than 15% of round 1 choices of medium, renegotiation does not identify the choices made by subjects. Furthermore, there is no evidence in the 3-action games for the strategy predicted by the hypothesis.

The equilibrium selection implicit in the hypothesis is, however, very demanding: it requires that subjects, just through a process of introspection, incorporate the possibility to renegotiate and thus adjust the super-game equilibrium they coordinate on. It seems more reasonable to expect that subjects might need to experience such renegotiation first, and that coordination on a renegotiation-proof equilibrium might be more difficult. In this spirit, we identify three channels through which such renegotiation experiences might affect behavior in less transparent ways.

First, we examine whether histories associated with such renegotiation experiences result in longer phases of communication. Second, we look at whether experiencing such renegotiation histories in a prior match results in the greater use of the medium, or even Nash, action. While there is some support for the first conjecture during the first five matches, by the end of the experiment both the length and content of agreements during renegotiation phases do not appear to differ from the pattern observed for all histories. On the second conjecture, there is no evidence that experiencing a renegotiation phase resulted in subjects learning to support the medium action instead, or even giving up on supporting cooperation altogether.

The third channel looks for changes in the type of strategies used to support cooperation when the medium action is included. The logic of the renegotiation prediction can be thought of in two parts: First, to avoid renegotiation, the players must increase the value in the punishment state. Second, since the first part results in a smaller reduction in value when going from the reward state to the punishment state, the value that can be supported in the reward state might need to be decreased. The first of these two points is achieved by the use of more forgiving strategies – by eventually moving away from the Nash outcome, the value in the punishment state rises. As can be seen in Table 7, amongst strategies that intend to cooperate from the start, i.e., do not start with Nash effort, the fraction of forgiving strategies increases when the third action is added; in both cases increasing by about 20 percentage points. In addition, in treatment B, going from 2 to 3 choices, the strategy T11 HN, which punishes for only one round, just like the renegotiation proof strategy, is played in 10% of the data, and is statistically significant. Finally, in that same treatment, B-3, the strategy 333 H is found to be played in 6% of the data, a small, but significant, proportion. That strategy is very close to the renegotiation proof strategy, the only difference being the first choice.

Hence, among subjects who have the intention of cooperating from the start of a match, there is a movement in the direction of using less severe punishments. However, few of them go all the way to using the least severe punishment, although some do. One possibility is that they do not go all the way to the minimal punishment because they trade off this concern with a desire to support higher effort in the reward phase. In addition, the renegotiation-proof strategy has a rather small basin of attraction against All Nash, smaller than Grim for instance.[45] As previously noted, the size of the basin of attraction of always defect has been found to be a good predictor of cooperation rates in perfect monitoring environments. It may, more generally, be a helpful predictor of strategy choice, although more research would be required to determine if this is indeed the case.

This shift in the direction of less severe punishments cannot mask the fact that the bulk of the evidence is that subjects in these experiments do not behave as predicted by renegotiation-proofness. In a different, but potentially related vein, Dal Bó & Fréchette (2013*a*) observe in a perfect monitoring environment that, although the vast majority of strategy choices correspond to Nash equilibrium strategies, a large fraction of them are not sub-game perfect. This may indicate that certain theoretically relevant considerations when it comes to what strategies should be selected are either beyond, or beside, the concerns of subjects. In particular, considerations having to do with decisions nodes further in the decision tree may be too remote to be decisive. Again, this is, at this point, a matter of conjecture, but could formulate an interesting hypothesis for future work.

[45]The basin of attraction of Grim HN is 0.267 and 0.167 in treatments A-3 and B-3 respectively; whereas it is 0.264 and 0.039 for T11 MN.

7 Conclusion

This paper presents the results of a laboratory implementation of a repeated partnership game with a simple imperfect public monitoring environment. The experiments aim to understand how subjects support and maintain cooperation in a setting whose key strategic features underlie many economic interactions. The stage game has a prisoner's dilemma structure that provides the tension between individual incentives and optimality for the group. The imperfect monitoring environment, while of interest in itself, results in naturally occurring punishment phases that provide a rich source of data to study the strategies subjects use. The monitoring structure highlights potential renegotiation concerns, the resolution of which gives an equilibrium refinement with sharp predictions for super-game behavior. Our implementation has the flexibility to test this hypothesis by changing the number of actions available to subjects in the stage game.

Despite the complexity of the environment, our results suggest that a significant number of subjects attempt to support cooperation, and do so by conditioning their behavior on the outcome of the signal. Cooperation rates observed in our imperfect monitoring implementation compare favorably to those from prior experimental studies in which subject observe the action choices of their partners. Efforts to cooperate appear to be aided by the communication stage that is repeated before each round. Most pairs make symmetric, short-term agreements that coordinate on the most efficient outcome using a minimal number of message iterations. Even though they carry no commitment, these messages are informative. Agreements made in round 1 or after a success are generally honored, whereas stages that end without an agreement mostly result in noncooperative choices.

The strategy analysis reveals a degree of heterogeneity in approaches to the repeated game. Although more use conditional strategies, some 20-30% of subjects play the repeated one-shot Nash equilibrium. Of those exhibiting conditional behavior, we find both forgiving monotone punishments, unforgiving Grim punishments, and a counting strategy, which we refer to as Sum, are common. In addition, subjects display more lenient and forgiving behavior than is typical in perfect monitoring environments. However, when using Grim or tit-for-tat-like strategies , our subjects either adopt a lenient stance or a forgiving one, but not both in any significant numbers. If they want to be both lenient and forgiving, they do so using the Sum-type strategies.

On the equilibrium refinement front, despite the initial promise at an aggregate level, our results do not suggest renegotiation-concerns as the driver for inefficient super-game behavior. Instead, we find more support for the idea that, in a setting with heterogeneous behavior, subjects' overwhelming concern is to establish cooperation, rather than to maintain it, even with a communication technology to facilitate coordination. For example, after experiencing a punishment phase, we find subjects are more likely to start again with a high level of cooperation than to switch to either the medium or Nash level. An interest contrast, however, is that, when renegotiation-proofness predicts a forgiving strategy, we do indeed observe that subjects who use cooperative strategies are more likely to use forgiving strategies, albeit with longer punishment phases.

References

Abreu, D., Pearce, D. & Stacchetti, E. (1986), 'Optimal cartel equilibria with imperfect monitoring', *Journal of Economic Theory* **39**(1), 251–269.

Abreu, D., Pearce, D. & Stacchetti, E. (1993), 'Renegotiation and symmetry in repeated games', *Journal of Economic Theory* **60**(2), 217–240.

Andersson, O. & Wengström, E. (2012), 'Credible communication and cooperation: Experimental evidence from multi-stage games', *Journal of Economic Behavior & Organization* **81**(1), 207–219.

Aoyagi, M. & Fréchette, G. R. (2009), 'Collusion as public monitoring becomes noisy: Experimental evidence', *Journal of Economic Theory* **144**(3), 1135–1165.

Benoit, J.-P. & Krishna, V. (1985), 'Finitely repeated games', *Econometrica: Journal of the Econometric Society* pp. 905–922.

Bernheim, D. B. & Ray, D. (1989), 'Collective dynamic consistency in repeated games', *Games and Economic Behavior* **1**(4), 295–326.

Bigoni, M., Potters, J. & Spagnolo, G. (2011), Flexibility and cooperation with imperfect monitoring.

Camera, G., Casari, M. & Bigoni, M. (2012), 'Cooperative strategies in anonymous economies: An experiment', *Games and Economic Behavior* **75**(2), 570 – 586.

Cason, T. N. & Khan, F. U. (1999), 'A laboratory study of voluntary public goods provision with imperfect monitoring and communication', *Journal of Development Economics* **58**(2), 533–552.

Chamberlain, G. (1982), 'Multivariate regression models for panel data', *Journal of Econometrics* **18**(1), 5–46.

Cooper, D. & Kuhn, K.-U. (2009), 'Communication, renegotiation, and the scope for collusion'.
URL: *http://ssrn.com/abstract=1533168*

Cooper, R., DeJong, D. V., Forsythe, R. & Ross, T. W. (1992), 'Communication in coordination games', *The Quarterly Journal of Economics* **107**(2), 739–771.

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. & Fréchette, G. R. (2011), 'The evolution of cooperation in infinitely repeated games: Experimental evidence', *American Economic Review* .

Dal Bó, P. & Fréchette, G. R. (2013*a*), Strategy choice in the infinitely repeated prisoners' dilemma.

Dal Bó, P. & Fréchette, G. R. (2013*b*), Survey on infinitely repeated games in the laboratory. Manuscript in preparation.

Davis, D. D. & Holt, C. A. (1999), 'Equilibrium cooperation in two-stage games: Experimental evidence', *International Journal of Game Theory* **28**(1), 89–109.

Dreber, A., Fudenberg, D. & Rand, D. G. (2011), 'Who cooperates in repeated games?', *Available at SSRN* .
URL: *http://ssrn.com/abstract=1752366*

Duffy, J. & Ochs, J. (2009), 'Cooperative behavior and the frequency of social interaction', *Games and Economic Behavior* **66**(2), 785–812.

Farrell, J. & Maskin, E. (1989), 'Renegotiation in repeated games', *Games and Economic Behavior* **1**(4), 327–360.

Feinberg, R. M. & Husted, T. A. (1993), 'An experimental test of Discount-Rate effects on collusive behaviour in duopoly markets', *The Journal of Industrial Economics* **41**(2), 153–160.

Feinberg, R. & Snyder, C. (2002), 'Collusion with secret price cuts: an experimental investigation', *Economics Bulletin* **3**(6), 111.

Fischbacher, U. (2007), 'z-tree: Zurich toolbox for ready-made economic experiments.', *Experimental Economics* **10(2)**, 171–178.

Fréchette, G. R. (2012), 'Session-effects in the laboratory', *Experimental Economics* **15**(3), 485–498.

Fréchette, G. R. & Yuksel, S. (2013), Infinitely repeated games in the laboratory: Four perspectives on discounting and random termination.

Fudenberg, D. & Maskin, E. (1986), 'The folk theorem in repeated games with discounting or with incomplete information', *Econometrica* **54**(3), 533–554.

Fudenburg, D., Rand, D. G. & Dreber, A. (2012), Slow to anger and fast to forgive: Cooperation in an uncertain world.

Heckman, J. J. (1981), Heterogeneity and state dependence, *in* 'Studies in labor markets', University of Chicago Press, pp. 91–140.

Holcomb, J. H. & Nelson, P. S. (1997), 'The role of monitoring in duopoly market outcomes', *Journal of Socio-Economics* **26**(1), 79–93.

Holt, C. A. (1985), 'An experimental test of the Consistent-Conjectures hypothesis', *The American Economic Review* **75**(3), 314–325.

Hopenhayn, H. A. & Hauser, C. (2008), Trading favors: optimal exchange and forgiveness, Carlo Alberto Notebooks 88, Collegio Carlo Alberto.
URL: *http://www.carloalberto.org/assets/working-papers/no.88.pdf*

Imhof, L. A., Fudenberg, D. & Nowak, M. A. (2007), 'Tit-for-tat or win-stay, lose-shift?', *Journal of theoretical biology* **247**(3), 574–580.

Mailath, G. J. & Samuelson, L. (2006), *Repeated games and reputations: long-run relationships*, Oxford University Press, USA.

Möbius, M. (2001), Trading favors.

Murnighan, J. K. & Roth, A. E. (1978), 'Large group bargaining in a characteristic function game', *The Journal of Conflict Resolution* **22**(2), 299–317.

Murnighan, J. K. & Roth, A. E. (1983), 'Expecting continued play in prisoner's dilemma games: A test of several models', *The Journal of Conflict Resolution* **27**(2), 279–300.

Palfrey, T. R. & Rosenthal, H. (1994), 'Repeated play, cooperation and coordination: An experimental study', *The Review of Economic Studies* **61**(3), 545–565.

Pearce, D. G. (1987), 'Renegotiation-Proof equilibria: Collective rationality and intertemporal cooperation'.
URL: *http://cowles.econ.yale.edu/P/cd/d08b/d0855.pdf*

Rand, D. G., Fudenberg, D. & Dreber, A. (2013), It's the thought that counts: The role of intentions in reciprocal altruism.
URL: *http://papers.ssrn.com/abstract=2259407*

Rand, D. G. & Nowak, M. A. (2013), 'Human cooperation', *Trends in cognitive sciences* .

Roth, A. E. & Murnighan, J. K. (1978), 'Equilibrium behavior and repeated play of the prisoner's dilemma', *Journal of Mathematical Psychology* **17**(2), 189–198.

Roy, N. (2012), Cooperation without immediate reciprocity: An experiment in favor exchange.

Skrzypacz, A. & Hopenhayn, H. (2004), 'Tacit collusion in repeated auctions', *Journal of Economic Theory* **114**(1), 153–169.

Vespa, E. (2013), Cooperation in dynamic games: An experimental investigation.

Wooldridge, J. M. (2002), *Econometric analysis of cross section and panel data*, The MIT press.