Negotiating Cooperation under Uncertainty: Communication in Noisy, Indefinitely Repeated Interactions*

Fabian Dvorak
University of Konstanz,
Center for the Advanced Study
of Collective Behavior
fabian.dvorak@uni-konstanz.de

Sebastian Fehrler
University of Bremen,
SOCIUM – Research Center on
Inequality and Social Policy
sebastian.fehrler@uni-bremen.de

December 2020

Abstract

Case studies of cartels and recent theory suggest that communication is a key factor for cooperation under imperfect monitoring, where actions can only be observed with noise. We present results from a laboratory experiment which show that communication facilitates cooperation by reducing two types of uncertainty. Pre-play communication reduces strategic uncertainty, which boosts cooperation at the beginning of an interaction. Under perfect monitoring, this is sufficient to reach a high and stable cooperation rate. However, repeated communication is necessary to maintain a high level of cooperation under imperfect monitoring, where players face additional uncertainty about the history of play.

Keywords: infinitely repeated games, monitoring, communication, cooperation, strategic uncertainty, prisoner's dilemma

JEL Classification: C72, C73, C92, D83

*We would like to thank Pietro Battiston, Maria Bigoni, Sam Bowles, Yves Breitmoser, Marco Casari, Anna Dreber Almenberg, Matthew Embrey, Simon Gächter, Susanne Goldlücke, Yuval Heller, Melis Kartal, David Miller, Wieland Müller, Ariel Rubinstein, Martin Sefton, Susumu Shikano, Bauke Visser, Georg Weizsäcker, Nick Zubanov, participants at the "Repeated Games: Theory and Experiments" workshop in Konstanz and several conferences and seminars, and the research group at the Thurgau Institute of Economics, Urs Fischbacher, Moritz Janas, and Irenaeus Wolff in particular, for helpful comments. We are also very grateful to Guillaume Fréchette for sharing his programming insights into the estimation of finite mixture models, and to Yongping Bao, Felix Beck and Malin Pimper for excellent research assistance. All remaining errors are our own.

1 Introduction

Many social and economic relationships are characterized by repeated interactions in which the behavior of partners is observable only with noise. Two examples are teamwork arrangements in which workers repeatedly produce goods for each other, and cartels with repeated price-setting by its members. How much effort a worker exerts in the production of the good cannot be observed directly but only inferred from the good itself – a noisy signal (Sekiguchi, 1997; Compte and Postlewaite, 2015). Likewise, whether or not other cartel members stick to a collusive agreement cannot be observed directly but only inferred from noisy signals like own sales in Stigler's (1964) or the market price in Green and Porter's (1984) seminal treatments of oligopolies. The former is the classic example for imperfect private monitoring – own sales can be observed only by the firm itself; the latter is the classic example for imperfect public monitoring, the market price being publicly observable.

Sustaining cooperation under imperfect monitoring has been the central topic in the theory of infinitely repeated games for the last three decades. This literature identifies communication as a key factor. However, empirical evidence on how communication promotes cooperation under imperfect monitoring is so far missing in the literature. This is unfortunate not only from a scientific but also from a policy perspective as evidence of cartel communication is an important source of evidence for antitrust authorities. Hence, it is important to know if communication is necessary and, if so, how much of it is needed to sustain a collusive agreement. Several experimental studies with perfect monitoring have demonstrated that the mere existence of cooperative equilibria is by no means sufficient for the emergence of cooperation (see, e.g., Dal Bó, 2005; Blonski et al., 2011; Breitmoser, 2015; Dal Bó and Fréchette, 2018a). Therefore, it is important to learn from data and not only from theory how communication affects cooperation and how this depends on the monitoring structure. Hence, we design a laboratory experiment, which allows us to study this interaction in a tightly controlled setting. Our results suggest that communication promotes cooperation by reducing two types of uncertainty. First, communication before the first round of the game reduces *strategic uncertainty*, that is uncertainty about the strategy their partner will follow in the game. Second, communication between rounds reduces *uncertainty about the history of play*, which stems from the noisy signals. As a result, participants' play becomes more lenient after bad signals, which facilitates relationships with high and stable cooperation rates.

Our laboratory experiment follows a (3×3) design varying both the communication and the monitoring structure of the game. We study the following communication structures: (i) no communication; (ii) pre-play communication, where subjects can chat with their partner before the first round of an interaction; and (iii) repeated communication, where subjects can chat with their partner before every round of the interaction. The second treatment variable is the monitoring structure. We study (i) perfect monitoring, (ii) imperfect public montoring, and (iii) imperfect private monitoring. The game that we ask subjects to play is an indefinitely repeated noisy prisoner's dilemma, similar to that studied theoretically by Sekiguchi (1997) and Compte and Postlewaite (2015), and experimentally, but without communication, by Aoyagi et al. (2018). In this variant of the prisoner's dilemma, signals are independent conditional on actions. Payoffs depend on own actions and received signals, which are noisy reflections of the other player's actions. Under perfect monitoring, signals and actions are observed; under imperfect public monitoring, sent and received signals are observed; and under private monitoring, only the received signals are observed.[1]

In their comprehensive review of experimental studies of repeated games without communication, Dal Bó and Fréchette (2018a) show that the level of strategic uncertainty is a good predictor for cooperation under perfect monitoring (see also Blonski et al., 2011; Breitmoser, 2015). We extend their measure of strategic uncertainty – the basin of attraction for playing a defective strategy – to the imperfect monitoring cases. According to this measure, strategic uncertainty is high in all of our treatments and higher in the treatments with imperfect monitoring than in those with perfect monitoring. Pre-play communication might reduce this uncertainty and thus increase cooperation rates. Under imperfect monitoring, full cooperation is not possible in equilibrium and simple grim-trigger strategies lead to lower efficiency than more lenient and forgiving strategies. Bad signals can also occur after a history of full cooperation, which adds another type of uncertainty, as the history of play becomes becomes uncertain. If communication makes subjects play more lenient and forgiving, this could boost cooperation rates. As we will see, the opportunity to communicate after the occurrence of a bad signal is crucial in this respect.

We choose open chat as the mode of communication to avoid reducing the potential roles that communication might play. Free-form communication is also the most natural type and allows us to study its use and content. We set game parameters which guarantee the existence of cooperative and evolutionarily stable

[1] We use the same stage game payoffs, the same continuation probability of $\delta = 0.8$ and the same error probability of $\epsilon = 0.1$ in all treatments.

equilibria under private monitoring with repeated communication, in which players truthfully reveal their private signals (Heller, 2017). As this is the main role ascribed to communication in the theoretical literature (e.g., Matsushima, 1991; Ben-Porath and Kahneman, 1996; Compte, 1998; Kandori and Matsushima, 1998; Obara, 2009; Awaya and Krishna, 2016), we conduct a detailed analysis of the communication content with a special focus on the exchange of private information. Moreover, our parameterization allows us to predict a unique response pattern of memory-one belief-free (M1BF) equilibrium strategies, an important class of strategies under private monitoring (e.g., Ely and Välimäki, 2002; Ely et al., 2005; Piccione, 2002). To be able to estimate the prevalence of these strategies in the data, we adapt and extend the widely used strategy frequency estimation method (SFEM) by Dal Bó and Fréchette (2011). Our approach, which is based on the EM-algorithm (Dempster et al., 1977), allows us to estimate finite mixtures of pure and behavior strategies (for a similar approach see Breitmoser, 2015; Backhaus and Breitmoser, 2018).

Our main results are the following: Cooperation rates are much higher in all communication treatments than in the no-communication treatments (Result 1). With repeated communication, cooperation rates are high and stable under all monitoring conditions, whereas they start high but decline rapidly over rounds with pre-play communication when monitoring is imperfect (Result 2). As bad signals occur with positive probability even after mutual cooperation under both imperfect monitoring structures, subjects need to coordinate their behavior for more contingencies than under perfect monitoring. While subjects do use pre-play communication to coordinate behaviour (Result 3a), we also find for all three monitoring structures that most subjects merely coordinate on mutual cooperation in the first round. If successful, they continue to cooperate. While this is enough to lead to stable coordination under perfect monitoring, it is not when signals are noisy. To reduce uncertainty about the history of the game under imperfect monitoring, subjects use repeated communication phases to exchange private information (Result 3b), and to revisit their incomplete pre-play agreements. Our strategy estimations show that lenient and forgiving strategies are chosen much more frequently with pre-play communication, and even more frequently with repeated communication, than in the absence of communication opportunities (Result 4). These results corroborate the importance of communication for high cooperation rates and of repeated communication for the stability of cooperation in noisy environments.

In the theoretical literature, communication has played a particularly prominent role in combination with private monitoring. For this monitoring structure, it has been shown that repeated communication opportunities can enlarge the set of achievable equilibria (Matsushima, 1991; Ben-Porath and Kahneman, 1996; Compte, 1998; Kandori and Matsushima, 1998; Obara, 2009; Awaya and Krishna, 2016). Moreover, *truthful communication equilibria*, in which players reveal their private signals, are evolutionarily stable, while the cooperative equilibria without repeated communication that have been analyzed in the literature are not (Heller, 2017). Moreover, any cooperative equilibrium without repeated communication requires complicated mixing, which makes coordination with or without pre-play communication extremely difficult; this led Compte and Postlewaite (2015) to characterize them as "*unrealistically complex and fragile*" (p. 45). These considerations suggest low cooperation rates in the private monitoring treatments without communication; this is, indeed, what we observe. They also suggests low cooperation rates with pre-play communication; instead, we observe high cooperation rates at the beginning of the interactions. This suggests that pre-play communication is effective in reducing strategic uncertainty through improved coordination on cooperation. However, subjects' mere coordination on mutual cooperation in the beginning of the interaction is insufficient to maintain a high and stable cooperation level both under imperfect private monitoring and under imperfect public monitoring, where efficient equilibria require the use of complex strategies as well. In our experiment, subjects fail to identify and communicate these strategies, which further increases the uncertainty about the history of play when bad signals occur for the first time.

Our key finding that repeated communication is important for stable cooperation in noisy environments is consistent with evidence from a number of case studies of cartels, which point to different roles for repeated communication. Genesove and Mullin (2001) note in their account of the sugar-refining cartel that its weekly *"[m]eetings were used to interpret and adapt the agreement, coordinate on jointly profitable actions, and determine whether cheating had occurred"* (p. 379); put differently: meetings were used to reduce strategic uncertainty and uncertainty about the history of play. Levenstein and Suslow (2006) review the empirical literature on cartels and identify repeated communication as a key ingredient of successful cartel organizations – *"not only to provide flexibility in the details of the agreement, but to build trust as well"* (p. 67). Finally, Harrington and Skrzypacz (2011), who study various cartel agreements, conclude that repeated truthful communication of sales is an important property of all of them. However, relying on field data has its limitations for the study of cooperation under imperfect monitoring since the noisy signals are often not observable for researchers. Likewise, most cartel communication is not documented as such documents could be used as evidence in legal cases.[2] To overcome these limitations, several laboratory experiments have been conducted to explore the effects of communication and test predictions from renegotiation-proofness refinements (Pearce, 1987; Farrell and Maskin, 1989) in experiments without noise (Andersson and Wengström, 2012; Fonseca and Normann, 2012; Cooper and Kühn, 2014a) or imperfect public monitoring (Embrey et al., 2013; Arechar et al., 2017). While they offer important insights, which we discuss further in Section 2.3, these experiments do not allow for a comparison of the use and the effects of communication across monitoring structures.[3] Moreover, we are not aware of any previous empirical study of communication under private monitoring. Given the important role that communication plays for this monitoring structure in the theoretical literature, this is quite surprising. Our study makes a first step to fill these gaps, and provides novel insights into how communication reduces uncertainty under all three monitoring structures.

The rest of the paper is structured as follows. In the next section, we present the game and its theoretical properties, and extend the theoretical predictors of cooperation to the imperfect monitoring cases. In Section 3, we present the experimental design, state our research questions, and discuss the methods used to tackle them. We turn to the experimental results in Section 4. In Section 5, we discuss our key findings and our methodological contributions. We conclude in Section 6.

2 The Repeated Prisoner's Dilemma with Noise

Two players interact with each other in indefinitely many rounds of an interaction, henceforth called a supergame. Let δ denote the fixed continuation probability after any given round. In every round, each of the two players can choose between two actions C or D. After both players have chosen an action $a \in \{C, D\}$, a noisy process translates both actions into conditionally independent signals. Each signal $\omega \in \{c, d\}$ corresponds to the chosen action with probability $(1-\epsilon)$. With probability ϵ, an error occurs and the action is translated into the wrong signal (C to d and D to c). All aspects of this process, the conditional independence of signals as well as

[2] See Genesove and Mullin (2001), Andersson and Wengström (2007), and Cooper and Kühn (2014a) for further discussion and examples of cartel cases.

[3] Camera et al. (2013) study communication in a setting with random re-matching within groups after every round of the repeated game. Kamei and Nesterov (2020) study endogenous reports of opponents' choices after a supergame to their next interactions partners. Vespa and Wilson (2018) study an indefinitely repeated version of a sender-receiver game (Crawford and Sobel, 1982). The experimental literature on indefinitely repeated games with imperfect monitoring but without communication further includes the papers by Aoyagi and Fréchette (2009) and Fudenberg et al. (2012) who study public monitoring.

the probability of an error are common knowledge. The payoff π_i of player i from the current round is defined by player i's own action a_i and the signal of the other player's action ω_{-i}.[4] We consider the following normalized expected stage-game payoffs of action profiles which take the noise into account:

	C	D
C	1,1	$-l$,1+g
D	1+g,$-l$	0,0

With $g > 0$ and $l > 0$ the stage game has the form of a prisoner's dilemma. The restriction $1 + l > g$ prevents that coordinated switching between cooperation and defection yields a greater expected payoff than mutual cooperation. We consider three different monitoring structures. Under perfect monitoring, each player i is informed about the actions $\{a_i, a_{-i}\}$ and the signals $\{\omega_i, \omega_{-i}\}$. Under imperfect *public* monitoring (Green and Porter, 1984), players cannot observe the action of the other player and the information set reduces to $\{a_i, \omega_i, \omega_{-i}\}$. Under imperfect *private* monitoring (Stigler, 1964), players also remain uninformed about ω_i, the signal the other player receives, as the information set reduces to $\{a_i, \omega_{-i}\}$.

2.1 Strategic Uncertainty

The conditions for cooperative subgame-perfect equilibria (SPE) under perfect and imperfect public monitoring are well-known results of the theoretical literature (see, e.g., Mailath and Samuelson, 2006). With perfect monitoring, players can condition on the intended actions and support full cooperation using pure strategies if the continuation probability δ is greater or equal to $\delta^{SPE} = \frac{g}{1+g}$. With public monitoring and strategies conditioning only on the public signals, the stricter condition $\delta^{SPE} = \frac{g}{1-\epsilon+(1-\epsilon)^2 g}$ applies with reduced efficiency.[5] With private monitoring, cooperation

[4]This might reflect the interaction of two workers where each worker exerts low or high effort on the production of a good for the other worker, and where whether the good is useful for the partner or not is a noisy signal of effort (Sekiguchi, 1997). For an alternative but similar interpretation, see Compte and Postlewaite (2015).

[5]The continuation probability must be high enough to defer a deviation in the cooperative state of the strategy with the strictest punishment scheme, the grim-trigger strategy. With perfect monitoring, grim-trigger defects for the remaining rounds of the interaction when at least one player defected in the previous period. With imperfect public monitoring, grim-trigger defects for the remaining rounds of the interaction when at least one signal indicates defection. The long-run incentives of cooperation must be as least as big as the immediate gains from defection in the

cannot be supported by an SPE based on pure strategies and players have to rely on mixed (Bhaskar and Obara, 2002; Sekiguchi, 1997) or behavior strategies (Ely and Välimäki, 2002; Piccione, 2002).[6]

Experimental evidence suggests that the SPE conditions are necessary but insufficient to observe high levels of cooperation in the indefinitely repeated prisoner's dilemma (see Dal Bó and Fréchette, 2018a). An obstacle for the emergence of cooperation is that mutual defection remains an equilibrium of the repeated game when cooperative equilibria exist. Without the possibility to coordinate strategies, the uncertainty about the strategy choice of the other player makes cooperation risky. Two different predictors for the emergence of cooperation under SPE conditions have been proposed that both take into account how risky cooperation is under strategic uncertainty. Dal Bó and Fréchette (2011) propose the basin of attraction of defection (BAD) as a predictor for cooperation. In a mixed population of grim-trigger (GRIM) and always-defect (ALLD) players, the BAD is defined as the share of GRIM which makes players indifferent between the two strategies. Let π^{DF} denote the probability of playing GRIM. Under perfect monitoring, indifference between GRIM and ALLD requires $\pi/(1-\delta) - (1-\pi)l = \pi(1+g)$. Hence, the BAD is defined as:

$$\pi^{DF} \quad = \quad \frac{l}{l - g + \frac{\delta}{1-\delta}} \tag{1}$$

In contrast to the SPE condition, the BAD also takes the *sucker* payoff $-l$ into account. The BAD is inversely related to the frequency of cooperation observed in laboratory experiments with perfect monitoring (Dal Bó and Fréchette, 2018a). Blonski et al. (2011) use an axiomatic approach to derive a critical value of δ for the emergence of cooperation which is also related to strategic uncertainty.[7] The δ-threshold of Blonski et al. (2011) turns out to be the value of δ that makes cooperation risk-dominant in the sense of Harsanyi and Selten (1988).

cooperative state. This yields $1 + \frac{\delta}{1-\delta} \geq 1 + g$ under perfect monitoring, and $\frac{1}{1-\delta(1-\epsilon)^2} \geq \frac{1+g}{1-\delta\epsilon(1-\epsilon)}$ under imperfect public monitoring.

[6]Under private monitoring, players lack a public history of play to coordinate behavior in such a way that defection is the mutual best-response after a defection signal. Instead, a player who believes that the other player is still in the cooperative state and obtains a defection signal would want to ignore the defection signal to keep the partner in the cooperative state. The incentive to ignore defection signals undermines the necessary responsiveness of the cooperative strategy to defection.

[7]Blonski et al. (2011) use five axioms to identify a critical value of δ as a function of the three incentives: the long-run incentive to cooperate ($\frac{\delta}{1-\delta}$), the short-run incentive to defect if the opponent cooperates (g), and the short-run incentive to defect if the opponent defects as well (l). The δ-threshold equals $\frac{g+l}{1+g+l}$, and corresponds to critical value of δ that makes cooperation risk-dominant. This can be verified by setting $\pi^{DF} = 0.5$ and solving for the continuation probability.

The role of strategic uncertainty for the emergence of cooperation under imperfect monitoring is not well understood. We analyze the robustness of cooperative strategies to uncertainty under imperfect monitoring. In Appendix A, we derive lower bounds for the BAD under public and private monitoring. To derive the lower bounds, we use variants of GRIM that are most robust to strategic uncertainty as they defect forever after a defection signal.[8] Lenient or forgiving versions of GRIM are more vulnerable to defection and result in higher values of the BAD. This highlights the tradeoff between the efficiency of cooperative strategies and their robustness to strategic uncertainty under imperfect monitoring. The lower bounds of the BAD under public and private monitoring suggest that, in the frequently studied cases $g = l$ and $1 + g = l$, the negative impact of strategic uncertainty on cooperation is amplified under imperfect monitoring.

2.2 Strategies

The strategy to defect in every round of the repeated game is a SPE under all three monitoring conditions. Under perfect monitoring, players can condition on the intended actions and the strategy GRIM is sufficient to support full cooperation. Under imperfect public monitoring, players can rely on a variant of GRIM that conditions on the noisy, public signals. This implies reduced efficiency since defection occurs with positive probability on the equilibrium path. Lenient and (or) forgiving strategies counteract the efficiency loss caused by the monitoring imperfections. Fudenberg et al. (2012) show that participants of a laboratory experiment use lenient and forgiving strategies under imperfect public monitoring despite the fact that these strategies are more risky under strategic uncertainty.

For imperfect private monitoring, the theoretical literature has repeatedly pointed out the the role of mixed (Bhaskar and Obara, 2002; Sekiguchi, 1997) or behavior strategies (Ely and Välimäki, 2002; Piccione, 2002). Yet, these strategies also play a role under perfect monitoring. In fact, an alternative way of constructing predictors of cooperation is to identify conditions under which cooperative behavior strategies can be supported as SPE. Breitmoser (2015) shows that the δ-threshold for the emergence of cooperation identified by Blonski et al. (2011) is the existence condition of equilibria in semi-grim strategies. Semi-grim strategies are a sub-class of M1BF strategies (Ely and Välimäki, 2002; Ely et al., 2005; Piccione, 2002) which are studied in both the theoretical and the experimental literature on repeated games

[8]Since playing GRIM cannot be supported as symmetric pure-strategy equilibrium under private monitoring, the BAD should rather be interpreted as the mixing probability in Sekiguchi's (1997) construction of a belief-based equilibrium.

(Breitmoser, 2015; Aoyagi et al., 2018; Heller, 2017). Breitmoser (2015) provides empirical evidence that under perfect monitoring, behavior on both the aggregate, and the individual level is well summarized by semi-grim strategies.

The multiplicity of equilibrium strategies in the indefinitely repeated prisoner's dilemma prompts the question which strategies provide good descriptions of actual behavior. To answer this question, strategies can be estimated from experimental data (see Dal Bó and Fréchette, 2018b, for a review). Behavior strategies create a problem in this context. As behavior strategies have more free parameters, they have an advantage over pure strategies to explain observed behavior. One solution is to penalize the strategy models estimated from data with information criteria that take the number of free model parameters into account (see, for example, Breitmoser, 2015).

We propose an alternative solution based on theory. In Appendix A, we derive conditions that make it possible to pinpoint the behavior of M1BF equilibrium strategies for all three monitoring structures. We show that all M1BF equilibria predict the same behavior after round one if the continuation probability δ is at the threshold of the existence condition of those equilibria. The common pattern of M1BF behavior is then defined by the stage-game parameters. If $l > g$, the response is a lenient form of tit-for-tat. If $g > l$, it is a forgiving form of GRIM. In the frequently studied case $g = l$, it is the tit-for-tat response. We call this behavior the T1BF response, for threshold memory-one belief free. The T1BF pattern allows us to estimate M1BF strategies from our data which have the same number of free parameters as pure strategies.

2.3 Communication

Renegotiation-proofness refinements (Pearce, 1987; Farrell and Maskin, 1989) are the most widely used tools to restrict the usually large set of equilibria in repeated games that allow for cooperation. Weak renegotiation proofness (Farrell and Maskin, 1989) requires that an equilibrium strategy profile must not have continuation values in any subgames that are Pareto-dominated by the continuation values in another subgame – the idea being that subjects would otherwise renegotiate away from the former to the latter. Equilibria that support cooperation with strongly symmetric strategies, such as equilibria where both players defect in the punishment state, do not survive this refinement because players would otherwise renegotiate in this state and restart the game. However, weakly renegotiation-proof cooperative equilibria often exist in the indefinitely repeated prisoner's dilemma (van Damme, 1989). They require more complex behavior in the punishment phase, where players have to play asymmetrically. The player that has deviated must play C while the punisher plays D. After a certain number of rounds, the punishment phase ends and play restarts with mutual cooperation.[9] Such an equilibrium is arguably more complicated to coordinate on, which has led some authors to restrict attention to strongly-symmetric strategies. Embrey et al. (2013), for example, adapt Pearce's (1987) slightly different renegotiation-proofness concept to derive predictions for a game with imperfect monitoring. In their variant of renegotiation-proofness *"a candidate equilibrium would survive potential renegotiation if there is no other perfect public equilibrium, using strongly symmetric two-state automata, that has a larger expected value in the punishment state"* (p.11). In addition to considering only strongly symmetric strategies, they restrict their attention to perfect public equilibria. The reason for this additional restriction is that renegotiation concepts rely on the existence of multiple subgames, which requires that subjects condition their play on the public history. If subjects instead condition on private histories, as they do in the belief-based equilibrium construction by Sekiguchi (1997) or in belief-free equilibria (Piccione, 2002; Ely and Välimäki, 2002), the only subgame is the entire game.

Predictions that are based on these refinements have been tested in a number of experimental studies with repeated communication. The results are mixed. Cooper and Kühn (2014a), who study two-stage games, and Embrey et al. (2013) find no reduction in cooperation when renegotiation-proofness predicts less cooperation.[10] Andersson and Wengström (2012), also studying a two-stage game with structured communication, find that pre-play messages are more effective if renegotiation between the two periods is not possible. They observe slightly lower cooperation rates with repeated as compared to pre-play communication. Cooper and Kühn (2014a and 2014b) compare treatments with structured and free-form communication via a chat interface, and find that cooperation rates are higher with free-form communication (see also Bigoni et al., 2018; Kartal and Müller, 2018).

As discussed above, in the absence of communication there are only complicated equilibrium constructions under private monitoring. However, when players can communicate repeatedly, private signals can be reported, which creates a public history and thereby allows for simpler and more stable equilibria (Heller, 2017).

[9] For the existence of such equilibria l must not be too large relative to the value of returning to the CC state, which obviously depends on δ. Otherwise, the return to CC is not attractive enough for the punished to agree on playing C and receiving $-l$ in the punishment phase. We derive renegotiation-proof equilibria in Appendix A.

[10] Fonseca and Normann (2012), also studying a two-stage game, and Camera et al. (2013), studying a game with random re-matching in groups after every round, find a positive rather than a negative effect of repeated communication on cooperation. Neither of the two studies explicitly tests renegotiation-proofness predictions.

Such *truthful communication* equilibria can exist if certain revelation constraints are fulfilled (Compte, 1998). The punishment stage is constructed in a way that makes every player indifferent between truthfully reporting the private signal and misreporting or staying silent. This requires that no player benefits or suffers from entering the punishment phase in which the other player is punished.[11] The stability of these equilibria stems from the fact that they provide strict incentives for cooperation, whereas the other equilibrium constructions by Sekiguchi (1997), Piccione (2002), or Ely and Välimäki (2002) do not (see Heller, 2017).[12]

3 Experimental Design

Our experiment follows a 3 (monitoring: perfect, imperfect public, imperfect private) $\times$ 3 (communication: none, pre-play, repeated) between-subject design with 9 experimental treatments. We implement the three different monitoring conditions following Aoyagi et al. (2018). Under *perfect* monitoring both players are informed about the intended actions (a_i, a_{-i}) and the signals (ω_i, ω_{-i}). Under *public* monitoring, players are given the reduced information set $(a_i, \omega_i, \omega_{-i})$. Under *private* monitoring, players are informed only about (a_i, ω_{-i}).

In addition to the three different monitoring conditions, we implement three different communication conditions. The benchmark case is that of *no communication* (as in Aoyagi et al., 2018). In the *pre-play communication* condition, subjects enter an open-chat communication stage before the first round of a supergame. The chat can be used by both players of the current match to exchange messages for 120 seconds. In the *repeated communication* condition, players additionally enter a communication stage before each of the following rounds where they can exchange messages for 40 seconds.

To keep the length of the supergames constant between treatments, we generate two sequences of supergames beforehand using a series of random numbers to determine the length of each supergame.[13] Both sequences are implemented for

[11] We show how such an equilibrium can be constructed in Appendix A.

[12] If signals are correlated, which is not the case in our set-up, *truthful communication* equilibria with strict revelation constraints can be constructed (Kandori and Matsushima, 1998), and higher levels of efficiency might be achievable by exploiting the informational content of the correlation (Awaya and Krishna, 2016). Awaya and Krishna (2016) study a set-up with a fixed discount rate, whereas other studies have focused on proving Folk theorems (Ben-Porath and Kahneman, 1996; Compte, 1998; Kandori and Matsushima, 1998; Obara, 2009).

[13] We use the Stata random number generator with seeds 1 and 2 to create two series of uniformly distributed random numbers between 0 and 1. The first supergame had x rounds if the xth random number was less than or equal to 0.2 and all previous numbers were greater than 0.2. Then the first x random numbers were deleted and the following numbers determined the length of the second

all treatments in different sessions. At the end of every round of a supergame, subjects receive feedback about their earnings and additional information which allows them to (imperfectly) monitor others' decisions. The realized random number, which determines whether the supergame continues or not, is also displayed at the end of each round, and could thus be used as a public randomization device. To allow for learning, each participant in our experiment plays seven supergames with different partners. The matching proceeds as follows: we divide the subjects of an experimental session into matching groups of 8–12 subjects. For the first supergame, each subject is then randomly matched with another participant from their matching group. After the termination of a supergame, participants are re-matched with a new partner from their matching group who they did not interact with before. Subjects were informed about this matching procedure. Before the start of the treatment, participants had to answer control questions to check their understanding of the instructions (see Appendix D).

Table 1: Summary Statistics for the Experimental Treatments

	Perfect			Public			Private		
	No	Pre	Rep	No	Pre	Rep	No	Pre	Rep
Sessions	2	2	2	3	3	4	2	3	3
Matching groups	6	6	6	6	6	6	6	6	6
Subjects	52	54	54	48	52	50	48	50	50
Mean group size	8.7	9.0	9.0	8.0	8.7	8.3	8.0	8.3	8.3
Mean supergame length	5.6	5.6	5.6	5.6	5.6	5.6	5.6	5.6	5.6

Notes: Mean group size indicates the average number of subjects who formed a matching group. The modal size of a matching group was eight (44 groups). Seven groups were of size 10 and three of size 12. Subjects did not know the exact size of their matching group. Mean supergame length indicates the average number of rounds of all supergames played in a treatment.

We collected data from three matching groups per sequence-treatment combination, that is from six matching groups per treatment. A total of 458 participants (average age 22, 60% female) participated between January and April of 2016 in the 24 sessions of our experiment at the LakeLab of the University of Konstanz.[14] The average earning was EUR 18, and the session length 75–90 minutes. Table

supergame, and so forth. We used the two series to determine the lengths of seven supergames each. The length of the two resulting sequences of supergames are: SQ1 (11 3 5 1 5 2 11) and SQ2 (2 5 5 7 13 4 4). Average supergame lengths were moderately longer than the expected length of five of the underlying geometric distribution (SQ1: 5.4; SQ2: 5.7). Random termination is the most widely used way of implementing infinitely repeated games in the lab. See Fréchette and Yuksel (2017) for a study of other implementation methods.

[14]The experiment was programmed in z-Tree (Fischbacher, 2007) and subjects were recruited via ORSEE (Greiner, 2015).

Figure 1: Stage-Game Parameters and Predictors of Cooperation

	c	d
C	30	0
D	37	17

	C	D
C	27,27	3,35
D	35,3	19,19

	Perfect	Public	Private
δ	0.80	0.80	0.80
ϵ	0.10	0.10	0.10
δ^{SPE}	0.50	0.65	–
π^{DF}	0.40	0.76	0.77
δ^{SG}	0.75	0.86	0.86
δ^{BF}	0.67	0.80	0.80

Notes: The payoffs depicted in the left panel are in experimental currency units. The exchange rate was 50 ECU = EUR 1. Subjects saw both representations of the stage-game at all times when making their decisions. The right panel shows the continuation and error probability, the cooperative SPE condition for pure strategies δ^{SPE}, and the basin of attraction π^{DF} for all three monitoring conditions. For public and private monitoring the values of π^{DF} are the lower bounds derived in Appendix A. The lower part of the right panel shows the existence conditions for semi-grim δ^{SG}, and M1BF equilibria δ^{BF} (see Appendix A for details).

1 summarizes the distribution of sessions, subjects, and matching groups across experimental treatments and depicts the average size of a matching group as well as the average length of the supergames.

3.1 Experimental Parameters

The upper left panel of Figure 1 shows the payoff in experimental currency units that a subject receives in a round of a supergame as a function of the action and the signal received about the other player's action. We use the same payoff structure, the same continuation probability of $\delta = 0.8$ and the same error probability of $\epsilon = 0.1$ in all treatments. These values translate into expected stage-game payoffs for actions depicted in the lower left panel.

The right panel of Figure 1 shows the values of the predictors of cooperation which result from the parameters. We choose the parameters for the following reasons: First, the parameters are such that without communication we expect low levels of cooperation under imperfect monitoring and a slightly higher but still low level under perfect monitoring. These expectations are formed on the basis of our analysis of the BAD and the cooperation rates in other studies with different levels of BAD as reviewed in (Dal Bó and Fréchette, 2018a). This leaves scope for higher cooperation levels in the communication treatments.[15] Second, we

[15]Our no-communication treatments complement the treatments of Aoyagi et al. (2018) where the BAD takes lower values of 0.03 (0.06), 0.15 (0.53), 0.13 (0.43) for perfect, public, private

want to focus on the main difference between the public and private monitoring treatments identified in the theoretical literature. This is the possibility of supporting cooperation based on pure strategies with public signals. We choose parameters, that lead to a similar BAD with public and private monitoring. The parameters also rule out that the set of M1BF equilibria is different between public and private monitoring since no M1BF equilibria exists, where strategies condition on the public signals (see Appendix A for details). Third, to pinpoint the behavior of M1BF strategies, the parameters predict the same T1BF pattern for M1BF equilibria under perfect and imperfect monitoring. This pattern is a lenient version of tit-for-tat with choice probabilities $\sigma = (\sigma_\emptyset, 1, 0.5, 1, 0)$ conditional on the memory-one action-signal histories $(\sigma_\emptyset, Cd, Cd, Dc, Dd)$. Finally, we are interested in whether subjects use communication to transform the game with private monitoring into one with public signals. Our parameters assure that equilibria exist where players truthfully reveal their private signals under private monitoring. They also assure the existence of renegotiation-proof cooperative equilibria under perfect and public monitoring (see Appendix A).

3.2 Research Questions and Methods

We state three main research questions. As we explain below, we expect different answers to them for the three monitoring regimes.

Question 1: *Does pre-play communication effectively increase cooperation rates?*

According to our measures, strategic uncertainty is high in our parametrization in all three monitoring structures. However, while strategic uncertainty has been shown to matter, at least under perfect monitoring without communication, it has also long been recognized that communication can help coordination (e.g., Cooper et al., 1992; Rabin, 1994; Ellingsen and Östling, 2010) and that coordination on a cooperative equilibrium would decrease strategic uncertainty (Kartal and Müller, 2018). Therefore, we expect pre-play communication to facilitate coordination and thereby to lower strategic uncertainty. However, while efficient equilibria are easy to find in the perfect monitoring case, this task becomes a lot more difficult under imperfect public monitoring. Even if players cooperate, bad signals occur with positive probability and thus players will likely have to enter a phase of punishment at some point. For this reason, simple punishments, such as "defect forever" after

monitoring, in their low (high) noise treatment, respectively.

a bad signal, are inefficient and players have to coordinate on lenient or forgiving strategy profiles to reap a greater share of the potential gains of cooperation.

With private monitoring it becomes even more complicated. The equilibria that have been found and analyzed in the literature are all mixed (or behavior) strategy profiles, which are extremely hard to find, and coordination on these equilibria seems highly unlikely (Compte and Postlewaite, 2015). So, while we expect a positive effect of pre-play communication on cooperation rates across all monitoring structures, compared to the no communication treatments, the effect might be more pronounced under perfect monitoring than under imperfect monitoring, and more pronounced under public than under private monitoring. To address *Question 1*, we compare the average frequency of cooperation and the average stability of cooperation over rounds between no communication and pre-play communication within the same monitoring structure. In addition, we measure coordination by the frequency of choices of the same action within a matched pair in round one of a supergame, and compare these frequencies between treatments.

Question 2: *Is repeated communication crucial for stable cooperation?*

For the case of private monitoring, Heller (2017) shows that only defection can be sustained, by any of the mechanisms discussed in the literature, as an equilibrium that survives his weak stability criterion. He further shows that if players can communicate repeatedly, there typically are cooperative *truthful communication equilibria*, which are weakly stable and even survive the stronger criterion of evolutionary stability. In our parametrization, this is the case. Moreover, stable cooperative equilibria also exist without communication under public and perfect monitoring. We would thus expect a large positive effect of repeated communication on cooperation under private monitoring, whereas high cooperation is already achievable in stable equilibria without communication under public and perfect monitoring. In the latter two monitoring structures, weak renegotiation-proofness (Farrell and Maskin, 1989) eliminates cooperative equilibria in strongly symmetric strategies and repeated communication might thus have a negative effect. However, there are cooperative equilibria that are weakly renegotiation-proof. Finally, repeated communication might have an additional benefit under imperfect monitoring where coordination on an efficient equilibrium is difficult and players might need to revisit incomplete agreements after round one, in particular when a bad signal occurs for the first time.

To answer *Question 2*, we execute the same test sequence outlined for *Question 1* for the differences between pre-play and repeated communication treatments. Further, we compare the stability of cooperation across different monitoring structures.

Question 3: *What are the mechanisms through which communication affects behavior?*

To answer this question we analyse the communication content and subjects' strategy choices.

Communication Content We expect pre-play communication to be used for coordination to reduce strategic uncertainty. However, other uses are conceivable and this analysis will therefore mainly be of an exploratory nature. Under imperfect private monitoring, we expect a very specific and important role for repeated communication. The sharing of private information is the key role ascribed to communication under private monitoring in the recent theoretical literature (e.g., Compte, 1998; Kandori and Matsushima, 1998; Awaya and Krishna, 2016; Heller, 2017). The reduction of uncertainty regarding the history of play is crucial in this context, which could also play an important role under imperfect public monitoring. Two research assistants coded the content of communication based on 72 sub-categories, from which we created five main categories (Table B1 in Appendix B illustrates the assignment of sub-categories to categories). The five main categories are Coordination, Deliberation, Relationship, Information and Trivia. The coding was done on the sub-category level for subject-round observations and multiple coding was possible. We consider a coding as valid only if both raters independently indicated the same sub-category for a subject-round observation. To get an overview of the use of communication across treatments, we focus on the category level. For the more detailed analyses on information sharing and communication after different histories we study the sub-categories.

Strategies To answer the question how communication opportunities affect strategies, we adapt and use the SFEM by Dal Bó and Fréchette (2011) to explore the use of strategies in our treatments (see Appendix C for details). We estimate the strategy shares of a set of pure strategies and three behavior strategies in our treatments using standard SFEM. A graphical illustration of each strategy can be found in Tables C1-C4 of Appendix C. The set of pure strategies is taken from Fudenberg et al.'s (2012) study on the indefinitely repeated prisoner's dilemma with imperfect public monitoring. Two of the three behavior strategies are motivated by Backhaus and Breitmoser's (2018) analysis, who present evidence suggesting that subjects play semi-grim M1BF strategies, and further find that a small share of (noise) players randomize 50–50 in all states. The third behavior strategy is a M1BF equilibrium strategy we derive theoretically in Appendix A.

To account for the fact that participants gain experience over supergames, we also study the evolution of strategy frequencies over the first three supergames using a new class of strategy estimation models which we develop borrowing from the literature on latent-class regression (Dayton and Macready, 1988; Bandeen-Roche et al., 1997). The new model class assumes that subjects' strategy choices over supergames reflect repeated independent draws from a probability distribution over a fixed set of candidate strategies. The probability distribution over strategies is modeled as a function of covariates that explain the selection of strategies by individuals. We describe the new model class in more detail in Appendix C.[16]

A disadvantage of imposing a set of candidate strategies to describe subjects behavior is that the results might be sensitive to the composition of the candidate set. To assess the robustness of our strategy estimation results, we also infer behavior and pure strategies from the data and compare the results with those of the SFEM. Finally, we assign subjects to predefined strategies following Camera et al.'s (2012) classification approach.[17]

4 Experimental Results

A common result in the experimental literature is that participants need a few supergames to adapt their behavior to the experimental environment (e.g., Dal Bó, 2005). We also observe a considerable amount of learning over supergames (see Figure 2). For our analyses of cooperation rates, we will therefore report the results of the last three supergames, where participants' behavior has largely stabilized, in addition to the results of all supergames. Our strategy estimations rely on the standard assumption of a stable distribution of strategies across supergames. Therefore, we mainly focus on the last three supergames for these analyses. However, we also study the evolution of strategy choices in the beginning of an experimental session by fitting strategy estimation models with time as a covariate of strategy choice.

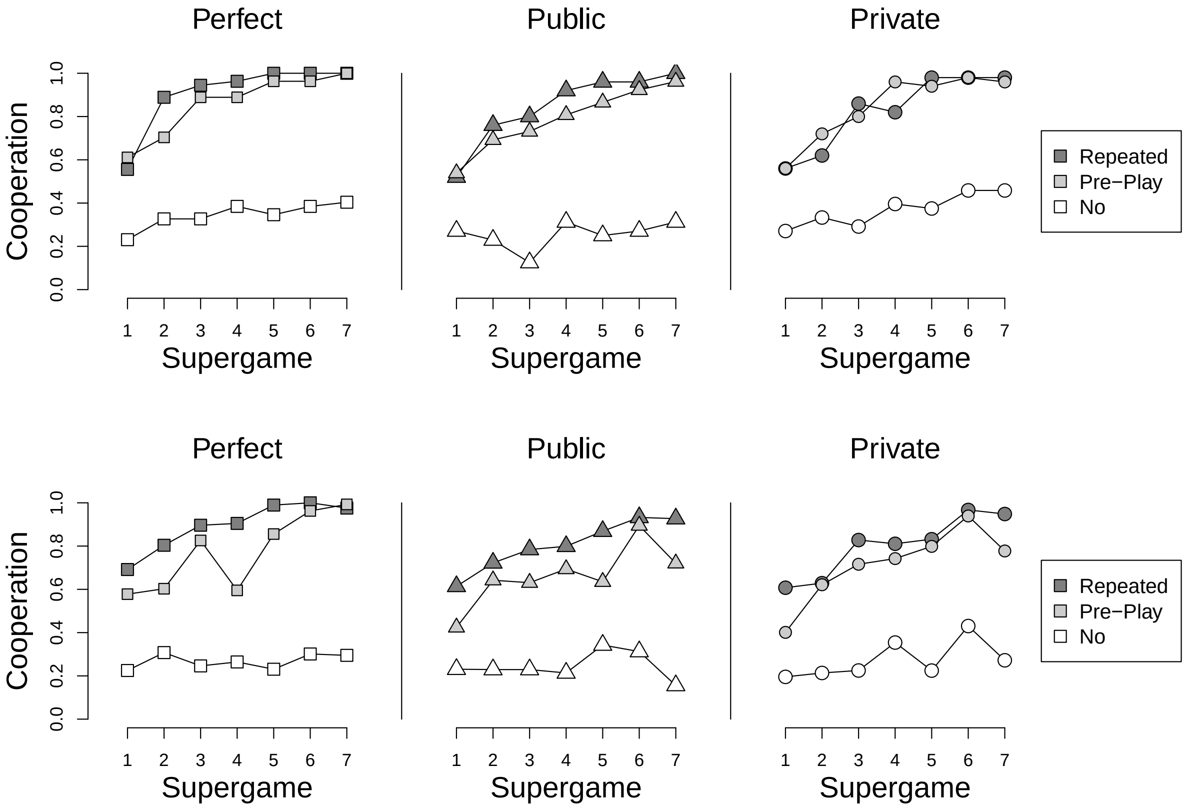
[16] See Dvorak (2020) for an introduction to the R package *stratEst* which implements the method.

[17] Another approach to understanding strategy choices in indefinitely repeated games is direct elicitation of strategies (Bruttel and Kamecke, 2011; Dal Bó and Fréchette, 2018b; Romero and Rosokha, 2018).

4.1 Cooperation

Figures 3 and 4 present two measures of cooperation: the average frequency of cooperation, and the average stability of cooperation over rounds. We provide answers to Questions 1 and 2 based on these two figures. The reported p-values, p_{all} (p_{l3}), result from one-sided t-tests of logistic regression coefficients with two-way clustered standard errors at the participant-match level (Cameron et al., 2011), including all (the last three) supergames.

Figure 2: Evolution of Cooperation over Supergames

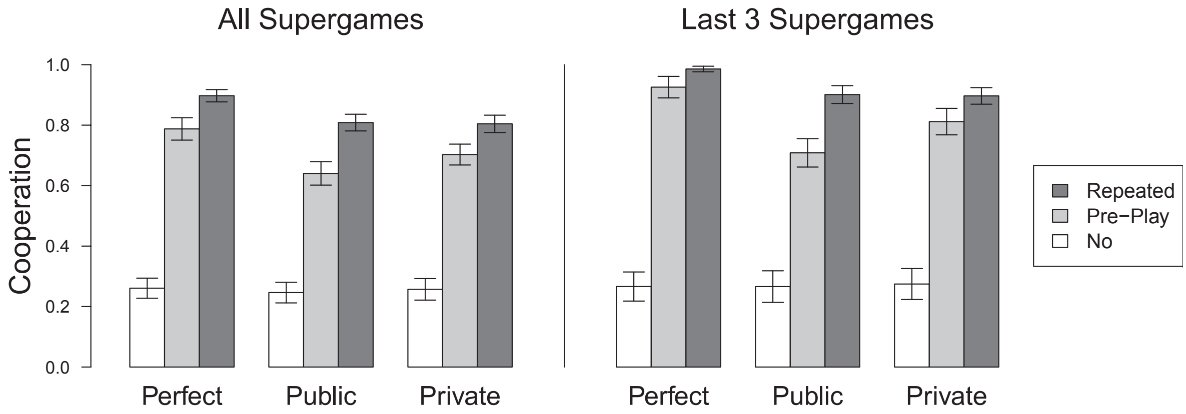


Notes: The upper three panels display average cooperation rates in round one over the seven supergames. The lower three panels display overall average cooperation rates in the seven supergames.

Question 1: *Does pre-play communication effectively increase cooperation rates?*

Figure 3 shows the average frequency of cooperation across the nine experimental treatments. The depicted levels of cooperation mostly reflect the amount of cooperation observed in the first four rounds, where each participants contributes two or three observations depending on the length of the supergames played. The bars

Figure 3: Average Frequency of Cooperation Across Treatments



Notes: Bars show the relative frequency of cooperation. Whiskers depict two-way clustered standard errors of the mean (clustered on subject and match).

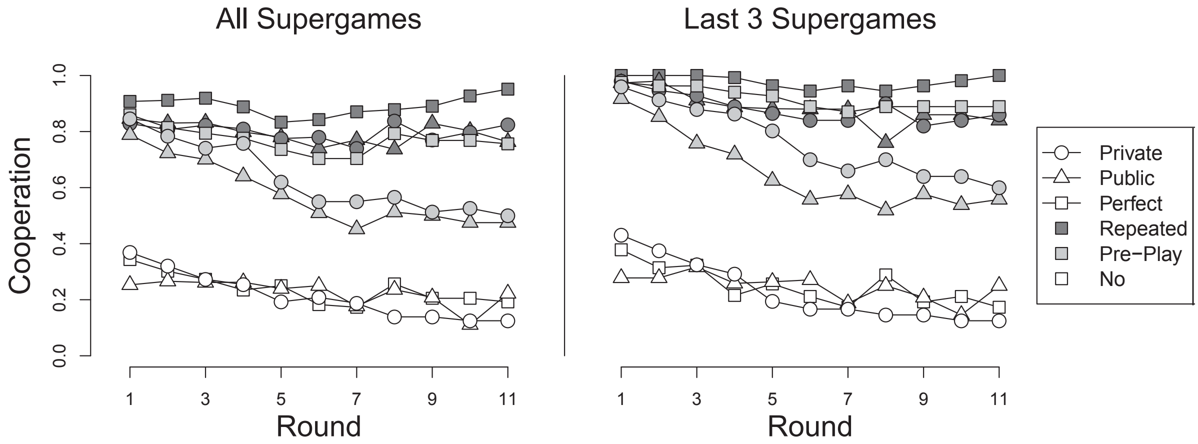
indicate that the mean cooperation level in the treatments with pre-play communication is substantially higher compared to the treatments without communication ($p_{all} < 0.01$, $p_{l3} < 0.01$). The effect of pre-play communication on the average cooperation rate is largest under perfect monitoring. Under perfect monitoring, the average cooperation rate is 53 percentage points higher with pre-play communication (66 pp, last three sg). Under imperfect public monitoring, the average cooperation rate is 39 percentage points higher with pre-play communication (44 pp, last three sg). Under imperfect private monitoring, the average cooperation rate is 44 percentage points higher with pre-play communication (54 pp, last three sg). Difference in differences tests indicate that the effect of pre-play communication is larger under perfect compared to imperfect public monitoring (perfect vs. public: $p_{all} = 0.04$, $p_{l3} < 0.01$; perfect vs. private: $p_{all} < 0.12$, $p_{l3} = 0.06$).

Result 1: *Pre-play communication leads to a large increase in cooperation rates under all three monitoring structures.*

Question 2: *Is repeated communication crucial for stable cooperation?*

Figure 3 shows that the mean cooperation level in treatments with repeated communication is higher compared to the treatments with pre-play communication (perfect: $p_{all} < 0.01$, $p_{l3} = 0.02$; public: $p_{all} < 0.01$, $p_{l3} < 0.01$; private: $p_{all} = 0.01$, $p_{l3} = 0.04$). The size of the effect is largest under public monitoring where the mean cooperation level is 17 percentage points higher (19 percentage points in

Figure 4: Stability of Cooperation over Rounds



Notes: The graph depicts the frequency of cooperation over rounds averaged over all supergames (the last three supergames). The average number of observations per treatment decreases from 355 (153) in round one to 77 (51) in round eleven because of the different lengths of the supergames (see footnote 16).

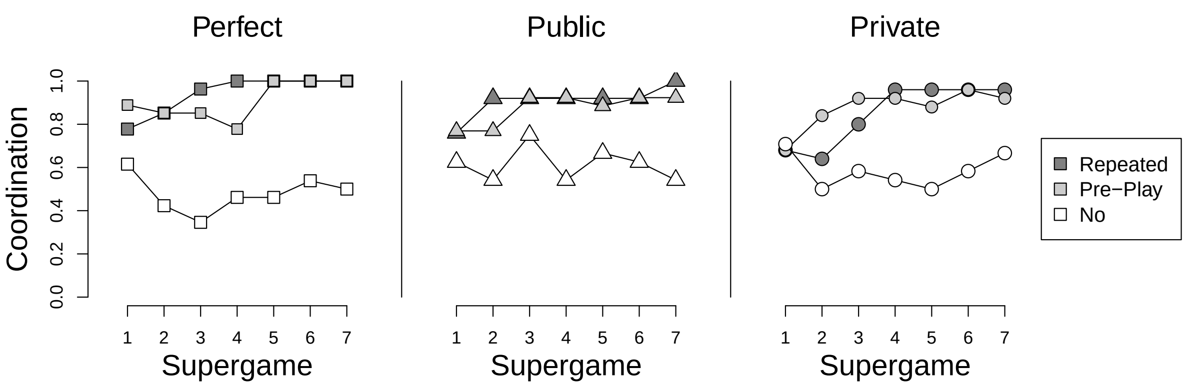
the last three supergames) with repeated communication. Difference in differences tests indicate that the effect of repeated communication is similar across all three monitoring structures (perfect vs. public: $p_{all} = 0.42$, $p_{l3} = 0.49$; perfect vs. private: $p_{all} = 0.26$, $p_{l3} = 0.19$).

Figure 4 shows the mean cooperation level over rounds. The mean cooperation level is depicted up to round 11 to assure that each participant contributes at least one observation to every estimate. The slopes of the lines indicate that the decline of cooperation over rounds varies between treatments. With communication, the cooperation rate in round one is around 85% (95% in the last three supergames) and does not differ much between the monitoring structures. In the treatments with repeated communication, the cooperation rate is generally more stable over rounds compared to the treatments with only pre-play communication. The differences in the stability of cooperation between the repeated and the pre-play communication treatments are more pronounced under imperfect monitoring. Under imperfect monitoring with pre-play communication, the average cooperation rate is between 31 to 36 percentage points lower after 10 rounds without communication opportunities, but only between 1 to 13 percentage points lower with repeated communication. In contrast, if monitoring is perfect, the average cooperation rate only reduces by around 10 percentage points with pre-play communication, and does not decline at all with repeated communication. Without communication, cooperation declines over the rounds of a supergame at a similar rate under all three monitoring structures.

To test whether the stability of cooperation over rounds differs between the pre-play and repeated communication treatments, we regress cooperation on the round number in probit regressions and perform tests on the treatment specific coefficients. The results indicate that cooperation is only more stable with repeated communication if monitoring is imperfect (perfect: $p_{all} = 0.15$, $p_{l3} = 0.34$; public: $p_{all} < 0.01$, $p_{l3} = 0.18$; private: $p_{all} < 0.01$, $p_{l3} = 0.08$). Moreover, if we compare the treatments with pre-play communication, we find that the decline of the average cooperation rate is steeper under the imperfect monitoring (perfect vs. public: $p_{all} = 0.01$, $p_{l3} = 0.13$; perfect vs. private: $p_{all} < 0.01$, $p_{l3} = 0.06$). Finally, Figure 4 shows no differences in the stability of cooperation between the perfect and the imperfect monitoring structures without communication.

Result 2: *While pre-play communication is sufficient for reaching a high and stable cooperation rate under perfect monitoring, repeated communication is necessary for stable cooperation under both imperfect monitoring structures.*

Figure 5: Evolution of Coordinated First Round Choices



Notes: The graph depicts the evolution of coordination on the same action in round one over supergames. The lines indicate the share of pairs, in which both participants choose the same action in round one.

Figure 5 depicts the evolution of coordinated first round choices over supergames. The evolution of coordinated first round choices illustrates that, with communication, participants eventually manage to coordinate their first round choices under every monitoring condition. This indicates that the differences in the effect of pre-play communication between the monitoring structures do not stem from an inability to coordinate behavior in round one. Instead, the differences in the effect of pre-play communication across the monitoring structures stem from the decline of cooperation depicted in Figure 4. Given the complexity of cooperative equilibria under public and, especially under private monitoring, it is rather surprising that participants perfectly coordinate their first round choices under such conditions. In the treatments without communication, the level of coordination of first round choices is rather low and does not increase over supergames. These results suggest that communication is, in the long run, equally effective in reducing strategic uncertainty under every monitoring condition. This interpretation is further substantiated by our communication content analysis (see Section 4.2).

4.2 Mechanisms

Question 3: *What are the mechanisms through which communication affects behavior?*

Communication Content Our two raters made an average of 2.65 classifications into 72 sub-categories per participant-round observation, resulting in 18,678 and 18,984 classifications in total. To make the interpretation of the communication content easier, we collapse the 72 sub-categories into five main categories: Coordination, Deliberation, Relationship, Information and Trivia.[18] The average Cohen's κ across treatments is above 0.7 for all five main categories, which indicates a high level of agreement between our raters. Table 2 reports their relative frequency in the last three supergames. Frequencies are very similar when all supergames are considered (see Table B1, Appendix B). Overall, we observe that the frequencies of the categories in round one of the repeated communication treatments (column Rep-f) are similar to those of pre-play communication. This indicates that the communication phase before the first round is used similarly by the participants of the pre-play and repeated communication treatments.

The category Coordination includes all attempts by participants to coordinate behavior in future rounds. The category also includes implicit or explicit announcements of choices since such announcements could also be used to coordinate behavior. The category occurs in the vast majority of participant-round observations of the pre-play phase. Its relative frequency in the later rounds of the repeated communication treatments is lower, which suggests that coordination predominantly occurs before the first round. The category Deliberation includes all instances in which participants discuss choices or strategies. Our raters indicate content related to deliberation in roughly every second participant-round observation with pre-play communication.

[18]See Tables B1 and B2 in Appendix B for the mapping of sub-categories to main categories, the frequency of occurrence of messages in the (sub-)categories and the average Cohen's κ (across treatments) of all categories and sub-categories.

In the repeated communication treatments, content related to deliberation becomes less frequent after round one. All content that concerns the relationship of a matched pair of participants is included in the category Relationship. The category also covers motivational talk and positive feedback that we find to be quite common. Content related to this category is more frequent under imperfect monitoring. In contrast to the categories Coordination and Deliberation, the category Relationship does not become less frequent after round one. The category Information includes all statements that contain reports of actions, signals or payoffs from the current supergame. It is not possible to report such information before round one. The category ranges among the most frequent with repeated communication. Note that content related to the Information category does not imply that participants reveal private information. In order to assess whether they report private information, we will look at data from sub-category level in the following. Our last main category, Trivia, contains content that is off-topic or classified as small talk by our raters. In contrast to the Relationship category, the content does not have an obvious relation to the game. The Trivia category is always among the most frequent in all treatments.

Table 2: Frequency of Codings per Individual-Round Observation

	Perfect			Public			Private		
	Pre	Rep-f	Rep-l	Pre	Rep-f	Rep-l	Pre	Rep-f	Rep-l
Coordination	0.98	1.00	0.11	0.99	0.99	0.21	1.00	0.97	0.28
Deliberation	0.60	0.51	0.09	0.72	0.69	0.13	0.63	0.72	0.12
Relationship	0.14	0.23	0.28	0.31	0.15	0.45	0.36	0.19	0.42
Information	–	–	0.31	–	–	0.50	–	–	0.49
Trivia	0.99	0.99	0.57	0.94	0.93	0.38	0.85	1.00	0.55

Notes: Level of the analysis are individual-round observations. "Rep-f" (Rep-l) indicates the first (later) rounds in the repeated communication treatments. The data is from the last three supergames. A coding is considered as valid if both raters indicated the same category for a participant-round observation.

Looking at the sub-categories for coordination before the first round of an interaction, we see that the suggestion to play CC is made by roughly half of all participants and in almost all pairs of participants in all communication treatments. Some participants suggest DD but these suggestions occur at a frequency below 10% in all treatments. More complex suggestions for coordinated play or explicit or implicit threats of punishment in the case of defection occur at even lower frequencies. These observations highlight that most pairs of participants enter the game without an agreed-upon plan for how to deal with defections or bad signals in the imperfect monitoring treatments. It seems plausible that this incomplete coordination on an efficient equilibrium explains the decline in cooperation in the pre-play communication treatments under imperfect monitoring.

Table 3: Communication after First Defection Signal

	Public Repeated			Private Repeated		
Category	$\omega \neq \{c, c\}$	$\omega = \{c, c\}$	Δ	$\omega_j = d$	$\omega_j = c$	Δ
Coordination	0.32	0.17	0.15	0.52	0.24	0.27
Deliberation	0.11	0.10	0.01	0.01	0.09	-0.08
Relationship	0.33	0.36	-0.02	0.27	0.32	-0.04
Information	0.67	0.32	0.34	0.73	0.36	0.36
Trivia	0.49	0.61	-0.12	0.44	0.66	-0.22
Report of action	0.46	0.01	0.45	0.54	0.12	0.43
Report of C	0.46	0.01	0.45	0.54	0.12	0.43
Report of D	0.00	0.00	0.00	0.00	0.00	0.00
Report of signal	0.54	0.32	0.22	0.70	0.34	0.36
Report of c	0.07	0.32	-0.25	0.00	0.33	-0.33
Report of d	0.46	0.00	0.46	0.70	0.00	0.69

Notes: Frequency of communication categories for subject-round observations with cooperative history up to round t. A Subject has a cooperative history if her previous actions were C and all signals she observed in rounds $< t$ were c. Frequencies illustrate the use of categories dependent on signals in round t. Frequency indicates the probability that both raters indicated the category for a text unit. Frequencies < 0.001 omitted (-).

To see what happens under imperfect monitoring when participants have the opportunity to talk repeatedly after bad news, we compare the communication content in the last three supergames after an interruption of a perfectly cooperative history, that is after the first bad signal (crisis), with the content after an uninterrupted perfectly cooperative history (when things go well) in Table 3 (see Table B4, Appendix B for all supergames). Participants make more proposals regarding future play (mostly CC) in the crisis situations than when things go well, suggesting that the first defection signal generates some demand for (re)coordination. The frequency of communication related to deliberation and the relationship of the matched participants does not change in crisis situations. However, an analysis of the subcategory level reveals that the tone of the communication about the relationship of the matched participants is negatively affected by the first defection signal. We see an increase in the frequency of expressions of disappointment, and an increase in the frequency of accusations of cheating (see Tables B5 and B6 in Appendix B). At the same time, we see a drop in off-topic talk in crisis situations, a substantial increase in information exchange about signals and payoffs. Table 3 also reveals that communication about signals increases after the first defection signal. Moreover, participants frequently respond to the uncertainty triggered by the first defection signal by reporting that their previous action was C, which is a truthful exchange of private information. In many cases, this appears to be sufficient to decrease the uncertainty triggered by the defection signal to a level that prevents participants from switching to defection.

Table 4: Frequency and Truthfulness of Private Information Exchange

	Private		Public	
	p(report)	p(true)	p(report)	p(true)
Actions				
Report of action	0.15	0.93	0.08	0.94
Report of C	0.15	0.93	0.08	0.95
Report of D	0.00	1	0.01	0.86
Report of C if $\omega_i = d$	0.31	0.75	0.15	0.84
Signals				
Report of signal	0.37	0.96	-	-
Report of c	0.27	0.99	-	-
Report of d	0.10	0.86	-	-
Report of d if $\omega_{-i} = d$	0.45	-	-	-

Notes: Frequencies of coding in all participant-round observations after round one of the last three supergames. A coding is considered valid if both raters indicated the same sub-category for a participant-round observation. Values might not add up as expected due to rounding.

Table 4 takes a closer look at the exchange of private information under imperfect public and imperfect private monitoring. It depicts the frequency and the truthfulness of the exchange of private information in all communication opportunities of the last three supergames (see Table B1, Appendix B for all supergames). Under public monitoring, this concerns the actions which cannot be observed by the other player. The right columns show that an action is reported in only 8% of all participant-round observations after round one. The vast majority of reports indicate cooperation in the last round, which is true in 94% of all cases. Table 4 also lists the frequency of C reports if the signal was d. In 15% of the cases where a d signal occurs, it is followed by a report of C (truthful in 84% of cases). The left columns of Table 4 show that a similar pattern exists under private monitoring but the frequency of action reports double. One important difference concerns the interpretation of reporting C when the signal is d. Under private monitoring, the difference compared to the baseline frequency of C reports suggests that their partners reported the d signal in the first place. This indirect evidence is supported by the values in the lower part of the table. A signal is reported in 37% of all participant-round interaction after round one. Most of the reports reveal a c signal truthfully. In 10% of all participant-round interactions participants report a d signal. To put this value into perspective, remember that d signals occur very seldom because of the high level of cooperation. The last line shows the frequency of d reports when a d signal actually occurred: it is 0.45. Summarizing the results reported in Table 4, we can say that participants make use of repeated communication to exchange private information. Actions are communicated less often than signals but both reports are usually credible.

Result 3: *Communication opportunities are mainly used to* ***(a)*** *coordinate behavior before the start of the interaction, which reduces strategic uncertainty, and* ***(b)*** *to exchange information about the history of play in later rounds under imperfect monitoring, which reduces uncertainty about what has happened.*

Strategies To explore heterogeneity in strategy choices, we perform a treatment-wise strategy frequency estimation (Dal Bó and Fréchette, 2011) of a standard candidate set of pure strategies (Fudenberg et al., 2012), augmented by three M1BF strategies.[19] We assume that all strategies of the same model condition on the same information and report the model with the highest likelihood. The strategies fitted to the data of the perfect monitoring treatments condition on the action profile $\{a_i, a_{-i}\}$ observed in the previous round. The strategies fitted to the data of the imperfect monitoring treatments condition on the action-signal profile $\{a_i, \omega_{-i}\}$ observed in the previous round. For each of the nine experimental treatments, the estimation procedure selects the subset of strategies which describes participants' behavior best according to the integrated completed likelihood criterion (ICL, Biernacki et al., 2000). Table 5 depicts the estimated strategy shares and standard errors. The main result of the strategy estimation is that the shares of lenient and forgiving strategies increase substantially with communication under all three monitoring structures. Under imperfect monitoring, repeated communication further increases the use of lenient and forgiving strategies, and the strategy that always cooperates (ALLC) is used frequently.

[19] The 23 strategies are illustrated in Tables C1-C4 of Appendix C. The set of strategies by Fudenberg et al. (2012) has been used in a number of other studies (reviewed in Dal Bó and Fréchette, 2018a).

Table 5: SFEM with Behavior Strategies and Strategy Selection

	Perfect			Public			Private		
	No	Pre	Rep	No	Pre	Rep	No	Pre	Rep
ALLD	0.42	-	-	0.62	-	-	0.52	0.02	-
	(0.07)			(0.07)	-		(0.07)	(0.02)	
FC	-	-	-	-	0.09	-	-	-	-
					(0.04)				
TFT	0.11	-	-	-	-	-	-	-	-
	(0.07)		-						
PTFT	-	-	0.17	-	-	-	-	-	-
	-		(0.15)						
TF2T	0.04	-	-	-	-	-	-	-	-
	(0.03)								
T2F2T	-	-	-	-	0.18	-	-	0.44	0.16
					(0.11)			(0.09)	(0.11)
LGRIM2	-	-	0.83	-	0.29	0.22	0.21	-	0.15
			(0.15)		(0.11)	(0.12)	(0.07)		(0.09)
LGRIM3	-	-	-	0.04	-	0.32	-	-	-
				(0.03)		(0.17)			
DTFT	0.13	-	-	-	-	-	-	-	-
	(0.06)								
DTF2T	0.02	-	-	0.07	-	-	-	-	-
	(0.02)			(0.04)					
ALLC	-	-	-	-	0.18	0.31	-	-	0.43
					(0.08)	(0.17)			(0.13)
$M1BF_{eq}$	-	1	-	0.08	-	0.07	-	-	0.16
		(0.00)		(0.06)		(0.05)			(0.08)
SGRIM	0.22	-	-	0.13	0.16	-	0.27	0.43	0.10
	(0.08)			(0.05)	(0.08)		(0.07)	(0.09)	(0.07)
RAND	0.05	-	-	0.06	0.09	0.08	-	0.11	-
	(0.04)			(0.04)	(0.05)	(0.04)		(0.05)	
γ	0.05	0.01	0.01	0.07	0.07	0.03	0.06	0.02	0.05
ICL	357.09	78.11	86.95	362.93	424.45	259.33	287.42	253.86	297.30
$\ln L$	-321.48	-74.12	-59.19	-339.39	-371.71	-203.37	-273.65	-229.34	-240.35

Notes: The table reports maximum-likelihood shares of a candidate set of 23 strategies listed in Tables C1-C4 of Appendix C. Estimates are obtained assuming constant strategy use over the last three supergames. Strategies condition on action profiles in perfect treatments, and on action-signal profiles in public and private treatments. γ indicates the probability of a tremble. The table shows the shares of strategies that occur in at least one of the nine treatments after strategy selection based on the ICL information criterion. Omitted shares (-) indicate that a strategy is not among the selected strategies of the treatment. Analytic standard errors in parentheses. Values might not add up as expected due to rounding.

We analyze the evolution of strategy choices over the first three supergames by extending SFEM in the spirit of latent-class regression (Dayton and Macready, 1988; Bandeen-Roche et al., 1997). The models relax the traditional assumption of SFEM that each individual uses the same strategy across all supergames. Instead, individual strategy use is assumed to be the result of repeated independent draws from a fixed set of candidate strategies. This assumption allows to model the prior probability of using a strategy as a function of the supergame number. As for SFEM, we use the candidate set of 23 strategies and identify the subset of strategies which describes behavior in a treatment best according to the ICL criterion. Table C5 in Appendix C shows the estimated parameters of the latent-class regression models. As in Table 5, the shares of lenient and forgiving strategies are higher in the treatments with communication. Yet, the strategy that always defects, ALLD, receives significant shares in the communication treatments in the first three supergames. The latent class regression coefficients allow us to quantify and test time trends in the relative frequency of strategies over the first three supergames.[20] The coefficients in Table C5 reveal that participants generally switch to more lenient strategies, such as lenient variants of grim-trigger or tit-for-tat. However, these trends are only statistically significant in the treatments with communication.

A potential problem with SFEM is that observed behavior can be attributed only to the candidate strategies considered, which leads to a misrepresentation if participants play strategies that are not part of the set of candidate strategies (Dal Bó and Fréchette, 2018b). To check the robustness of our results, we adapt our strategy inference method to infer rather than impose a set of strategies which describe participants' behavior best for each experimental treatment.[21] This can be done for behavior as well as pure strategies. Tables C6 and C7 in Appendix C show the inferred strategies and their shares for the nine experimental treatments. The broad picture is the same as the one we saw with the other approaches.[22] We also classified behavior into strategies following Camera et al.'s approach (2012). As these results do not lead to new qualitative insights either, we report them only in Appendix C and provide no detailed discussion (see Table C8). Overall, the results of all strategy estimation procedures indicate that participants' play is substantially more lenient and forgiving with communication in the last three supergames.

[20]In Appendix C we explain how these coefficients can be transformed into the strategy shares for each round and into the changes of the shares between rounds in percentage points.

[21]See Appendix C for a description of the estimation approach.

[22]Note that most inferred behavior strategies do not show the semi-grim structure. This is even the case for the perfect-monitoring-no-communication treatment where the average M1BF suggests that subjects play a semi-grim strategy.

Result 4: *With more communication opportunities, subjects' play becomes more lenient and forgiving.*

5 Summary and Discussion

In the following, we briefly summarize our analysis, the tools developed to conduct it, and discuss our key findings.

To characterize the theoretical properties of the game with respect to strategic uncertainty, we extend existing work by Dal Bó and Fréchette (2011), Blonski et al. (2011), and Breitmoser (2015), which applies only to the perfect monitoring case. We derive new measures of strategic uncertainty for the imperfect public and private monitoring cases. For the experiment, we choose parameters that make cooperation riskier under imperfect monitoring than under perfect monitoring, and under which cooperation seems to be unlikely without communication. Our continuation probability $\delta = 0.8$ coincides with the M1BF-threshold at which a unique M1BF equilibrium exists under imperfect monitoring.[23] Moreover, stable *truthful communication* equilibria also exist with these parameters under private monitoring, while all known cooperative equilibrium constructions for this monitoring structure without repeated communication are unstable (Heller, 2017). In contrast, under perfect and public monitoring stable cooperative equilibria exist under all communication structures. These design choices allow us to address a number of important open empirical questions. The first two questions concern aggregate cooperation rates.

Question 1: *Does pre-play communication effectively increase cooperation rates?*

Question 2: *Is repeated communication crucial for stable cooperation?*

We find that pre-play communication is very effective in increasing cooperation rates under all three monitoring regimes (Result 1). We also find that cooperation under imperfect monitoring is more stable with repeated communication (Result 2). Instead of a difference between perfect and public monitoring on the one side, and private on the other, we observe a clear difference between the two imperfect monitoring structures and perfect monitoring in our pre-play communication treatments. Aggregate cooperation rates are higher under perfect than under both imperfect monitoring structures. Under perfect monitoring, cooperation rates are high and

[23] It is unique up to the probability of cooperation in the first period, which is always a free parameter in strategies that are played in belief-free equilibria.

stable over all rounds of an interaction in the pre-play as well as in the repeated-communication treatments. Under imperfect monitoring, cooperation starts on a high level with pre-play communication but then declines much more strongly over time as compared to the repeated-communication treatments. These are the key aggregate level results of our investigation.

Question 3: *What are the mechanisms through which communication affects behavior?*

Most subjects use the communication opportunity to coordinate on mutual cooperation in the first period of the game (Result 3a). This decreases strategic uncertainty and makes subjects choose cooperative strategies more frequently. However, subjects do not coordinate on complex efficient equilibria and fail to talk through all contingencies beforehand. Coordination on mutual cooperation in the first round is sufficient for a stable cooperative interaction under perfect monitoring, where mutual cooperation is never cast into doubt by a bad signal, as subjects observe the other player's action. This is not the case under imperfect monitoring, and as subjects fail to coordinate their responses to bad signals beforehand, they need repeated communication phases to deal with these situations, exchange information about the history of play (Result 3b), and thus reduce the additional uncertainty. As they are absent in the pre-play communication treatments, cooperation rates begin to decline as soon as bad signals arrive and raise doubts about the history of play. In our repeated-communication treatment with private monitoring, we further find that subjects frequently (and mostly truthfully) exchange information about their private signals, as suggested by recent theory.

To study strategy choices, we use and extend Dal Bó and Fréchette's (2011) strategy frequency estimation method (SFEM). Rather than having to rely on a predefined set of candidate strategies, our extension allows us to infer the strategies from the data. In addition to the strategies uncovered by this approach, we report results of the SFEM with a standard set of strategies plus additional (behavior) strategies, and of Camera et al.'s (2012) classification approach. With all three approaches, we find that subjects' play becomes substantially more lenient and forgiving with communication (Result 4). This effect is particularly strong with repeated communication under imperfect monitoring. We further develop an extension of SFEM based on latent-class regression, which allows us to study the evolution of strategy choices over time, and which can be used in future studies to analyze the correlation between strategy choices and other covariates (e.g., those studied in Proto et al., 2019). Our latent-class regressions show that subjects switch very quickly from less cooperative to more cooperative strategies in the first three supergames of the treatments with communication.

In addition to these results, which are all related to communication, our findings from the no-communication treatments complement Aoyagi et al.'s (2018) results on cooperation without communication under the three different monitoring structures. Their parametrization is characterized by lower levels of strategic uncertainty and our cooperation rates are, indeed, lower than theirs. The levels of strategic uncertainty are quite different between perfect and imperfect monitoring structures in their high noise treatment, as they are in our parametrization – higher with imperfect monitoring and similar between the two imperfect monitoring structures. Somewhat surprisingly, neither study finds a difference between the monitoring structures with respect to cooperation. However, within any monitoring structure cooperation is lower when strategic uncertainty is higher. This suggests that the level of strategic uncertainty, above which cooperation rates decline, is lower under perfect than under imperfect monitoring.

6 Conclusion

We set out to answer the central question how much communication is needed to sustain cooperation in noisy long-term interactions, such as cartels, teams or friendships. Our results give a comprehensive overview of how communication is used and affects cooperation and strategy choices. They demonstrate that communication can have an enormous impact on cooperation and its stability. The controlled laboratory environment allows us not only to cleanly identify and separate the effect of communication, but also to pinpoint the mechanisms through which communication affect cooperation. Our results show that communication fosters cooperation by reducing two types of uncertainty, strategic uncertainty and uncertainty about the history of play, and thereby reveal an important interplay between communication opportunities and the quality of monitoring. Most importantly, we find that repeated communication opportunities are crucial for sustaining cooperation under imperfect monitoring where uncertainty about the history of play becomes important. This finding is consistent with case study evidence on cartels and corroborates that cracking down on communication is a reasonable strategy for antitrust authorities. Without repeated communication opportunities, it becomes very difficult to sustain cooperation even in the relatively simple setting of our laboratory experiment.

We would finally like to point to some interesting avenues for future research. Communication affects choices and vice-versa. Ideally, we would thus like to estimate strategies that treat communication content as a choice and condition behavior not only on past actions and signals but also on past communication. To have a chance to recover such strategies from the data, one would have to strongly limit the message space, as do Arechar et al. (2017), who allow for communication only about intended actions. To gain more insights into the role of information exchange under private monitoring, it could be useful to limit communication to the reporting of private signals in future studies. However, while that would help to gain insights into this important role of communication, our results, and those from other recent studies of communication in repeated games, clearly suggest that thinking about communication as a mere exchange of information is insufficient. Kartal and Müller (2018) make a first step in broadening this narrow theoretical view of communication by modeling how communication reduces strategic uncertainty. Taking further steps in this direction, for example, by combining the two key roles of communication under imperfect monitoring – the reduction of strategic uncertainty and the reduction of uncertainty about the history of play – in one framework, promises to be a fruitful agenda for future research.

References

Agresti, A. (2003): "Logit Models for Multinomial Responses," in *Categorical Data Analysis*, Wiley-Blackwell, chap. 7, 267–313, 2 ed.

Andersson, O. and E. Wengström (2007): "Do Antitrust Laws Facilitate Collusion? Experimental Evidence on Costly Communication in Duopolies," *Scandinavian Journal of Economics*, 109, 321–339.

——— (2012): "Credible communication and cooperation: Experimental evidence from multi-stage Games," *Journal of Economic Behavior and Organization*, 81, 207–219.

Aoyagi, M., V. Bhaskar, and G. R. Fréchette (2018): "The Impact of Monitoring in Infinitely Repeated Games: Perfect, Public, and Private," *American Economic Journal: Microeconomics*, forthc.

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

Arechar, A. A., A. Dreber, D. Fudenberg, and D. G. Rand (2017): "I'm just a soul whose intentions are good: The role of communication in noisy repeated games," *Games and Economic Behavior*, 104, 726–743.

Awaya, Y. and V. Krishna (2016): "On Communication and Collusion," *American Economic Review*, 106, 285–315.

Backhaus, T. and Y. Breitmoser (2018): "God does not play dice, but do we? On the determinism of choice in long-run interactions," *Mimeo.*

Bandeen-Roche, K., D. L. Miglioretti, S. L. Zeger, and P. J. Rathouz (1997): "Latent variable regression for multiple discrete outcomes," *Journal of the American Statistical Association*, 92, 1375–1386.

Ben-Porath, E. and M. Kahneman (1996): "Communication in Repeated Games with Private Monitoring," *Journal of Economic Theory*, 70, 281–297.

Bhaskar, V., G. J. Mailath, and S. Morris (2008): "Purification in the infinitely-repeated prisoners' dilemma," *Review of Economic Dynamics*, 11, 515–528.

Bhaskar, V. and I. Obara (2002): "Belief-based equilibria in the repeated prisoners' dilemma with private monitoring," *Journal of Economic Theory*, 102, 40–69.

Biernacki, C., G. Celeux, and G. Govaert (2000): "Assessing a mixture model for clustering with the integrated completed likelihood," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22, 719–725.

Bigoni, M., J. Potters, and G. Spagnolo (2018): "Frequency of Interaction , Communication and Collusion: An Experiment," *Mimeo*.

Blonski, M., P. Ockenfels, and G. Spagnolo (2011): "Equilibrium Selection in the Repeated Prisoner's Dilemma: Axiomatic Approach and Experimental Evidence," *American Economic Journal: Microeconomics*, 3, 164–192.

Bolck, A., M. Croon, and J. Hagenaars (2004): "Estimating latent structure models with categorical variables: One-step versus three-step estimators," *Political Analysis*, 12, 3–27.

Breitmoser, Y. (2015): "Cooperation, but no reciprocity: Individual strategies in the repeated prisoner's dilemma," *American Economic Review*, 105, 2882–2910.

Bruttel, L. and U. Kamecke (2011): "Infinity in the lab. How do people play repeated games?" *Theory and Decision*, 72, 205–219.

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

——— (2013): "Experimental markets with frictions," *Journal of Economic Surveys*, 27, 536–553.

Cameron, a. C., J. B. Gelbach, and D. L. Miller (2011): "Robust Inference With Multiway Clustering," *Journal of Business & Economic Statistics*, 29, 238–249.

Compte, O. (1998): "Communication in Repeated Games with Imperfect Private Monitoring," *Econometrica*, 66, 587–626.

Compte, O. and A. Postlewaite (2015): "Plausible cooperation," *Games and Economic Behavior*, 91, 45–59.

Cooper, D. J. and K.-U. Kühn (2014a): "Communication, renegotiation, and the scope for collusion," *American Economic Journal: Microeconomics*, 6, 247–278.

——— (2014b): "Cruel to be Kind: An Experimental Study of Communication, Collusion, and Threats to Punish," *Mimeo.*

Cooper, R. W., D. V. DeJong, R. Forsythe, and T. W. Ross (1992): "Communication in Coordination Games," *Quarterly Journal of Economics*, 107, 739–771.

Crawford, V. and J. Sobel (1982): "Strategic information transmission," *Econometrica*, 50, 1431–1451.

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

Dal Bó, P. and G. R. Fréchette (2011): "The evolution of cooperation in infinitely repeated games: Experimental evidence," *American Economic Review*, 101, 411–429.

——— (2018a): "On the Determinants of Cooperation in Infinitely Repeated Games: A Survey," *Journal of Economic Literature*, 56, 60–114.

——— (2018b): "Strategy choice in the infinitely repeated prisoners' dilemma," *Mimeo.*

Dayton, M. C. and G. Macready (1988): "Concomitant-Variable Latent-Class Models," *Journal of the American Statistical Association*, 83, 173–178.

Dempster, A., N. Laird, and D. B. Rubin (1977): "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society Series B*, 39, 1–38.

Dvorak, F. (2020): "stratEst: An R Package for Strategy Estimation," *TWI Working Paper Series*, 119.

Ellingsen, T. and R. Östling (2010): "When Does Communication Improve Coordination?" *American Economic Review*, 100, 1695–1724.

Ely, J. C., J. Hörner, and W. Olszewski (2005): "Belief-Free Equilibria in Repeated Games," *Econometrica*, 73, 377–415.

Ely, J. C. and J. Välimäki (2002): "A Robust Folk Theorem for the Prisoner's Dilemma," *Journal of Economic Theory*, 102, 84–105.

Embrey, M., G. Fréchette, and E. Stacchetti (2013): "An Experimental Study of Imperfect Public Monitoring: Efficiency Versus Renegotiation-Proofness," *Mimeo.*

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

Fischbacher, U. (2007): "Z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental Economics*, 10, 171–178.

Fonseca, M. a. and H. T. Normann (2012): "Explicit vs. tacit collusion-The impact of communication in oligopoly experiments," *European Economic Review*, 56, 1759–1772.

Fréchette, G. R. and S. Yuksel (2017): "Infinitely repeated games in the laboratory: four perspectives on discounting and random termination," *Experimental Economics*, 20, 279–308.

Fudenberg, D., D. G. Rand, and A. Dreber (2012): "Slow to Anger and Fast to Forgive: Cooperation in an Uncertain World," *American Economic Review*, 102, 720–749.

Genesove, D. and W. P. Mullin (2001): "Rules, communication, and collusion: Narrative evidence from the sugar institute case," *American Economic Review*, 91, 379–398.

Green, E. J. and R. H. Porter (1984): "Noncooperative Collusion under Imperfect Price Information," *Econometrica*, 52, 87.

Greiner, B. (2015): "Subject pool recruitment procedures: organizing experiments with ORSEE," *Journal of the Economic Science Association*, 1, 114–125.

Harrington, J. E. and A. Skrzypacz (2011): "Private Monitoring and Communication in Cartels: Explaining Recent Collusive Practices," *American Economic Review*, 101, 2425–2449.

Harsanyi, J. C. and R. Selten (1988): *A General Theory of Equilibrium Selection in Games*, Cambridge, MA: MIT Press.

Heller, Y. (2017): "Instability of belief-free equilibria," *Journal of Economic Theory*, 168, 261–286.

Kalai, E., D. Samet, and W. Stanford (1988): "A note on reactive equilibria in the discounted prisoner's dilemma and associated games," *International Journal of Game Theory*, 17, 177–186.

Kamei, K. and A. Nesterov (2020): "Endogenous Monitoring Through Gossiping in an Infinitely Repeated Prisoner's Dilemma Game: Experimental Evidence," *Working Paper*.

Kandori, M. and H. Matsushima (1998): "Private Observation, Communication and Collusion," *Econometrica*, 66, 627–652.

Kartal, M. and W. Müller (2018): "A New Approach to the Analysis of Cooperation Under the Shadow of the Future: Theory and Experimental Evidence," *SSRN Working Paper*.

Levenstein, M. C. and V. Y. Suslow (2006): "What Determines Cartel Success?" *Journal of Economic Literature*, 44, 43–95.

Mailath, G. J. and L. Samuelson (2006): *Repeated Games and Reputations*, Oxford University Press: Oxford.

Matsushima, H. (1991): "On the theory of repeated games with private information. Part II: revelation through communication," *Economics Letters*, 35, 257–261.

Obara, I. (2009): "Folk Theorem with Communication," *Journal of Economic Theory*, 144, 120–134.

Pearce, D. G. (1987): "Renegotiation-Proof Equilibria: Collective Rationality and Intertemporal Cooperation," *Cowles Foundation Discussion Paper*.

Piccione, M. (2002): "The Repeated Prisoner's Dilemma with Imperfect Private Monitoring," *Journal of Economic Theory*, 102, 70–83.

Proto, E., A. Rustichini, and A. Sofianos (2019): "Intelligence, Personality, and Gains from Cooperation in Repeated Interactions," *Journal of Political Economy*, 127, 1351–1390.

Rabin, M. (1994): "A Model of Pre-game Communication," *Journal of Economic Theory*, 63, 370–391.

Romero, J. and Y. Rosokha (2018): "Constructing strategies in the indefinitely repeated prisoner's dilemma game," *European Economic Review*, 104, 185–219.

Schwarz, G. (1978): "Estimating the Dimension of a Model," *The Annals of Statistics*, 6, 461–464.

Sekiguchi, T. (1997): "Efficiency in repeated prisoner's dilemma with private monitoring," *Journal of Economic Theory*, 76, 345–361.

Stigler, G. J. (1964): "A Theory of Oligopoly," *Journal of Political Economy*, 72, 44–61.

van Damme, E. (1989): "Renegotiation-proof equilibria in repeated prisoners' dilemma," *Journal of Economic Theory*, 47, 206–217.

Vespa, E. and A. J. Wilson (2018): "Information Transmission under the shadow of the future: An experiment," *Mimeo.*

[FOR ONLINE PUBLICATION]

Appendix A Theoretical Appendix

In A.1, we present two extensions of the BAD to the imperfect monitoring structures. In A.2, we derive existence conditions for equilibria in memory-one belief-free strategies in general, and for the subset of semi-grim memory-one belief-free equilibria. The latter give us the SG-thresholds. Further, we provide a characterization of these equilibria. In A.3, we construct renegotiation-proof equilibria for perfect and imperfect public monitoring and a truthful communication equilibrium for the case of imperfect private monitoring. It will be useful to recall the normalized stage game parameters:

	C	D
C	1,1	$-l$,1+g
D	1+g,$-l$	0,0

A.1 BAD under imperfect monitoring

Extending the BAD to imperfect monitoring requires to adapt the GRIM strategy to the imperfect monitoring structures. To derive lower bounds of the BAD, we use the adaptation of GRIM which is most robust to strategic uncertainty. This adaptation prescribes that players play D if they already played D in the previous round or when the last signal was not cc (c) under public (private) monitoring.

A.1.1 Public Monitoring

With public monitoring, indifference between GRIM and ALLD requires

$$\pi\frac{1}{1-\delta(1-\epsilon)^2}-(1-\pi)\frac{l}{1-\delta\epsilon(1-\epsilon)}=\pi\frac{(1+g)}{1-\delta\epsilon(1-\epsilon)}.$$

Hence, the BAD is

$$\pi^{DF}=\frac{l}{l-g+\frac{\delta((1-\epsilon)^2-\epsilon(1-\epsilon))}{1-\delta(1-\epsilon)^2}}. \qquad (2)$$

If $g = l$, the lower bound is

$$\pi^{DF} = \frac{1-\delta(1-\epsilon)^2}{\delta((1-\epsilon)^2 - \epsilon(1-\epsilon))} l,$$

and $\partial\pi^{DF}/\partial\epsilon = l\delta(3-4\epsilon-\delta(1-\epsilon)^2)/(\delta(1-2\epsilon)^2(\epsilon-1)^2) > 0$ for $\delta < 1$ and $\epsilon \leq 0.5$.
If $1+g = l$, the lower bound is

$$\pi^{DF} = \frac{1-\delta(1-\epsilon)^2}{1-\delta\epsilon(1-\epsilon)} l,$$

and $\partial\pi^{DF}/\partial\epsilon = l\delta(3-4\epsilon-\delta(1-\epsilon)^2)/(1-\delta\epsilon(1-\epsilon)^2 > 0$ for $\delta < 1$ and $\epsilon \leq 0.5$. Note that for $\epsilon = 0$ the equations above yield the BAD of perfect monitoring defined in Equation 1.

A.1.1 Private Monitoring

With private monitoring, indifference requires

$$\pi \frac{1+\delta\epsilon(1-\epsilon)(1+g-l)/(1-\delta\epsilon)}{1-\delta(1-\epsilon)^2} - (1-\pi)\frac{l}{1-\delta\epsilon} = \pi\frac{(1+g)}{1-\delta\epsilon},$$

and the BAD is given by

$$\pi^{DF} = \frac{l}{l-g+\frac{\delta((1-2\epsilon)-\epsilon(1-\epsilon)(l-g))}{1-\delta(1-\epsilon)^2}}. \tag{3}$$

If $g = l$, the lower bound is

$$\pi^{DF} = \frac{1-\delta(1-\epsilon)^2}{\delta(1-2\epsilon)} l,$$

and $\partial\pi^{DF}/\partial\epsilon = 2l(1-\delta(\epsilon(1-\epsilon))/(\delta(1-2\epsilon)^2) > 0$ for $\delta < 1$ and $\epsilon \leq 0.5$.
If $1+g = l$, the lower bound is

$$\pi^{DF} = \frac{1-\delta(1-\epsilon)^2}{1-\delta\epsilon(1-\epsilon)} l,$$

and $\partial\pi^{DF}/\partial\epsilon = l\delta(3-2\epsilon-\delta(1-\epsilon)^2)/(1-\delta\epsilon)^2 > 0$ for $\delta < 1$ and $\epsilon \leq 0.5$. For $\epsilon = 0$, the equations above yield the BAD of perfect monitoring defined in Equation 1. Note that under private monitoring (GRIM, GRIM) is not an equilibrium in pure strategies but π^{DF} equals the mixing probability in Sekiguchi's (1997) construction of a belief-based equilibrium.

A.2 Belief-Free Equilibria

Depending on the monitoring structure, different versions of memory-one belief-free strategies exist. We consider three cases: (1) M1BF strategies which condition on (a_i, a_{-i}), (2) M1BF strategies which condition on (ω_i, ω_{-i}), and (3) M1BF strategies which condition on (a_i, ω_{-i}). Under perfect monitoring, all three cases are possible. Under public monitoring, only cases 2 and 3 are possible while case 3 is the only possible case under private monitoring. The existence conditions of semi-grim strategies which condition on public signals and action-signal combinations are defined in Propositions 1.1.2, 1.2.2 and 1.3.2.

A.2.1 Actions (Perfect Monitoring)

Proposition 2.1.1 [Memory-One Belief-Free Equilibria Conditioning on Actions]

(i) *If strategies condition on actions, the existence condition for symmetric memory-one belief-free equilibria depends on the larger of the two values g and l. Let ϕ denote the larger of the two values. The existence condition is:*

$$\delta \geq \delta_{aa}^{BF} = \frac{\phi}{1+\phi} \tag{4}$$

(ii) *Above the threshold, a two-dimensional manifold of memory-one belief-free equilibria exists given by*

$$\sigma_{cd} = \sigma_{cc} + \left(\sigma_{cc} - \sigma_{dd} - \frac{1}{\delta}\right) g \tag{5}$$

and

$$\sigma_{dc} = \sigma_{dd} - \left(\sigma_{cc} - \sigma_{dd} - \frac{1}{\delta}\right) l \tag{6}$$

(iii) *For $\delta = \delta_{aa}^{BF}$ all memory-one belief-free equilibrium strategies have the same cooperation probabilities after nonempty memory-one histories and are $\sigma = (\sigma_\emptyset, 1, (1 - g/l), 1, 0)$ if $l > g$, $\sigma = (\sigma_\emptyset, 1, 0, (l/g), 0)$ if $g > l$ and $\sigma = (\sigma_\emptyset, 1, 0, 1, 0)$ if $g = l$. We call this the threshold memory-one belief-free equilibrium T1BF.*

Since g and l are both positive values these equilibria exist for high enough values of δ. Note that if $g \geq l$ the δ threshold corresponds to the one for cooperative subgame-perfect equilibria of the repeated game with perfect monitoring. However, if $l > g$ as in our case, the conditions differ with $\delta_{aa}^{BF} > \delta^{SPE}$. The condition applies for belief-free equilibria in reactive strategies

(Kalai et al., 1988) which condition on the other player's action and require $g = l$ which yields $\delta_{aa}^{BF} = \delta^{SPE}$.

Proof of Proposition 1.1.1. Let $V_{a_j a_i}^{a_i}$ denote player i's expected payoff for playing a_i if player j observed the action profile $\{a_j, a_i\}$ in the previous round (we say player j is in state $a_j a_i$). If $\sigma_{a_i a_j}$ denotes the probability to play c for any player i after $\{a_i, a_j\}$, we have:

$$V_{aa}^{c} = (1-\delta)(\sigma_{aa} - (1-\sigma_{aa})l) + \delta(\sigma_{aa}V_{cc} + (1-\sigma_{aa})V_{dc}) \tag{7}$$

$$V_{aa}^{d} = (1-\delta)(\sigma_{aa}(1+g) + (1-\sigma_{aa})0) + \delta(\sigma_{aa}V_{cd} + (1-\sigma_{aa})V_{dd}) \tag{8}$$

Following Bhaskar et al. (2008), we derive conditions for V_{cd} and V_{cc} which assure the strategies are belief-free, that is, for any $\sigma_{aa} \in (0,1)$, player i is indifferent between playing c or d independent of player j's state. Subtracting (8) from (7) gives:

$$0 = \sigma_{aa}\left\{(1-\delta)(l-g) + \delta\left(V_{cc} - V_{cd} - V_{dc} + V_{dd}\right)\right\} - (1-\delta)l + \delta\left(V_{dc} - V_{dd}\right)$$

The equation holds independent of σ_{aa} if the terms in curly brackets and the last part are both zero. Solving the the condition resulting from the last part for $V_{dc} - V_{dd}$ and inserting the solution into the condition derived from the terms in curly brackets gives

$$V_{cc} = V_{cd} + \frac{(1-\delta)g}{\delta}$$

and

$$V_{dc} = V_{dd} + \frac{(1-\delta)l}{\delta}$$

Solving (7) for σ_{cc} using the condition on V_{dc} above and rearranging for V_{cc} yields

$$V_{cc} = \frac{(1-\delta)\sigma_{cc} + \delta(1-\sigma_{cc})V_{dd}}{1-\delta\sigma_{cc}}$$

Solving (7) for σ_{dd} using the condition on V_{cd} and V_{cc} above gives

$$V_{dd} = \frac{\sigma_{dd}}{1 + \delta\sigma_{dd} - \delta\sigma_{cc}}$$

Now, all V_{aa} can be eliminated from (7) solved for σ_{dd} and σ_{dc} this yields (5) and (6) which proofs (ii). Note that $\partial\sigma_{cd}/\partial\delta > 0$, $\partial\sigma_{cd}/\partial\sigma_{cc} > 0$ and $\partial\sigma_{cd}/\partial\sigma_{dd} < 0$. The question is, how big δ must be at least in order to assure that $\sigma_{cd} \geq 0$ if $\sigma_{cc} = 1$ and $\sigma_{dd} = 0$. Inserting these values into (5) and rearranging gives $\delta > \delta_{aa}^{BF}$ with $\phi = g$. Note that $\sigma_{cd} \leq 1$ is true even if $\sigma_{cc} = 1$ and $\sigma_{dd} = 0$ for all feasible values of δ, g and l. At the same time $\partial\sigma_{dc}/\partial\delta < 0$,

$\partial\sigma_{dc}/\partial\sigma_{cc} < 0$ and $\partial\sigma_{dc}/\partial\sigma_{dd} > 0$. The question here is, how big δ must be at least in order to assure that $\sigma_{dc} \leq 1$ if $\sigma_{cc} = 1$ and $\sigma_{dd} = 0$. Inserting these values into (6) and rearranging gives $\delta > \delta^{BF}_{aa}$ with $\phi = l$. At the same time, $\sigma_{dc} \geq 0$ true even if $\sigma_{cc} = 1$ and $\sigma_{dd} = 0$ for all feasible values of δ, g and l. Hence, the larger of the values g and l imposes the stricter condition on δ which proofs (i). To complete the proof, insert (4) together with $\sigma_{cc} = 1$ and $\sigma_{dd} = 0$ into (5) and (6) to obtain the structure of the T1BF response defined by g and l. $\square$

Next, we derive the δ threshold, above which semi-GRIM equilibria exist. See Breitmoser (2015) for an alternative derivation.

Proposition 1.1.2 [Semi-Grim M1BF Equilibria Conditioning on Actions]

(i) *If strategies condition on actions, the existence condition for symmetric semi-grim memory-one belief-free equilibria is:*

$$\delta \geq \delta^{SG}_{aa} = \frac{g+l}{1+g+l} \tag{9}$$

(ii) *Above the threshold a continuum* $\sigma_{cc} \in (\frac{g+l}{\delta(1+g+l)}, 1)$ *of memory one belief-free equilibria in semi-grim strategies exists, given by:*

$$\sigma_{dd} = \sigma_{cc} - \frac{g+l}{\delta(1+g+l)} \tag{10}$$

and

$$\sigma_{cd} = \sigma_{dc} = \sigma_{cc} - \frac{g}{\delta(1+g+l)} \tag{11}$$

(iii) *For* $\delta = \delta^{SG}_{aa}$ *all semi-grim memory-one belief-free equilibrium strategies have the same cooperation probabilities after nonempty memory-one histories and are* $\sigma = (\sigma_{\emptyset}, 1, 1 - g/(g+l), , 0)$. *If* $l = g$, *then* $\sigma = (\sigma_{\emptyset}, 1, 0.5, 0.5, 0)$.

Proof of Proposition 1.1.2. Using (5) and (6) yields (10) and (11). Note that $\sigma_{dd} < \sigma_{cd} < 1$ for $\sigma_{cc} \in (0, 1)$ and any $\delta \in (0, 1)$. For existence σ_{dd} must be positive. Rearranging yields the SG-threshold. Note that the condition on δ is always stricter than the condition on δ, which results from $\sigma_{cd} = \sigma_{dc} \geq 0$, and is $\delta \geq g/(1+g+l)$. $\square$

Note that the condition for semi grim equilibria is a mixture of the two possible conditions based on the different values of ϕ with equal weight on g and l as required by axiom 5 in Blonski et al. (2011) while (4) gives full weight on the larger of the two values.

A.2.2 Public Signals (Perfect and Public Monitoring)

Proposition 2.2.1 [M1BF Equilibria Conditioning on Public Signals]

(i) *If strategies condition on the ϵ-noisy public signals, the existence condition for symmetric memory-one belief-free equilibria depends on the larger of the two values g and l. Let ϕ denote the larger and ψ the smaller of the two values. The existence condition is:*

$$\delta \geq \delta_{ss}^{BF} = \frac{(1-\epsilon)\phi - \epsilon\psi}{(1-2\epsilon)(1-2\epsilon+(1-\epsilon)\phi-\epsilon\psi)} \tag{12}$$

(ii) *Above the threshold, a two-dimensional manifold of memory-one belief-free equilibria exists given by*

$$\sigma_{cd} = \sigma_{cc} + \frac{\sigma_{cc} - \sigma_{dd} - \frac{1}{\delta(1-2\epsilon)}}{1-2\epsilon}((1-\epsilon)g - \epsilon l) \tag{13}$$

and

$$\sigma_{dc} = \sigma_{dd} - \frac{\sigma_{cc} - \sigma_{dd} - \frac{1}{\delta(1-2\epsilon)}}{1-2\epsilon}((1-\epsilon)l - \epsilon g) \tag{14}$$

(iii) *For $\delta = \delta_{ss}^{BF}$ all memory-one belief-free equilibrium strategies have the same cooperation probabilities after nonempty memory-one histories and are $\sigma = (\sigma_\emptyset, 1, (1-g/l), 1, 0)$ if $l > g$, $\sigma = (\sigma_\emptyset, 1, 0, (l/g), 0)$ if $g > l$ and $\sigma = (\sigma_\emptyset, 1, 0, 1, 0)$ if $g = l$. We call this the threshold memory-one belief-free equilibrium T1BF.*

In contrast to result for actions, combinations of the parameters g, l and ϵ exists for which $\delta_{ss}^{BF} > 1$.

Proof of Proposition 2.2.1. The proof follows the same steps as for actions. Let $V_{s_j s_i}^{a_i}$ denote player i's expected payoff for playing a_i if player j observed $\{s_j, s_i\}$ in the previous round (which means player j is in state $s_j s_i$). If $\sigma_{s_i s_j}$ denotes the (universal) probability of player i

to play c after $\{s_i, s_j\}$, we get:

$$\begin{aligned} V_{ss}^c = (1-\delta)(\sigma_{ss} - (1-\sigma_{ss})l) + \delta[(1-\epsilon)(\sigma_{ss}(1-\epsilon) + (1-\sigma_{ss})\epsilon)V_{cc} \\ +\epsilon(\sigma_{ss}(1-\epsilon) + (1-\sigma_{ss})\epsilon)V_{cd} \\ +(1-\epsilon)(\sigma_{ss}\epsilon + (1-\sigma_{ss})(1-\epsilon))V_{dc} \\ +\epsilon(\sigma_{ss}\epsilon + (1-\sigma_{ss})(1-\epsilon))V_{dd}] \qquad (15) \\ V_{ss}^d = (1-\delta)(\sigma_{ss}(1+g) + (1-\sigma_{ss})0) + \delta[\epsilon(\sigma_{ss}(1-\epsilon) + (1-\sigma_{ss})\epsilon)V_{cc} \\ +(1-\epsilon)(\sigma_{ss}(1-\epsilon) + (1-\sigma_{ss})\epsilon)V_{cd} \\ +\epsilon(\sigma_{ss}\epsilon + (1-\sigma_{ss})(1-\epsilon))V_{dc} \\ +(1-\epsilon)(\sigma_{ss}\epsilon + (1-\sigma_{ss})(1-\epsilon))V_{dd}] \qquad (16) \end{aligned}$$

Again we derive conditions for V_{cd} and V_{cc} which together assure the belief-free property following Following Bhaskar et al. (2008), that is, for any $\sigma_{ss} \in (0,1)$, player i is indifferent between playing c or d independent of player j's state. First, subtracting (16) from (15) gives:

$$\begin{aligned} 0 = \sigma_{ss}\left\{(1-\delta)(l-g) + \delta\left((1-2\epsilon)^2 V_{cc} - (1-2\epsilon)^2 V_{cd} - (1-2\epsilon)^2 V_{dc} + (1-2\epsilon)^2 V_{dd}\right)\right\} \\ -(1-\delta)l + \delta\left((1-2\epsilon)\epsilon V_{cc} - (1-2\epsilon)\epsilon V_{cd} + (1-2\epsilon)(1-\epsilon)V_{dc} - (1-2\epsilon)(1-\epsilon)V_{dd}\right) \end{aligned}$$

Note that he expression holds independent of σ_{ss} if the terms in curly brackets and the terms in the second line are both zero. Solving the the condition on the second line for $V_{dc} - V_{dd}$ and inserting into the other condition gives

$$V_{cc} = V_{cd} + \frac{(1-\delta)((1-\epsilon)g - \epsilon l)}{\delta(1-2\epsilon)^2}$$

and

$$V_{dc} = V_{dd} + \frac{(1-\delta)((1-\epsilon)l - \epsilon g)}{\delta(1-2\epsilon)^2}$$

Solving (15) for σ_{cc} and rearranging for V_{cc} yields

$$V_{cc} = \frac{(1-\delta)(\sigma_{cc} - l) + \delta(1-\epsilon-\sigma_{cc}(1-2\epsilon))V_{dd} + \frac{(1-\delta)(1-\epsilon)((1-\epsilon)l-\epsilon g)}{(1-2\epsilon)^2} - \frac{(1-\delta)\epsilon l}{1-2\epsilon}}{1-\delta(\sigma_{cc}(1-2\epsilon)+\epsilon)}.$$

Solving (15) for σ_{dd} and inserting V_{cc} yields an expression for V_{dd} (omitted here) that does not depend on any other V_{ss}. Now, all V_{ss} can be eliminated from (15) and we can solve for σ_{cd} and σ_{dc} which leads to (ii). For existence we need to assure that $\sigma_{cd} \in (0,1)$ and $\sigma_{dc} \in (0,1)$

for a feasible combination of values σ_{cc}, σ_{dd} and δ. First assume $(1-\epsilon)\psi - \epsilon\phi > 0$ and consider σ_{cd} (note that $(1-\epsilon)\phi - \epsilon\psi > 0$ always holds for $\epsilon < 0.5$). In this case $\partial\sigma_{cd}/\partial\sigma_{cc} > 0$ and $\partial\sigma_{cd}/\partial\sigma_{dd} < 0$. Note that $\sigma_{cd} \leq 1$ for any $\delta \in (0,1)$ even if $\sigma_{cc} = 1$ and $\sigma_{dd} = 0$. To establish $\sigma_{cd} \geq 0$ we use $\sigma_{cc} = 1$ and $\sigma_{dd} = 0$. Solving for δ shows gives the condition $\delta > \delta_{ss}^{BF}$ with $\phi = g$. Next, we consider σ_{dc} still assuming $(1-\epsilon)\psi - \epsilon\phi > 0$. Hence $\partial\sigma_{dc}/\partial\sigma_{cc} < 0$ and $\partial\sigma_{dc}/\partial\sigma_{dd} > 0$. Again $\sigma_{dc} \geq 0$ for any $\delta \in (0,1)$ even if $\sigma_{cc} = 1$ and $\sigma_{dd} = 0$. To establish $\sigma_{dc} \leq 1$ we use $\sigma_{cc} = 1$ and $\sigma_{dd} = 0$ which gives $\delta > \delta_{ss}^{BF}$ with $\phi = l$. Therefore, if $(1-\epsilon)\psi - \epsilon\phi > 0$ the stricter condition on δ results from the larger of the two values g or l as in (12). Note that $(1-\epsilon)\psi - \epsilon\phi < 0$ also requires $\delta > \delta_{ss}^{BF}$ to make the probabilities interior. On the other hand, it implies $\phi > \frac{1-\epsilon}{\epsilon}\psi$ and $\delta_{ss}^{BF} > 1$. To see this we can rearrange $\delta_{ss}^{BF} < 1$ to $\phi < \frac{(1-2\epsilon)^2+2\epsilon^2\psi}{2\epsilon-2\epsilon^2}$ and show that this contradicts $\phi > \frac{1-\epsilon}{\epsilon}\psi$ for $\epsilon \in (0, 0.5)$. This proofs (i). To complete the proof, insert (12) together with $\sigma_{cc} = 1$ and $\sigma_{dd} = 0$ into (13) and (14) to obtain the structure of the T1BF response defined by g and l. □

Proposition 2.2.2 [Semi-Grim M1BF Equilibria Conditioning on Public Signals]

(i) *If players condition on the ϵ-noisy public signals, the existence condition for semi-GRIM equilibria is:*

$$\delta \geq \delta_{ss}^{SG} = \frac{g+l}{(1-2\epsilon)(1+g+l)} \tag{17}$$

(ii) *Above this threshold, a continuum $\sigma_{cc} \in (\frac{g+l}{\delta(1-2\epsilon)(1+g+l)}, 1)$ of semi-grim equilibria exists given by:*

$$\sigma_{dd} = \sigma_{cc} - \frac{g+l}{\delta(1-2\epsilon)(1+g+l)} \tag{18}$$

and

$$\sigma_{cd} = \sigma_{dc} = \sigma_{cc} - \frac{g}{\delta(1-2\epsilon)(1+g+l)} \tag{19}$$

(iii) *For $\delta = \delta_{ss}^{SG}$ all semi-grim memory-one belief-free equilibrium strategies have the same cooperation probabilities after nonempty memory-one histories and are $\sigma = (\sigma_\emptyset, 1, 1-g/(g+l), 1-g/(g+l), 0)$. If $l = g$, then $\sigma = (\sigma_\emptyset, 1, 0.5, 0.5, 0)$.*

Proof of Proposition 2.2.2. Using the semi-grim property $\sigma_{cd} = \sigma_{dc}$ for (13) and (14) yields (18) and (19). Observe that $\sigma_{dd} < \sigma_{cd} < 1$ for $\sigma_{cc} \in (0,1)$ and for existence σ_{dd} must be positive which can be rearranged to yield (17). □

A.2.3 Action-Signal Combinations (All Monitoring Structures)

Proposition 2.3.1 [M1BF Equilibria Conditioning on Action-Signal Combinations]

(i) *If players condition on their own action and the ϵ-noisy signal of the other player's action, the existence condition for symmetric memory-one belief-free equilibria also depends on the larger of the two values g and l. Let ϕ denote the larger of the two values and ψ the smaller of the two. The existence condition is:*

$$\delta \geq \delta_{as}^{BF} = \frac{\phi}{1 - 2\epsilon + (1-\epsilon)\phi - \epsilon\psi} \tag{20}$$

If $g = l$ the condition is the same as for private signals.

(ii) *Above the threshold, a two-dimensional manifold of memory-one belief-free equilibria exists given by*

$$\sigma_{cd} = \sigma_{cc} + \frac{\sigma_{cc} - \sigma_{dd} - \frac{1}{\delta}}{1 - 2\epsilon - \epsilon(g+l)} g \tag{21}$$

and

$$\sigma_{dc} = \sigma_{dd} - \frac{\sigma_{cc} - \sigma_{dd} - \frac{1}{\delta}}{1 - 2\epsilon - \epsilon(g+l)} l \tag{22}$$

(iii) *For $\delta = \delta_{as}^{BF}$ all memory-one belief-free equilibrium strategies have the same cooperation probabilities after nonempty memory-one histories and are $\sigma = (\sigma_{\emptyset}, 1, (1-g/l), 1, 0)$ if $l > g$, $\sigma = (\sigma_{\emptyset}, 1, 0, (l/g), 0)$ if $g > l$ and $\sigma = (\sigma_{\emptyset}, 1, 0, 1, 0)$ if $g = l$. We call this the threshold memory-one belief-free equilibrium T1BF.*

Proof of Proposition 2.3.1. Again the proof follows the same steps as for actions. Let $V_{a_j s_i}^{a_i}$ denote player i's expected payoff for playing a_i if player j played a_j and observed s_i in the previous round (which means player j is in state $a_j s_i$). If $\sigma_{a_i s_j}$ denotes the (universal) probability of player i to play c after $\{a_i, s_j\}$, we get:

$$\begin{aligned} V_{as}^{c} &= (1-\delta)(\sigma_{as} - (1-\sigma_{as})l) + \\ &\delta\left((1-\epsilon)\sigma_{as}V_{cc} + \epsilon\sigma_{as}V_{cd} + (1-\epsilon)(1-\sigma_{as})V_{dc} + \epsilon(1-\sigma_{as})V_{dd}\right) \end{aligned} \tag{23}$$

$$\begin{aligned} V_{as}^{d} &= (1-\delta)\sigma_{as}(1+g) + \\ &\delta\left((1-\epsilon)\sigma_{as}V_{cc} + \epsilon\sigma_{as}V_{cd} + (1-\epsilon)(1-\sigma_{as})V_{dc} + \epsilon(1-\sigma_{as})V_{dd}\right) \end{aligned} \tag{24}$$

Subtracting (24) from (23) gives:

$$0 = \sigma_{as} \left\{(1-\delta)(l-g) + \delta\left((1-2\epsilon)V_{cc} - (1-2\epsilon)V_{cd} - (1-2\epsilon)V_{dc} + (1-2\epsilon)V_{dd}\right)\right\} \\ -(1-\delta)l + \delta\left((1-2\epsilon)V_{dc} - (1-2\epsilon)V_{dd}\right)$$

The conditions on V_{cd} and V_{cc} based on the belief-free property are now:

$$V_{dc} = V_{dd} + \frac{(1-\delta)l}{\delta(1-2\epsilon)}$$

$$V_{cc} = V_{cd} + \frac{(1-\delta)g}{\delta(1-2\epsilon)}$$

Solving (23) for σ_{cc} and rearranging for V_{cc} yields

$$V_{cc} = \frac{(1-\delta)(\sigma_{cc} - (1-\sigma_{cc})l) + \delta(1-\sigma_{cc})V_{dd} - \delta\sigma_{cc}\frac{(1-\delta)((1-\epsilon)l+\epsilon g)}{\delta(1-2\epsilon)} + \delta(1-\epsilon)\frac{(1-\delta)l}{\delta(1-2\epsilon)}}{1-\delta\sigma_{cc}}$$

Solving (23) for σ_{dd} and inserting the solution for V_{cc} gives

$$V_{dd} = \frac{\sigma_{dd}\left(1 - \frac{(1-\delta)\epsilon l + \epsilon g}{1-2\epsilon}\right) + (1-\delta\sigma_{cc})\frac{\epsilon l}{1-2\epsilon}}{1+\delta\sigma_{dd} - \delta\sigma_{cc}}$$

Next, all V_{as} can be eliminated from (23) solved for σ_{dd} and σ_{dc} proofs (ii). For existence we need to assure that $\sigma_{cd} \in (0,1)$ and $\sigma_{dc} \in (0,1)$ for a feasible combination of values σ_{cc}, σ_{dd} and δ. First assume $1-2\epsilon-\epsilon(g+l) > 0$ and consider σ_{cd}. In this case $\partial\sigma_{cd}/\partial\sigma_{cc} > 0$ and $\partial\sigma_{cd}/\partial\sigma_{dd} < 0$. Note that $\sigma_{cd} \leq 1$ for any $\delta \in (0,1)$ even if $\sigma_{cc} = 1$ and $\sigma_{dd} = 0$. To establish $\sigma_{cd} \geq 0$ we use $\sigma_{cc} = 1$ and $\sigma_{dd} = 0$. Solving for δ shows gives the condition $\delta > \delta_{as}^{BF}$ with $\phi = g$. Next, we consider σ_{dc} still assuming $1-2\epsilon-\epsilon(g+l) > 0$. Hence $\partial\sigma_{dc}/\partial\sigma_{cc} < 0$ and $\partial\sigma_{dc}/\partial\sigma_{dd} > 0$. Again $\sigma_{dc} \geq 0$ for any $\delta \in (0,1)$ even if $\sigma_{cc} = 1$ and $\sigma_{dd} = 0$. To establish $\sigma_{dc} \leq 1$ we use $\sigma_{cc} = 1$ and $\sigma_{dd} = 0$ which gives $\delta > \delta_{as}^{BF}$ with $\phi = l$. Therefore, if $1-2\epsilon-\epsilon(g+l) > 0$ the stricter condition on δ results from the larger of the two values g or l as in (20).

If $1-2\epsilon-\epsilon(g+l) < 0$, $\partial\sigma_{cd}/\partial\sigma_{cc} < 0$ and $\partial\sigma_{cd}/\partial\sigma_{dd} > 0$. Using $\sigma_{cc} = 1$ and $\sigma_{dd} = 0$ we establish that $\sigma_{cd} \leq 1$ only if $\delta \geq 1$ (and the same can be shown for $\sigma_{dc} \geq 0$ when using $\sigma_{cc} = 0$ and $\sigma_{dd} = 1$). Note that (20) also requires $\delta \geq 1$ in this case. For the last case $1-2\epsilon-\epsilon(g+l) = 0$, σ_{cd} and σ_{dc} are not defined and (20) also requires $\delta \geq 1$. This proofs (i). To complete the proof, insert (20) together with $\sigma_{cc} = 1$ and $\sigma_{dd} = 0$ into (21) and (22) to obtain the structure of the T1BF response defined by g and l. □

Proposition 2.3.2 [Semi-Grim M1BF Equilibria Conditioning on Action-Signal Combinations]

(i) *If players condition on their own action and the ϵ-noisy signal of the other player's action, the existence condition for symmetric memory one belief-free equilibria in semi grim strategies is:*

$$\delta \geq \delta_{as}^{SG} = \frac{g+l}{1-2\epsilon+(1-\epsilon)(g+l)} \tag{25}$$

(ii) *Above this threshold, a continuum $\sigma_{cc} \in (\frac{g+l}{\delta(1-2\epsilon+(1-\epsilon)(g+l))}, 1)$ of semi-grim equilibria exists given by:*

$$\sigma_{dd} = \sigma_{cc} - \frac{g+l}{\delta(1-2\epsilon+(1-\epsilon)(g+l))} \tag{26}$$

and

$$\sigma_{cd} = \sigma_{dc} = \sigma_{cc} - \frac{g}{\delta(1-2\epsilon+(1-\epsilon)(g+l))} \tag{27}$$

(iii) *For $\delta = \delta_{as}^{SG}$ all semi-grim memory-one belief-free equilibrium strategies have the same cooperation probabilities after nonempty memory-one histories and are $\sigma = (\sigma_\emptyset, 1, 1-g/(g+l), 1-g/(g+l), 0)$. If $l = g$, then $\sigma = (\sigma_\emptyset, 1, 0.5, 0.5, 0)$.*

Proof of Proposition 2.3.2. Using the semi-grim property $\sigma_{cd} = \sigma_{dc}$ for (21) and (22) yields (26) and (27). Observe that $\sigma_{dd} < \sigma_{cd} < 1$ for $\sigma_{cc} \in (0, 1)$ and for existence σ_{dd} must be positive which can be rearranged to yield (25). □

A.3 Renegotiation-Proof and Truthful Communication Equilibria

We give examples for the construction of renegotiation-proof equilibria for the perfect and imperfect monitoring cases and for a truthful communication equilibrium under imperfect private monitoring. These equilibria can be described by two states each: (1) a reward stage, in which both players cooperate, and (2) a punishment stage; and transition rules between the states. Unlike in equilibria in strongly symmetric strategies, the punisher and the punished player have to play differently in the punishment stage to assure that this state is not Pareto-dominated by the reward state. Hence, the continuation values of the two players will be different once we enter the punishment state. We will use the following notation: V_r

for the continuation value of the reward state, and V_{pp} (V_{pd}) for the continuation value of the punisher (the punished player) in the punishment state. The following condition has to hold in any renegotiation-proof equilibrium:

$$V_{pp} \geq V_r \tag{28}$$

The following condition has to hold in any truthful communication equilibrium, where the revelation constraints require that the punisher must be indifferent between staying in the reward state or entering the punishment state as punisher:

$$V_{pp} = V_r \tag{29}$$

A.3.1 Perfect Monitoring

The most simple candidate equilibrium is the following. It starts in the reward state with both players cooperating. In case of a defection, they enter the punishment state, in which the player who defected plays C while the other player plays D for one period. After this period, the game returns to the reward state. For this to be a renegotiation-proof equilibrium, the following three conditions have to be fulfilled:

1. No player has an incentive to deviate in the reward stage:

$$1 \geq (1-\delta)(1+g) - \delta(1-\delta)l + \delta^2$$

2. In the punishment stage, the player being punished has no incentive to deviate:

$$-(1-\delta)l + \delta \geq -\delta(1-\delta)l + \delta^2$$

3. The punisher wants to enter the punishment stage:

$$(1-\delta)(1+g) + \delta \geq (1-\delta)l + \delta^2$$

For our experimental parameters it is easy to verify that all three conditions are satisfied. Hence, our candidate equilibrium is, indeed, an equilibrium.

A.3.2 Imperfect Public Monitoring

The construction becomes slightly more complicated under imperfect public monitoring. Renegotiation-proofness criteria can only be applied if players play public strategies, that is, strategies that condition only on the public history. A special case that has to be considered is the public signal *dd*, that occurs with positive probability even when both players cooperate.

The simplest candidate equilibrium is the following. It starts in the reward state with both players cooperating. In case of a *cc* or a *dd* signal, they stay in the reward state. In case of a *dc* or *cd* signal, they transition to the punishment state, in which the player who appears to have defected plays C, while the other player plays D for one period. In case the public signal contains a c for the punished, the game returns to the reward state. Otherwise, the punishment phase is repeated. Note that in comparison to the equilibrium under perfect monitoring, the incentive to comply as a punished player in the punishment state is weakened by the positive probability of getting away with playing D and still producing a c signal with probability ϵ. The continuation payoff of the reward stage of this candidate equilibrium is:

$$V_r = c + \delta(\epsilon^2 + (1-\epsilon)^2)V_r + \delta(\epsilon(1-\epsilon))V_{pd} + \delta((1-\epsilon)\epsilon)V_{pp}$$

where:

$$V_{pd} = s + \delta(1-\epsilon)V_r + \delta\epsilon V_{pd}$$

$$V_{pp} = b + \delta(1-\epsilon)V_r + \delta\epsilon V_{pp}$$

By plugging V_{pd} and V_{pp} into V_r and simplifying the equation we get:

$$V_r = \frac{c(1-\delta\epsilon) + \delta(1-\epsilon)\epsilon(b+s)}{(1+\delta-2\delta\epsilon)(1-\delta\epsilon) - 2\delta(1-\epsilon)^2}$$

The continuation payoff of deviating from cooperation is:

$$V_d = b + 2\delta\epsilon(1-\epsilon)V_r + \delta(1-\epsilon)^2 V_{pd} + \delta\epsilon^2 V_{pp}$$

By plugging V_{pd} and V_{pp} into V_d and simplifying the equation we get:

$$V_d = b + \frac{\delta\epsilon^2(b+s) - 2s\delta\epsilon}{1-\delta\epsilon} + \frac{\delta(1-\epsilon)[2\epsilon + \delta(1-\epsilon)^2 + \epsilon^2]V_r}{1-\delta\epsilon}$$

It is easy to verify that with the parameters of our paper, $V_r > V_d$, and thus no player has incentive to deviate in the reward stage.

However, the player who is punished in the punishment stage has an incentive to deviate in the punishment state. His continuation payoffs from complying and deviating are:

$$V^{punished}_{comply} = s + \delta(1-\epsilon)V_r + \delta\epsilon V_{pd}$$

$$V^{punished}_{deviate} = d + \delta\epsilon V_r + \delta(1-\epsilon)V_{pd}$$

Plugging V_{pd} and V_r into the two equations above and simplifying yields:

$$V^{punished}_{comply} = \frac{s}{1-\delta\epsilon} + \frac{c\delta(1-\epsilon)}{(1-\delta-2\delta\epsilon)(1-\delta\epsilon)-2\delta(1-\epsilon)^2} + \frac{\delta^2(1-\epsilon)^2\epsilon(b+s)}{(1-\delta-2\delta\epsilon)(1-\delta\epsilon)^2-2\delta(1-\epsilon)^2}$$

$$V^{punished}_{deviate} = \frac{d+\delta\epsilon-\delta\epsilon(d+s)}{1-\delta\epsilon} + \frac{\delta^2(1-\epsilon)\epsilon(b+s)(\epsilon+\delta-2\delta\epsilon)}{(1-\delta-2\delta\epsilon)(1-\delta\epsilon)^2-2\delta(1-\epsilon)^2} + \frac{c\delta(\delta+\epsilon-2\delta\epsilon)}{(1-\delta-2\delta\epsilon)(1-\delta\epsilon)-2\delta(1-\epsilon)^2}$$

With our experimental parameters, the condition $V^{punished}_{comply} \geq V^{punished}_{deviate}$ is violated, which means that the punished player has incentive to deviate in the punishment stage. Hence, this candidate equilibrium is not an equilibrium in our parametrization.

However, if we add a second round to the punishment state, in which both play D, we have found a renegotiation-proof equilibrium for our parametrization. The continuation payoff of the reward stage is still:

$$V_r = c + \delta(\epsilon^2 + (1-\epsilon)^2)V_r + \delta(\epsilon(1-\epsilon))V_{pd} + \delta((1-\epsilon)\epsilon)V_{pp}$$

Since we add a second punishment stage, V_{pd} and V_{pp} change to:

$$V_{pd} = d + \delta[s + \delta(1-\epsilon)V_r + \delta\epsilon V_{pd}]$$

$$V_{pp} = d + \delta[b + \delta(1-\epsilon)V_r + \delta\epsilon V_{pp}]$$

By plugging V_{pd} and V_{pp} into V_r and simplifying the equation we get:

$$V_r = \frac{c(1-\delta^2\epsilon) + \delta\epsilon(1-\epsilon)[2d + \delta(b+s)]}{[1-\delta(1-2\epsilon+2\epsilon^2)](1-\delta^2\epsilon) - 2\delta^3\epsilon(1-\epsilon)^2}$$

The (unchanged) continuation payoff of deviating from cooperation is:

$$V_d = b + 2\delta\epsilon(1-\epsilon)V_r + \delta(1-\epsilon)^2V_{pd} + \delta\epsilon^2V_{pp}$$

By plugging V_{pd} and V_{pp} into V_d and simplifying the equation we get:

$$V_d = \frac{\delta(1-2\epsilon+2\epsilon^2)d}{1-\delta^2\epsilon} + \frac{[1-\delta^2\epsilon(1-\epsilon)]b}{1-\delta^2\epsilon} + \frac{\delta^2(1-\epsilon)^2 s}{1-\delta^2\epsilon} + \frac{[\delta\epsilon(2-\delta^2\epsilon)+\delta^3(1-\epsilon)^2](1-\epsilon)V_r}{1-\delta^2\epsilon}$$

And it is easy to verify that under the parameterization of our paper, $V_r > V_d$, and thus no player has incentive to deviate in the reward stage.

Next, we have to check whether the punisher and the player who gets punished have an incentive to deviate in the punishment stage. The continuation payoff is the same as in the previous case. For the punisher it is obvious that there is no incentive to deviate in the punishment stage. For the player who gets punished, the continuation payoff is:

$$V_{comply}^{punished} = s + \delta(1-\epsilon)V_r + \delta\epsilon V_{pd}$$

$$V_{deviate}^{punished} = d + \delta\epsilon V_r + \delta(1-\epsilon)V_{pd}$$

Plugging V_{pd} and V_r into the two equations and simplifying yields:

$$\begin{aligned} V_{comply}^{punished} &= \frac{s+d\delta\epsilon}{1-\delta^2\epsilon} + \frac{c\delta(1-\epsilon)}{[1-\delta(1-2\epsilon+2\epsilon^2)](1-\delta^2\epsilon)-2\delta^3\epsilon(1-\epsilon)^2} + \\ &\quad \frac{\delta^2\epsilon(1-\epsilon)^2[2d+\delta(b+s)]}{[1-\delta(1-2\epsilon+2\epsilon^2)](1-\delta^2\epsilon)^2-2\delta^3\epsilon(1-\delta^2\epsilon)(1-\epsilon)^2} \end{aligned}$$

$$\begin{aligned} V_{deviate}^{punished} &= d + \frac{\delta(1-\epsilon)(d+s\delta)}{1-\delta^2\epsilon} + \\ &\quad \frac{\delta[c(1-\delta^2\epsilon)+\delta\epsilon(1-\epsilon)(2d+\delta(b+s))](\epsilon-2\delta^2\epsilon+\delta^2)}{[1-\delta(1-2\epsilon+2\epsilon^2)](1-\delta^2\epsilon)-2\delta^3\epsilon(1-\epsilon)^2} \end{aligned}$$

With our parameters, $V_{comply}^{punished} \geq V_{deviate}^{punished}$ is satisfied. Thus, this candidate equilibrium is, indeed, a renegotiation-proof equilibrium.

Note that renegotiation-proof equilibria can be constructed in a way that makes them substantially more efficient than the most efficient equilibrium in strongly-symmetric strategies. This requires the use of a public randomization device to determine whether or not the punishment stage is entered after *cd* or *dc* signals with a probability less than one, such that V_{pd} equals the continuation value of the punishment state with strong symmetry. Efficiency will then be higher because $V_{pp} \geq V_r > V_{pd}$. So, even if they are more complicated than equilibria in strongly-symmetric strategies, players have an incentive to coordinate on them, in addition to potential renegotiation concerns.

A.3.3 Imperfect Private Monitoring

Truthful communication equilibria have a similar structure as renegotiation-proof equilibria, but for a different reason. The condition $V_{pp} = V_r$ stems from the fact that players must not have an incentive to lie about their private signal. In other words, reporting a c must lead to the same continuation value as a report of d. An equilibrium can be constructed as follows. Players start in the reward state, where they cooperate and report their private signals truthfully every round, which essentially transforms the game into one of imperfect public monitoring. Instead of the public signal under public monitoring, the reported signals are used to determine whether the players stay in the reward state or enter the punishment state. Unlike under public monitoring, a dd (reported) signal combination cannot be treated as a cc signal, as this would create an incentive to report d. Instead, the probability of having to enter the punishment state as the punished player must be independent of the own report. To this end, the public randomization device can be used to determine which of the two reports is considered (if any), each with a probability $\pi \leq 1/2$, and never both at the same time. If a report is considered and the reported signal is c, the game stays in the reward state. Otherwise, it transitions to the punishment state, in which the player who appeared to have defected, according to the considered report, becomes the punished player.

The punishment state starts with one period of mutual defection. After this round, the public randomization device determines whether or not a second round of mutual defection is entered with probability ρ. In these one or two rounds of mutual defection, no reports are necessary. In the next and last round of the punishment phase, the punished player plays C while the punisher plays D. After this round, the punisher reports the signal. If the punisher reports a d, the punishment phase is repeated, otherwise the players return to the reward state. With our experimental parameters and $\pi = 0.5$ and $\rho = 0.0498$, it can easily be verified that this is, indeed, an equilibrium (see below). Moreover, it is an equilibrium with a strict incentive not to deviate in the reward state. Hence, it survives Heller's (2017) stability criteria.

The continuation payoff of the reward stage of the proposed equilibrium is:

$$V_r = c + \delta(\pi(1-\epsilon)^2 + (1-\pi))V_r + \delta(\pi(1-\epsilon)\epsilon)V_{pp} + \delta\pi\epsilon V_{pd}$$

Where:

$$V_{pd} = d + \rho[\delta d + \delta(\delta s + \delta(\delta(1-\epsilon)V_r + \delta\epsilon V_{pd}))] + (1-\rho)[\delta s + \delta(\delta(1-\epsilon)V_r + \delta\epsilon V_{pd})]$$

is the continuation payoff from being punished. The continuation payoff as a punisher is:

$$V_{pp} = d + \rho[\delta d + \delta(\delta b + \delta(\delta(1-\epsilon)V_r + \delta\epsilon V_{pp}))] + (1-\rho)[\delta b + \delta(\delta(1-\epsilon)V_r + \delta\epsilon V_{pp})]$$

Moreover, the truthful communication constraint has to hold:

$$V_{pp} = V_r$$

We get a solution for ρ by solving the system of equations. With our experimental parameters and $\pi = 0.5$ we get $\rho = 0.0498$. Moreover, we get:

$$V_{pp} = V_r = \frac{d + \delta b + \rho\delta(d - b + \delta b)}{1 - \rho\delta^3 - (1-\rho)\delta^2}$$
$$V_{pd} = \frac{(1 - \delta + \delta\pi\epsilon)[\delta(1 - \rho + \rho\delta)b + (1 + \rho\delta)d]}{\delta\pi\epsilon[1 - \rho\delta^3 - (1-\rho)\delta^2]} - \frac{c}{\delta\pi\epsilon}$$

Now, we are ready to check whether there are incentives to deviate from following the proposed equilibrium strategies. First, consider whether players have an incentive to deviate in the reward stage. The continuation payoff from deviating is:

$$V_d = b + \delta[\pi\epsilon + (1-\pi)]V_r + \delta\pi(1-\epsilon)V_{pd}$$

Plugging V_r, V_{pd} into the equation above yields:

$$V_d = b + \frac{[(1+\rho\delta)d + \delta(1 - \rho + \rho\delta)b][1 - \delta - \epsilon + 2\delta\epsilon]}{\epsilon[1 - \rho\delta^3 - (1-\rho)\delta^2]} - \frac{c(1-\epsilon)}{\epsilon}$$

Plugging in $\pi = 0.5$ and $\rho = 0.0498$ we see that $V_d < V_r$. Thus, there is no incentive to deviate in the reward stage.

For the punishment stage, we have to check that the punished player has no incentive to deviate. His continuation payoffs from deviating and complying are as follows:

$$V_{deviate}^{punished} = d + \delta(\epsilon V_r + (1-\epsilon)V_{pd})$$
$$V_{comply}^{punished} = s + \delta((1-\epsilon)V_r + \epsilon V_{pd})$$

Plugging V_r, V_{pd} into these equations, we can verify that the first condition $V_{comply}^{punished} > V_{deviate}^{punished}$ holds for our parameters and $\pi = 0.5$.

For the punisher it is obvious that there is no incentive to deviate in the punishment stage either. Thus, the proposed strategy profile is, indeed, a truthful communication equilibrium.

Appendix B Communication Content

Table B1: Categories Generated from Subcategories – All Supergames

Category	Subcategories	Freq.	Frequency in Treatment						$\bar{\kappa}$
			PerPre	PubPre	PrivPre	PerRep	PubRep	PrivRep	
All Supergames									
Coordination (C)	1-16,51,52,71,72	0.503	0.958	0.929	0.946	0.341	0.454	0.479	0.93
Deliberation (D)	17-26,34-41,57,70	0.274	0.643	0.643	0.606	0.192	0.219	0.218	0.72
Relationship (R)	30-33,42-45,47-50,58	0.228	0.103	0.181	0.200	0.219	0.270	0.236	0.71
Trivia (T)	53-55	0.605	0.886	0.810	0.711	0.633	0.515	0.552	1.00
Information (I)	27-29,46,56,59-69	0.215	-	-	-	0.184	0.297	0.285	0.81
Report of action	27,29,46,61,62,66-69	0.008	-	-	-	0.003	0.020	0.006	0.85
Report of action C	27,29,61,66,68	0.062	-	-	-	0.054	0.087	0.081	0.77
Report of action D	46,62,67,69	0.058	-	-	-	0.025	0.070	0.113	0.92
Report of signal	28,56,59,60,66-69	0.141	-	-	-	0.128	0.187	0.190	0.84
Report of signal c	59,68,69	0.066	-	-	-	0.028	0.091	0.118	0.91
Report of signal d	28,56,60,66,67	0.204	-	-	-	0.183	0.273	0.272	0.80
Last 3 Supergames									
Coordination (C)	1-16,51,52,71,72	0.404	0.975	0.974	0.973	0.241	0.328	0.381	0.95
Deliberation (D)	17-26,34-41,57,70	0.223	0.543	0.654	0.58	0.146	0.167	0.186	0.68
Relationship (R)	30-33,42-45,47-50,58	0.258	0.117	0.244	0.293	0.208	0.301	0.29	0.7
Trivia (T)	53-55	0.708	0.963	0.91	0.833	0.73	0.641	0.66	1
Information (I)	27-29,46,56,59-69	0.24	-	-	-	0.176	0.325	0.338	0.79
Report of action	27,29,46,61,62,66-69	0.003	-	-	-	0.001	0.007	0.002	0.8
Report of action C	27,29,61,66,68	0.066	-	-	-	0.06	0.083	0.086	0.75
Report of action D	46,62,67,69	0.064	-	-	-	0.012	0.076	0.139	0.91
Report of signal	28,56,59,60,66-69	0.161	-	-	-	0.112	0.219	0.232	0.82
Report of signal c	59,68,69	0.067	-	-	-	0.013	0.083	0.141	0.91
Report of signal d	28,56,60,66,67	0.227	-	-	-	0.175	0.301	0.318	0.78

Notes: Categories are 1 if the rater identified content related to at least one of the subcategories for a give text unit and 0 otherwise. Frequency indicates the probability that both raters indicated one of the respective subcategories for a randomly selected text unit. Frequencies < 0.001 omitted (-). $\bar{\kappa}$ is the average Cohen's Kappa over all treatments. Mean $\bar{\kappa}$ of all generated categories is 0.84.

Table B2: Battery of Subcategories for Coding – All Supergames

#	Subcategory	Category	Freq.	Frequency in Treatment						$\bar{\kappa}$
				PerPre	PubPre	PrivPre	PerRep	PubRep	PrivRep	
1	Proposal: both C	C	0.246	0.542	0.420	0.500	0.169	0.210	0.231	0.85
2	Proposal: both D	C	0.033	0.071	0.077	0.054	0.012	0.039	0.030	0.81
3	Proposal: alternate	C	0.013	0.024	0.058	0.066	0.005	0.001	0.013	0.75
4	Proposal: self D other C	C	0.010	0.013	0.047	0.031	0.006	0.004	0.008	0.72
5	Proposal: self C other D	C	0.005	0.008	0.008	0.009	0.001	0.001	0.010	0.56
6	Proposal: other coordination	C	0.006	0.029	0.044	0.017	-	0.005	0.002	0.41
7	Question: what action other	C	0.009	0.024	0.025	0.017	0.009	0.005	0.005	0.51
8	Announcement: C	C	0.009	0.016	0.047	0.006	0.006	0.006	0.008	0.59
9	Announcement: D	C	0.007	0.021	0.014	0.017	0.006	0.006	0.004	0.76
10	Rejection of proposal	C	0.004	0.005	0.005	0.017	0.002	0.004	0.002	0.59
11	Acceptance proposal	C	0.297	0.685	0.585	0.617	0.189	0.256	0.268	0.85
12	Implicit punishment threat for D	C	0.003	0.005	0.003	0.029	-	0.004	0.001	0.33
13	Punishment threat grim	C	0.003	0.005	0.014	0.003	0.005	-	-	0.57
14	Punishment threat lenient grim	C	-	-	-	-	-	-	-	-
15	Approval of punishment threat	C	0.002	-	-	0.014	0.002	0.001	0.001	0.41
16	Ask for coordination	C	0.041	0.119	0.115	0.120	0.011	0.031	0.041	0.79
17	Benefits of C	D	0.051	0.161	0.099	0.151	0.038	0.034	0.035	0.63
18	Benefits of D	D	0.007	0.013	0.027	0.023	0.002	0.005	0.005	0.53
19	Benefits of asymmetric play	D	0.003	0.003	0.008	0.011	0.002	0.001	0.003	0.50
20	Related to fairness discussion	D	0.009	0.040	0.025	0.031	0.002	0.002	0.010	0.66
21	Related to strategic uncertainty	D	0.050	0.095	0.206	0.100	0.026	0.042	0.036	0.56
22	Related to payoffs	D	0.055	0.188	0.181	0.154	0.029	0.035	0.036	0.71
23	Related to Prisoner's dilemma	D	0.004	0.058	0.003	-	0.002	-	-	0.84
24	Related to game theory	D	0.002	0.011	0.005	0.009	-	0.001	-	0.54
25	Future benefit of C	D	0.009	0.016	0.019	0.054	0.006	0.007	0.003	0.49
26	Short term incentives of D	D	-	0.005	-	-	-	-	-	0.05
27	Attribute other d to randomness	I	0.004	-	-	-	0.006	0.006	0.002	0.34
28	Attribute own d to randomness	I	0.006	-	-	-	0.010	0.007	0.005	0.36
29	Assurance to have played C	I	0.002	-	-	-	-	0.003	0.003	0.21
30	Promise	R	0.021	0.040	0.069	0.077	0.014	0.015	0.013	0.71
31	Distrust	R	0.002	0.005	-	-	0.002	0.001	0.002	0.27
32	Trust	R	0.012	0.016	0.019	0.023	0.011	0.010	0.012	0.63
33	Argue for trustworthy behavior	R	0.026	0.048	0.102	0.111	0.021	0.011	0.014	0.62
34	Report payoff from past games	D	0.028	0.063	0.022	0.006	0.030	0.025	0.027	0.72
35	Report signals of past games	D	0.013	0.042	-	0.009	0.013	0.014	0.011	0.42
36	Good past experience with CC	D	0.051	0.151	0.126	0.100	0.028	0.048	0.037	0.75
37	Good past experience with DD	D	0.001	0.003	0.003	0.003	-	0.002	0.001	0.43
38	Bad past experience with CC	D	0.008	0.021	0.060	0.014	0.002	0.001	0.007	0.44
39	Bad past experience with CC	D	-	-	0.003	-	-	0.001	0.001	0.24
40	Good past experience asym. play	D	0.001	0.005	0.011	0.003	-	-	0.001	0.53
41	Bad past experience asym. play	D	0.001	0.003	0.003	0.006	-	0.002	-	0.52
42	Positive feedback after CC	R	0.119	-	-	-	0.115	0.167	0.143	0.81
43	Positive feedback after DD	R	0.002	-	-	-	0.002	0.003	0.001	0.65
44	Positive feedback after asym. play	R	0.001	-	-	-	0.001	0.002	0.002	0.64
45	Empathy	R	0.016	-	0.003	-	0.014	0.022	0.020	0.57
46	Confess D	I	-	-	-	-	-	0.001	-	0.40
47	Apology	R	0.002	-	-	-	0.004	0.001	0.001	0.48
48	Justification of play	R	0.001	-	-	-	0.003	0.001	-	0.19
49	Accusation of cheating	R	0.007	-	-	-	0.004	0.008	0.014	0.55
50	Verbal punishment	R	0.001	-	-	-	0.001	0.001	-	0.57
51	Renegotiation	C	0.001	-	-	-	-	0.001	0.001	0.06
52	Argument against punishment	C	-	-	-	-	-	-	-	-
53	Small talk	T	0.247	0.820	0.739	0.583	0.176	0.141	0.168	0.70
54	Off topic	T	0.283	0.193	0.093	0.094	0.368	0.229	0.330	0.58
55	Boredom	T	0.011	0.021	-	0.014	0.012	0.012	0.010	0.57
56	Disappointed after d signal	I	0.024	-	-	-	0.029	0.030	0.025	0.55
57	Confusion	D	0.033	0.058	0.085	0.026	0.015	0.036	0.037	0.35
58	Motivational talk	R	0.026	-	-	-	0.030	0.041	0.022	0.51
59	Report: own signal c	I	0.004	-	-	-	0.001	0.006	0.008	0.65
60	Report: own signal d	I	0.012	-	-	-	0.005	0.021	0.016	0.82
61	Report: own action C	I	0.005	-	-	-	0.001	0.013	0.005	0.50
62	Report: own action D	I	0.003	-	-	-	-	0.009	0.001	0.78
63	Ask for others payoff	I	0.019	-	-	-	0.010	0.023	0.035	0.83
64	Ask for others signal	I	0.006	-	-	-	0.003	0.004	0.014	0.45
65	Ask for others action	I	0.006	-	-	-	0.003	0.011	0.007	0.85
66	Report: own payoff 0	I	0.025	-	-	-	0.012	0.032	0.047	0.95
67	Report: own payoff 17	I	0.004	-	-	-	0.002	0.009	0.003	0.90
68	Report: own payoff 30	I	0.022	-	-	-	0.011	0.016	0.051	0.96
69	Report: own payoff 37	I	0.001	-	-	-	0.001	0.002	0.001	0.73
70	Being cheated on in past games	D	0.005	-	-	0.003	0.003	0.007	0.006	0.45
71	Counter-proposal	C	-	-	-	-	-	0.001	0.001	0.46
72	Rejection of punishment	C	-	-	0.003	-	-	-	-	0.67

Notes: Subcategories are 1 if the rater identified content related to the subcategory for a given text unit and 0 otherwise. Category are Coordination (C), Deliberation (D), Relationship (R), Trivia (T) and Information (I). Frequency indicates the probability that both raters indicated the respective subcategory for a randomly selected text unit. Frequencies < 0.001 omitted (-). $\bar{\kappa}$ is the average Cohen's Kappa over all treatments. Mean $\bar{\kappa}$ of all subcategories with an overall frequency > 0.01 is 0.65.

Table B3: Battery of Subcategories for Coding – Last Three Supergames

#	Subcategory	Category	Freq.	Frequency in Treatment						$\bar{\kappa}$
				PerPre	PubPre	PrivPre	PerRep	PubRep	PrivRep	
1	Proposal: both C	C	0.224	0.673	0.487	0.613	0.131	0.177	0.195	0.88
2	Proposal: both D	C	0.01	0.012	0.058	0.013	0.004	0.011	0.005	0.78
3	Proposal: alternate	C	0.005	0.025	0.032	0.013	-	-	0.007	0.75
4	Proposal: self D other C	C	0.002	-	0.026	-	-	-	0.004	0.76
5	Proposal: self C other D	C	0.002	-	0.006	0.007	-	-	0.005	0.64
6	Proposal: other coordination	C	0.005	0.012	0.071	0.007	-	0.002	-	0.56
7	Question: what action other	C	0.003	-	0.026	0.007	0.001	-	0.005	0.44
8	Announcement: C	C	0.007	0.006	0.058	-	0.002	0.004	0.01	0.54
9	Announcement: D	C	0.001	0.006	0.019	-	-	-	0.001	0.83
10	Rejection of proposal	C	0.003	0.006	0.006	0.013	-	0.003	0.002	0.6
11	Acceptance proposal	C	0.246	0.747	0.59	0.66	0.15	0.185	0.207	0.88
12	Implicit punishment threat for D	C	0.003	0.006	-	0.033	0.001	0.003	-	0.28
13	Punishment threat grim	C	0.002	-	-	0.007	0.005	-	-	0.52
14	Punishment threat lenient grim	C	-	-	-	-	-	-	-	-
15	Approval of punishment threat	C	0.002	-	-	0.027	0.002	-	-	0.4
16	Ask for coordination	C	0.022	0.062	0.096	0.093	0.004	0.01	0.024	0.79
17	Benefits of C	D	0.04	0.123	0.122	0.167	0.024	0.025	0.026	0.62
18	Benefits of D	D	0.001	-	0.006	0.007	-	0.001	-	0.28
19	Benefits of asymmetric play	D	-	-	0.006	-	-	-	-	0.4
20	Related to fairness discussion	D	0.007	0.037	0.019	0.033	0.002	-	0.008	0.66
21	Related to strategic uncertainty	D	0.036	0.068	0.237	0.093	0.013	0.028	0.024	0.54
22	Related to payoffs	D	0.032	0.136	0.147	0.113	0.01	0.02	0.02	0.71
23	Related to Prisoner's dilemma	D	0.003	0.056	-	-	0.002	-	-	0.88
24	Related to game theory	D	0.001	0.012	-	0.013	0.001	-	-	0.71
25	Future benefit of C	D	0.007	0.006	0.013	0.067	0.006	0.006	0.001	0.54
26	Short term incentives of D	D	-	-	-	-	-	-	-	-
27	Attribute other d to randomness	I	0.004	-	-	-	0.005	0.006	0.002	0.31
28	Attribute own d to randomness	I	0.006	-	-	-	0.01	0.004	0.005	0.3
29	Assurance to have played C	I	0.002	-	-	-	-	0.003	0.005	0.22
30	Promise	R	0.026	0.062	0.103	0.12	0.015	0.017	0.012	0.72
31	Distrust	R	0.002	0.006	-	-	0.002	0.001	0.003	0.36
32	Trust	R	0.012	0.006	0.019	0.02	0.012	0.006	0.016	0.6
33	Argue for trustworthy behavior	R	0.029	0.062	0.135	0.18	0.014	0.012	0.015	0.61
34	Report payoff from past games	D	0.025	0.043	0.019	-	0.024	0.023	0.03	0.65
35	Report signals of past games	D	0.017	0.062	-	0.02	0.014	0.016	0.014	0.44
36	Good past experience with CC	D	0.055	0.142	0.179	0.167	0.029	0.048	0.039	0.73
37	Good past experience with DD	D	0.001	0.006	0.006	-	-	-	-	0.36
38	Bad past experience with CC	D	0.01	0.019	0.109	0.033	0.001	-	0.007	0.43
39	Bad past experience with CC	D	0.001	-	-	-	-	0.001	0.001	0.31
40	Good past experience asym. play	D	0.001	-	0.013	-	-	-	-	0.5
41	Bad past experience asym. play	D	0.001	-	-	-	-	0.002	-	0.67
42	Positive feedback after CC	R	0.14	-	-	-	0.11	0.201	0.178	0.8
43	Positive feedback after DD	R	0.001	-	-	-	0.001	-	0.001	0.44
44	Positive feedback after asym. play	R	-	-	-	-	-	-	-	-
45	Empathy	R	0.02	-	-	-	0.017	0.025	0.029	0.59
46	Confess D	I	-	-	-	-	-	0.001	-	1
47	Apology	R	-	-	-	-	0.001	-	-	0.15
48	Justification of play	R	0.001	-	-	-	0.001	0.001	-	0.12
49	Accusation of cheating	R	0.009	-	-	-	0.002	0.01	0.018	0.61
50	Verbal punishment	R	-	-	-	-	-	0.001	-	0.29
51	Renegotiation	C	0.001	-	-	-	-	-	0.002	0.05
52	Argument against punishment	C	-	-	-	-	-	-	-	-
53	Small talk	T	0.241	0.92	0.821	0.66	0.156	0.127	0.177	0.66
54	Off topic	T	0.394	0.315	0.122	0.14	0.473	0.342	0.455	0.58
55	Boredom	T	0.014	0.043	-	0.02	0.016	0.012	0.011	0.52
56	Disappointed after d signal	I	0.029	-	-	-	0.039	0.038	0.021	0.56
57	Confusion	D	0.022	0.031	0.006	0.027	0.012	0.023	0.031	0.25
58	Motivational talk	R	0.028	-	-	-	0.027	0.046	0.026	0.49
59	Report: own signal c	I	0.002	-	-	-	-	0.003	0.005	0.5
60	Report: own signal d	I	0.01	-	-	-	0.002	0.016	0.017	0.8
61	Report: own action C	I	0.005	-	-	-	-	0.011	0.005	0.43
62	Report: own action D	I	0.001	-	-	-	-	0.002	0.001	0.75
63	Ask for others payoff	I	0.018	-	-	-	0.006	0.017	0.04	0.77
64	Ask for others signal	I	0.002	-	-	-	0.002	0.002	0.003	0.2
65	Ask for others action	I	0.004	-	-	-	0.002	0.006	0.006	0.82
66	Report: own payoff 0	I	0.028	-	-	-	0.01	0.034	0.054	0.94
67	Report: own payoff 17	I	0.001	-	-	-	-	0.004	0.001	0.91
68	Report: own payoff 30	I	0.023	-	-	-	0.002	0.017	0.063	0.96
69	Report: own payoff 37	I	0.001	-	-	-	0.001	0.001	-	0.67
70	Being cheated on in past games	D	0.008	-	-	-	0.004	0.011	0.012	0.47
71	Counter-proposal	C	-	-	-	-	-	-	0.001	0.33
72	Rejection of punishment	C	-	-	-	-	-	-	-	-

Notes: See notes of Table B2. Data from last three supergames.

Table B4: Communication after First Defection Signal - All Supergames

	Public Repeated			Private Repeated		
Category	$\omega \neq \{c,c\}$	$\omega = \{c,c\}$	Δ	$\omega_j = d$	$\omega_j = c$	Δ
Coordination	0.43	0.28	0.14	0.50	0.29	0.21
Deliberation	0.11	0.13	-0.02	0.06	0.10	-0.04
Relationship	0.25	0.40	-0.15	0.23	0.29	-0.06
Information	0.64	0.34	0.31	0.66	0.32	0.34
Trivia	0.36	0.53	-0.17	0.37	0.55	-0.18
Report of action	0.40	0.02	0.38	0.42	0.10	0.32
Report of C	0.40	0.02	0.38	0.42	0.10	0.32
Report of D	0.00	0.00	0.00	0.00	0.00	0.00
Report of signal	0.55	0.33	0.22	0.65	0.31	0.34
Report of c	0.07	0.33	-0.26	0.03	0.30	-0.28
Report of d	0.49	0.00	0.49	0.63	0.00	0.62

Notes: Frequency of communication categories for subject-round observations with cooperative history up to round t. A Subject has a cooperative history if her previous actions were C and all signals she observed in rounds $< t$ were c. Frequencies illustrate the use of categories dependent on signals in round t. Frequency indicates the probability that both raters indicated the category for a text unit. Frequencies < 0.001 omitted (-).

Table B5: Communication after First Defection Signal – All Supergames

#	Subcategory	Public Repeated			Private Repeated		
		$\omega \neq \{c,c\}$	$\omega = \{c,c\}$	Δ	$\omega_j = d$	$\omega_j = c$	Δ
1	Proposal: both C	0.164	0.145	0.019	0.168	0.143	0.025
2	Proposal: both D	0.013	0.012	0.001	-	0.011	-0.011
3	Proposal: alternate	-	-	-	-	0.005	-0.005
4	Proposal: self D other C	-	-	-	0.017	0.003	0.014
5	Proposal: self C other D	0.007	-	0.007	-	-	-
6	Proposal: other coordination	-	0.004	-0.004	-	-	-
7	Question: what action other	-	-	-	-	-	-
8	Announcement: C	0.007	0.002	0.005	0.025	0.003	0.022
9	Announcement: D	0.007	-	0.007	-	-	-
10	Rejection of proposal	-	-	-	-	0.002	-0.002
11	Acceptance proposal	0.178	0.164	0.014	0.143	0.165	-0.022
12	Implicit punishment threat for D	-	-	-	-	0.002	-0.002
13	Punishment threat grim	-	-	-	-	-	-
14	Punishment threat lenient grim	-	-	-	-	-	-
15	Approval of punishment threat	-	-	-	-	0.002	-0.002
16	Ask for coordination	0.013	0.004	0.009	0.025	0.005	0.02
17	Benefits of C	0.007	0.008	-0.001	0.008	0.017	-0.009
18	Benefits of D	-	-	-	-	-	-
19	Benefits of asymmetric play	-	-	-	-	-	-
20	Related to fairness discussion	-	-	-	-	-	-
21	Related to strategic uncertainty	0.013	0.017	-0.004	0.025	0.011	0.014
22	Related to payoffs	0.013	0.006	0.007	0.017	0.016	0.001
23	Related to Prisoner's dilemma	-	-	-	-	-	-
24	Related to game theory	-	0.002	-0.002	-	-	-
25	Future benefit of C	0.007	0.002	0.005	0.008	0.002	0.006
26	Short term incentives of D	-	-	-	-	-	-
27	Attribute other d to randomness	0.033	-	0.033	-	0.002	-0.002
28	Attribute own d to randomness	0.053	-	0.053	0.042	-	0.042
29	Assurance to have played C	-	-	-	0.008	0.003	0.005
30	Promise	-	0.012	-0.012	0.008	-	0.008
31	Distrust	-	-	-	0.008	-	0.008
32	Trust	0.013	0.006	0.007	0.084	0.003	0.081
33	Argue for trustworthy behavior	0.013	-	0.013	-	0.003	-0.003
34	Report payoff from past games	-	0.019	-0.019	0.008	0.003	0.005
35	Report signals of past games	-	0.004	-0.004	-	0.005	-0.005
36	Good past experience with CC	-	0.017	-0.017	-	0.002	-0.002
37	Good past experience with DD	-	-	-	-	-	-
38	Bad past experience with CC	-	-	-	-	-	-
39	Bad past experience with CC	-	-	-	-	0.002	-0.002
40	Good past experience asym. play	-	-	-	-	-	-
41	Bad past experience asym. play	-	-	-	-	-	-
42	Positive feedback after CC	-	0.321	-0.321	0.017	0.233	-0.216
43	Positive feedback after DD	-	-	-	-	-	-
44	Positive feedback after asym. play	-	-	-	0.008	0.002	0.006
45	Empathy	0.132	-	0.132	-	0.027	-0.027
46	Confess D	-	-	-	-	-	-
47	Apology	-	0.002	-0.002	-	-	-
48	Justification of play	-	-	-	-	-	-
49	Accusation of cheating	0.046	-	0.046	0.143	-	0.143
50	Verbal punishment	0.007	-	0.007	-	-	-
51	Renegotiation	-	0.002	-0.002	-	-	-
52	Argument against punishment	-	-	-	-	-	-
53	Small talk	0.02	0.014	0.006	0.059	0.046	0.013
54	Off topic	0.118	0.269	-0.151	0.151	0.38	-0.229
55	Boredom	-	0.015	-0.015	-	0.008	-0.008
56	Disappointed after d signal	0.191	-	0.191	0.185	-	0.185
57	Confusion	0.059	0.044	0.015	-	0.027	-0.027
58	Motivational talk	0.033	0.089	-0.056	0.008	0.029	-0.021
59	Report: own signal c	0.007	0.004	0.003	0.008	0.008	-
60	Report: own signal d	0.151	-	0.151	0.16	0.002	0.158
61	Report: own action C	0.092	0.004	0.088	0.008	0.006	0.002
62	Report: own action D	-	-	-	-	-	-
63	Ask for others payoff	0.086	0.008	0.078	0.059	0.035	0.024
64	Ask for others signal	0.013	0.002	0.011	0.034	0.016	0.018
65	Ask for others action	0.066	-	0.066	0.042	-	0.042
66	Report: own payoff 0	0.197	-	0.197	0.395	0.003	0.392
67	Report: own payoff 17	-	-	-	-	-	-
68	Report: own payoff 30	0.066	0.015	0.051	-	0.076	-0.076
69	Report: own payoff 37	-	-	-	-	-	-
70	Being cheated on in past games	-	0.006	-0.006	-	0.003	-0.003
71	Counter-proposal	-	-	-	-	0.002	-0.002
72	Rejection of punishment	-	-	-	-	-	-

Notes: Frequency of subcategories for subject-round observations with cooperative history in round t. A Subject has a cooperative history if her previous actions were C and all signals she observed in rounds $< t$ were c. Frequencies illustrate the use of subcategories dependent on signals in round t. Frequency indicates the probability that both raters indicated the respective subcategory for a randomly selected text unit. Frequencies < 0.001 omitted (-).

Table B6: Communication after First Defection Signal – Last Three Supergames

		Public Repeated			Private Repeated		
#	Subcategory	$\omega \neq \{c, c\}$	$\omega = \{c, c\}$	Δ	$\omega_j = d$	$\omega_j = c$	Δ
1	Proposal: both C	0.136	0.094	0.042	0.182	0.112	0.07
2	Proposal: both D	-	0.01	-0.01	-	0.013	-0.013
3	Proposal: alternate	-	-	-	-	0.005	-0.005
4	Proposal: self D other C	-	-	-	-	0.005	-0.005
5	Proposal: self C other D	-	-	-	-	-	-
6	Proposal: other coordination	-	-	-	-	-	-
7	Question: what action other	-	-	-	-	-	-
8	Announcement: C	-	0.003	-0.003	0.03	0.005	0.025
9	Announcement: D	-	-	-	-	-	-
10	Rejection of proposal	-	-	-	-	0.003	-0.003
11	Acceptance proposal	0.123	0.094	0.029	0.121	0.142	-0.021
12	Implicit punishment threat for D	-	-	-	-	-	-
13	Punishment threat grim	-	-	-	-	-	-
14	Punishment threat lenient grim	-	-	-	-	-	-
15	Approval of punishment threat	-	-	-	-	-	-
16	Ask for coordination	-	-	-	0.045	0.003	0.042
17	Benefits of C	-	-	-	-	0.013	-0.013
18	Benefits of D	-	-	-	-	-	-
19	Benefits of asymmetric play	-	-	-	-	-	-
20	Related to fairness discussion	-	-	-	-	-	-
21	Related to strategic uncertainty	-	0.01	-0.01	-	0.003	-0.003
22	Related to payoffs	0.012	0.006	0.006	0.015	0.008	0.007
23	Related to Prisoner's dilemma	-	-	-	-	-	-
24	Related to game theory	-	-	-	-	-	-
25	Future benefit of C	0.012	0.003	0.009	-	-	-
26	Short term incentives of D	-	-	-	-	-	-
27	Attribute other d to randomness	0.037	-	0.037	-	-	-
28	Attribute own d to randomness	0.025	-	0.025	0.045	-	0.045
29	Assurance to have played C	-	-	-	0.015	0.005	0.01
30	Promise	-	0.01	-0.01	-	-	-
31	Distrust	-	-	-	0.015	-	0.015
32	Trust	0.025	0.003	0.022	0.136	0.005	0.131
33	Argue for trustworthy behavior	0.025	-	0.025	-	0.003	-0.003
34	Report payoff from past games	-	0.026	-0.026	-	-	-
35	Report signals of past games	-	0.003	-0.003	-	0.008	-0.008
36	Good past experience with CC	-	0.023	-0.023	-	0.003	-0.003
37	Good past experience with DD	-	-	-	-	-	-
38	Bad past experience with CC	-	-	-	-	-	-
39	Bad past experience with CC	-	-	-	-	0.003	-0.003
40	Good past experience asym. play	-	-	-	-	-	-
41	Bad past experience asym. play	-	-	-	-	-	-
42	Positive feedback after CC	-	0.314	-0.314	-	0.254	-0.254
43	Positive feedback after DD	-	-	-	-	-	-
44	Positive feedback after asym. play	-	-	-	-	-	-
45	Empathy	0.16	-	0.16	-	0.037	-0.037
46	Confess D	-	-	-	-	-	-
47	Apology	-	-	-	-	-	-
48	Justification of play	-	-	-	-	-	-
49	Accusation of cheating	0.074	-	0.074	0.182	-	0.182
50	Verbal punishment	0.012	-	0.012	-	-	-
51	Renegotiation	-	-	-	-	-	-
52	Argument against punishment	-	-	-	-	-	-
53	Small talk	0.025	-	0.025	0.091	0.064	0.027
54	Off topic	0.185	0.353	-0.168	0.197	0.479	-0.282
55	Boredom	-	0.01	-0.01	-	-	-
56	Disappointed after d signal	0.235	-	0.235	0.136	-	0.136
57	Confusion	0.062	0.036	0.026	-	0.035	-0.035
58	Motivational talk	0.049	0.071	-0.022	-	0.024	-0.024
59	Report: own signal c	-	0.003	-0.003	-	0.005	-0.005
60	Report: own signal d	0.111	-	0.111	0.121	0.003	0.118
61	Report: own action C	0.086	-	0.086	0.015	0.011	0.004
62	Report: own action D	-	-	-	-	-	-
63	Ask for others payoff	0.062	-	0.062	0.091	0.045	0.046
64	Ask for others signal	-	0.003	-0.003	-	0.003	-0.003
65	Ask for others action	0.049	-	0.049	0.045	-	0.045
66	Report: own payoff 0	0.21	-	0.21	0.5	0.003	0.497
67	Report: own payoff 17	-	-	-	-	-	-
68	Report: own payoff 30	0.074	0.006	0.068	-	0.091	-0.091
69	Report: own payoff 37	-	-	-	-	-	-
70	Being cheated on in past games	-	0.01	-0.01	-	0.005	-0.005
71	Counter-proposal	-	-	-	-	0.003	-0.003
72	Rejection of punishment	-	-	-	-	-	-

Notes: See notes of Table B5. Data from last three supergames.

Appendix C Strategy Inference

We use the strategy frequency estimation method (Dal Bó and Fréchette, 2011) and its adaptation to behavior strategies (Breitmoser, 2015) to analyze strategy choice across treatments. To study the evolution of strategy choice over supergames, we extend the existing methods and model strategy choices as a function of covariates in the spirit of latent class regression (Dayton and Macready, 1988; Bandeen-Roche et al., 1997). A documentation of the methods can be found in Dvorak (2020).

Model Definition

Let p_k denote the share of strategy $k \in \{1, \cdots, K\}$ in the population and $\pi_{s_k} \in [0, 1]$ the probability of cooperation prescribed by strategy k in state $s_k \in S_k$. When estimating pure strategies, we assume that there exists a pure underlying response probability $\xi_{s_k} \in \{0, 1\}$ to each π_{s_k}. The pure responses are confounded by a tremble which implements the wrong action and occurs with probability $\gamma \in [0, 0.5]$. We assume that the probability of a tremble is the same for all individuals, supergames and rounds and that the realizations of trembles are independent across these dimensions. The probability of cooperation for pure strategy k in state s_k is given by: $\pi_{s_k} = \xi_{s_k}(1 - \gamma) + (1 - \xi_{s_k})(1 - \gamma)$. Let y_{is_k} denote the number of times individual $i \in \{1, \cdots, N\}$ cooperates in n_{is_k} observations of state s_k of strategy k. We report the maximum-likelihood estimates of the parameters p_k, π_{s_k} (or alternatively ξ_{s_k} and γ) that maximize the log-likelihood

$$\ln L = \sum_{i=1}^{N} \ln \left(\sum_{k=1}^{K} p_k \prod_{s_k \in S_k} (\pi_{s_k})^{y_{is_k}} (1 - \pi_{s_k})^{n_{is_k} - y_{is_k}} \right).$$

To find the global optima of the parameters, we execute the EM-algorithm (Dempster et al., 1977) from multiple random starting points and use the Newton-Raphson method to check for convergence.

To obtain the results reported in Tables 5 and C5, we perform treatment-wise strategy estimation starting with the candidate set of 23 strategies listed in Tables C1-C4. For each treatment, the number of strategies is selected based the the ICL information criterion (Biernacki et al., 2000) that has been used to select the number of strategies before (Breitmoser, 2015). ICL is the Bayesian information criterion (Schwarz, 1978) penalized by the entropy of the data according to

$$\text{ICL} = -\ln L + \frac{df}{2} log(N) - \sum_{i=1}^{N} \sum_{K=1}^{K} \theta_{ik} log(\theta_{ik}),$$

where df represents the number of free parameters of the model and θ_{ik} is the posterior probability that individual i plays strategy k, given by

$$\theta_{ik} = \frac{p_k \prod_{s_k \in S_k} (\pi_{s_k})^{y_{is_k}} (1 - \pi_{s_k})^{n_{is_k} - y_{is_k}}}{\sum_{k=1}^{K} p_k \prod_{s_k \in S_k} (\pi_{s_k})^{y_{is_k}} (1 - \pi_{s_k})^{n_{is_k} - y_{is_k}}}.$$

Latent-Class Regression Models

An intuitive approach to analyze the effect of covariates for strategy choices is to assign individuals to strategies based on the posterior probability assignments θ_{ik} and use the assignments as the dependent variable in a multinomial model. However, this approach leads to downward biased coefficients for the effects of covariates as shown by Bolck et al. (2004). To circumvent this problem, we explore the evolution of strategy choice over supergames based on latent class regression models displayed in Table C5. The underlying models assume that subjects' strategy choices over supergames reflect repeated independent draws from a probability distribution over a fixed set of strategies. The probability distribution is modeled as a function of the supergame number. The log-likelihood of the latent-class regression model is

$$\ln L = \sum_{j=1}^{J} \ln \left(\sum_{k=1}^{K} p_{jk} \prod_{s_k \in S_k} (\pi_{s_k})^{y_{js_k}} (1 - \pi_{s_k})^{n_{js_k} - y_{js_k}} \right),$$

where the index $j \in \{1, \cdots, J\}$ enumerates the unique combinations of individuals $i \in \{1, \cdots, N\}$ and supergames $g \in \{1, \cdots, G\}$ as $J = N \cdot G$. The parameter p_{jk} reflects the prior probability that individual i uses strategy k in supergame g and the share p_k of strategy k in supergame g is the expected value of the prior p_{jk} in supergame g. To model the evolution of the strategy choices over supergames, we assume that the prior probabilities p_{jk} are a function of the supergame number. The latent-class regression approach suggests to model the log-odds of using strategy k as compared to the first strategy in the set based on the multinomial logit link function (Agresti, 2003)

$$\ln(p_{jk}/p_{i1}) = X_j \beta_k \ \forall \ k \in \{1, \cdots, K\},$$

where β_k is a column vector of coefficients for strategy k and X_j a row vector for observation j consisting of an intercept and the supergame number. Reformulation of the K equations yields

$$p_{jk} = \frac{e^{X_j \beta_k}}{\sum_{k=1}^{K} e^{X_i \beta_k}},$$

and the maximum-likelihood estimates of the parameters β_k and π_{s_k} (or alternatively ξ_{s_k} and γ) can be found based on a variant of the EM algorithm augmented by a Newton-Raphson step (Bandeen-Roche et al., 1997).

The coefficients β_0 displayed in Table C5 reflect the relative prevalence of the lenient and forgiving strategies in contrast to ALLD in the first supergame. Negative (positive) values indicate that a strategy is less (more) frequent than ALLD in the first supergame. The treatment with perfect monitoring and repeated communication is an exception. The ICL criterion suggests that the data of this treatment is best summarized by a single strategy (SGRIM). Since latent-class regression coefficients cannot be estimated for a model with only one strategy, we report the second best model with DTF3T and SGRIM. As ALLD is not included in this model, DTF3T serves as the reference strategy and β_0 reflects the relative prevalence of SGRIM in contrast to DTF3T in the first supergame. The coefficients β_{sg} displayed in Table C5 indicate the time trend in the relative frequency of strategies over the first three supergames. A negative (positive) coefficient indicates that the strategy becomes less (more) popular compared to ALLD in the first three supergames.

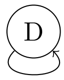
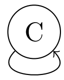
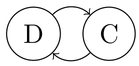
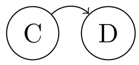
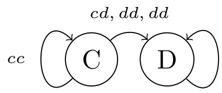
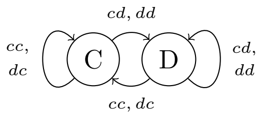
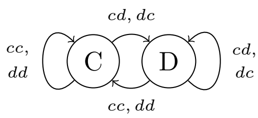
Adaptation of Strategies

Tables C1-C4 list the set of 23 strategies used to obtain the strategy estimation results reported in Table 5 and Table C5. Circles in Table C3 represent strategy states and arrows deterministic state transitions. In the treatments with perfect monitoring, the state traditions can in principle be triggered by action profiles, the two public signals or action-signal combinations. In the treatments with public monitoring, transitions can be triggered by the two public signals or action-signal combinations. We assume that all strategies in the set condition on the same information, run the estimation for the 3 (2) possibilities and report the results with the highest log-likelihood.

Strategies 1-20 and their descriptions are taken from Fudenberg et al. (2012). The remaining three strategies are behavior strategies. Two of the behavior strategies are motivated by Backhaus and Breitmoser's (2018) analysis, who present evidence suggesting that subjects play semi-grim M1BF strategies, and further find that a small share of (noise) players randomize 50–50 in all states. Taking these findings into account, we include a strategy RAND that predicts a 50% cooperation probability after all histories. We also include a semi-grim strategy SGRIM which starts with cooperation and cooperates with probability of 1 in the *cc*-state, probability 0 in the *dd*-state, and probability 0.35 in the *cd* and *dc* states. The value 0.35 is the average cooperation probability that Backhaus and Breitmoser (2018) report for these states in the lower panel of Table 1 of their paper. We choose this value instead of estimating the probability from our data, as this would give the strategy an additional free parameter and therefore an advantage over the other strategies in the set. In another strategy estimation exercise, we infer response probabilities from the data but keep the number of free parameters the same for all strategies.

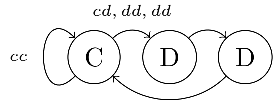
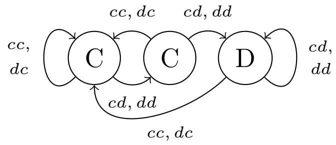
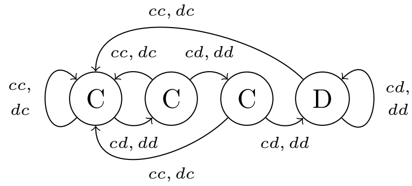
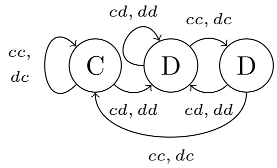
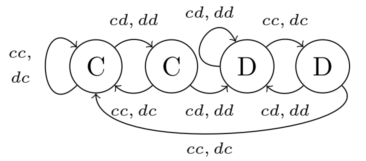
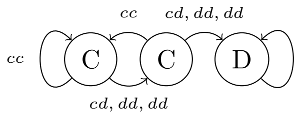
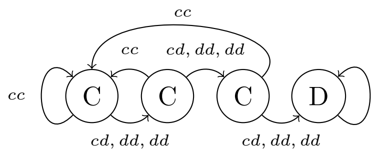
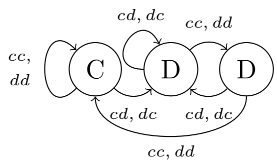
The third behavior strategy that we include is an equilibrium-M1BF strategy M1BF$_{eq}$ that starts with cooperation in the first round. For the imperfect monitoring structures this is the T1BF strategy ($\sigma_{\emptyset} = 1, \sigma_{cc} = 1, \sigma_{cd} = 0.5, \sigma_{dc} = 1, \sigma_{dd} = 0$) which conditions on the own action and the signal about the action of the partner in the previous round. For perfect monitoring, where subjects are more likely to condition behavior on observed actions rather than signals, it is ($\sigma_{\emptyset} = 1, \sigma_{cc} = 1, \sigma_{cd} = 0.75, \sigma_{dc} = 0.5, \sigma_{dd} = 0$), which is the only equilibrium M1BF strategy assuming that subjects cooperate after mutual cooperation, and defect after mutual defection (see Appendix A for the derivation of these equilibrium strategies).

Table C1: Strategies 1-7

Acronym	Description	Automaton
ALLD	Always play D.	
ALLC	Always play C.	
DC	Start with D, then alternate between C and D.	
FC	Play C in the first round, then D forever.	
Grim	Play C until either player plays D, then play D forever.	
TFT	Play C unless partner played D last round.	
PTFT (WSLS)	Play C if both players chose the same move last round, otherwise play D.	

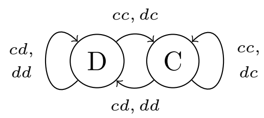
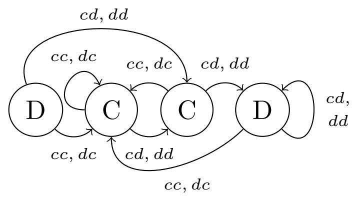
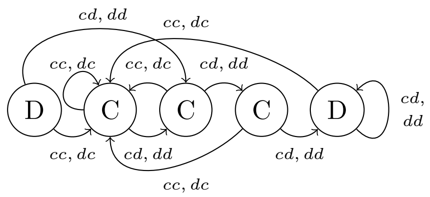
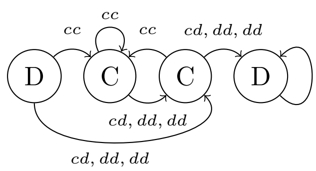
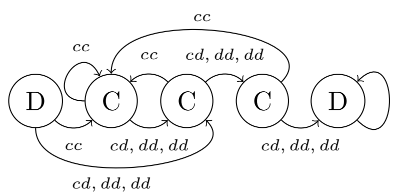
Notes: Circles represent the states of an automaton. The first state from the left is the start state. The labels *C* and *D* indicate whether the automaton prescribes cooperation or defection in the state. Arrows represent deterministic state transitions. The labels indicate the information profiles of the previous periods which trigger the transitions. An unlabeled arrows indicates an unconditional transition that occurs independent of the observed profile.

Table C2: Strategies 8-15

<table>
<tr><th>Acronym</th><th>Description</th><th>Automaton</th></tr>
<tr><td>T2</td><td>Play C until either player plays D, then play D twice and return to C (regardless of all actions during the punishment rounds).</td><td>
</td></tr>
<tr><td>TF2T</td><td>Play C unless partner played D in both of the last 2 rounds.</td><td>
</td></tr>
<tr><td>TF3T</td><td>Play C unless partner played D in all of the last 3 rounds.</td><td>
</td></tr>
<tr><td>T2FT</td><td>Play C unless partner played D in either of the last 2 rounds (2 rounds of punishment if partner plays D).</td><td>
</td></tr>
<tr><td>T2F2T</td><td>Play C unless partner played 2 consecutive Ds in the last 3 rounds (2 rounds of punishment if partner plays D twice in a row).</td><td>
</td></tr>
<tr><td>GRIM2</td><td>Play C until 2 consecutive rounds occur in which either player played D, then play D forever.</td><td>
</td></tr>
<tr><td>GRIM3</td><td>Play C until 3 consecutive rounds occur in which either player played D, then play D forever.</td><td>
</td></tr>
<tr><td>PT2FT</td><td>Play C if both players played C in the last 2 rounds, both players played D in the last 2 rounds, or both players played D 2 rounds ago and C last round. Otherwise play D.</td><td>
</td></tr>
</table>

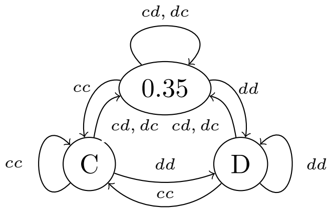
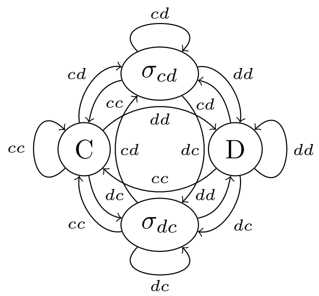
Notes: Circles represent the states of an automaton. The first state from the left is the start state. The labels *C* and *D* indicate whether the automaton prescribes cooperation or defection in the state. Arrows represent deterministic state transitions. The labels indicate the information profiles of the previous periods which trigger the transitions. An unlabeled arrows indicates an unconditional transition that occurs independent of the observed profile.

Table C3: Suspicious Strategies 16-20

Acronym	Description	Automaton
DTFT	Play D in the first round, then play TFT.	
DTF2T	Play D in the first round, then play TF2T.	
DTF3T	Play D in the first round, then play TF3T.	
DGRIM2	Play D in the first round, then play GRIM2.	
DGRIM3	Play D in the first round, then play GRIM3.	

Notes: Circles represent the states of an automaton. The first state from the left is the start state. The labels C and D indicate whether the automaton prescribes cooperation or defection in the state. Arrows represent deterministic state transitions. The labels indicate the information profiles of the previous periods which trigger the transitions. An unlabeled arrows indicates an unconditional transition that occurs independent of the observed profile.

Table C4: Behavior Strategies 21-23

Acronym	Description	Automaton
SGRIM	Play C if both players played C, and D if both players played D. If one player played D and the other C, play C with probability 0.35.	
M1BF$_{eq}$	Play C if both players played C, and D if both players played D. If the own action was C and the other player played D, play C with probability σ_{cd}. If the own action was D and the other player played C, play C with probability σ_{dc}. With perfect monitoring, $\sigma_{cd} = 0.75$ and $\sigma_{dc} = 0.5$ and the strategy conditions on actions. With imperfect monitoring, $\sigma_{cd} = 0.5$ and $\sigma_{dc} = 1$ and the strategy conditions on the own action and the received signal.	
RAND	Always randomize between C and D with $\sigma = 0.5$.	

Notes: Circles represent the states of an automaton. The first state from the left is the start state. The labels *C* and *D* indicate whether the automaton prescribes cooperation or defection in the state. The numbers in SGRIM indicate the probability of cooperation in the current state of the automaton. In the memory-one belief-free equilibrium strategy M1BF$_{eq}$, σ_{cd} and σ_{dc} are cooperation probabilities which depend on the monitoring structure. Arrows represent deterministic state transitions. The labels indicate the information profiles of the previous periods which trigger the transitions. An unlabeled arrows indicates an unconditional transition that occurs independent of the observed profile.

Appendix D Experimental Instructions and Quiz

[Below are the instructions for the perfect-monitoring treatment with repeated communication. Instructions for the other treatments were very similar and are therefore omitted here. They can be obtained from the authors upon request, along with the original instructions in German.]

Overview

Welcome to this experiment. We ask you not to speak with other participants during the experiment and to switch off your mobile phones and other mobile electronic devices.

For your participation in today's session, you will be paid in cash at the end of the experiment. The amount of the payout depends in part on your decisions, partly on the decisions of other participants and partly on chance. It is therefore important that you carefully read and understand the instructions before the start of the experiment.

In this experiment, every interaction between participants goes through the computers you are sitting in front of. You will interact with each other anonymously. Neither your name nor the names of other participants will be made public, either today or in future written evaluations.

Today's session includes several rounds. Your payout amount is the sum of the earned points in all rounds, converted into euros. The conversion of points into euros is done as follows. Each point is worth 2 cents, so the following applies: 50 points = EUR 1.00.

All participants will be paid privately, so that other participants will not be able to see how much they have earned.

Experiment

Interactions and Matching

This experiment comprises 7 identical interactions, each consisting of a randomly determined number of rounds.

At the very beginning, before the first interaction, you are randomly placed in a group with other participants. In each of the 7 interactions, you will interact with a different participant in your group.

In concrete terms, this is how it works: Before the first interaction, you are assigned to a person from your group with whom you interact in all rounds of the first interaction. In the second interaction, you are then assigned to a new person from your group, with whom you interact in all rounds of the second interaction, etc. In this way, you interact with each person assigned to your group in exactly one interaction, but in all

Table C5: Evolution of Strategy Choices in the First Three Supergames

		No			Pre-Play			Repeated		
		p_k	β_0	β_{sg}	p_k	β_0	β_{sg}	p_k	β_0	β_{sg}
Perfect	ALLD	0.75	-	-	0.25	-	-	-	-	-
		(0.04)			(0.04)					
	TFT	-	-	-	0.75	0.33	0.94	-	-	-
					(0.04)	(0.14)	(0.15)			
	T2F2T	0.25	-1.92	0.18	-	-	-	-	-	-
		(0.04)	(0.19)	(0.14)						
	DTF3T	-	-	-	-	-	-	0.10	-	-
								(0.03)		
	SGRIM	-	-	-	-	-	-	0.90	1.01	2.87
								(0.03)	(0.18)	(0.58)
	γ	0.16			0.07			0.09		
	ICL	446.51			310.15			316.86		
	ln$L$	-407.01			-284.15			-295.39		
Public	ALLD	0.88	-	-	0.33	-	-	0.09	-	-
		(0.03)			(0.05)			(0.03)		
	LGRIM3	0.12	-2.11	0.14	-	-	-	-	-	-
		(0.03)	(0.26)	(0.18)						
	T2F2T	-	-	-	-	-	-	0.53	0.91	0.96
								(0.05)	(0.22)	(0.18)
	M1BF$_{eq}$	-	-	-	0.67	0.21	0.54	-	-	-
					(0.05)	(0.17)	(0.13)			
	RAND	-	-	-	-	-	-	0.38	1.17	0.38
								(0.05)	(0.25)	(0.21)
	γ	0.17			0.16			0.03		
	ICL	402.17			488.87			421.42		
	ln$L$	-377.42			-444.55			-371.70		
Private	ALLD	0.71	-	-	0.37	-	-	0.18	-	-
		(0.05)			(0.04)			(0.04)		
	TF2T	-	-	-	0.63	-0.17	0.72	0.82	0.92	0.71
					(0.04)	(0.16)	(0.13)	(0.04)	(0.20)	(0.20)
	SGRIM	0.29	-1.12	0.24	-	-	-	-	-	-
		(0.05)	(0.19)	(0.14)						
	γ	0.11			0.14			0.18		
	ICL	367.31			412.54			441.74		
	ln$L$	-326.33			-380.98			-409.79		

Notes: The table reports maximum-likelihood estimates of strategy shares and latent-class regression coefficients for a candidate set of 23 strategies listed in Tables C1-C4 of Appendix C. Estimates are obtained assuming independent strategy choices over the first three supergames. Strategies condition on action profiles in perfect treatments, and on action-signal profiles in public and private treatments. γ indicates the probability of a tremble. The listed strategies persist after treatment-wise strategy selection based on the ICL information criterion. As ICL suggests a model with only one strategy (SGRIM) in the repeated communication treatment with perfect monitoring, we report the second best model with two strategies. Omitted shares (-) indicate that a strategy is not among the selected strategies of a treatment. Analytic standard errors in parentheses. Values might not add up as expected due to rounding.

Table C6: Memory-one Markov Strategies for the Last 3 Supergames

		Perfect						Public						Private					
		share	$\sigma_{\emptyset}$	σ_{cc}	σ_{cd}	σ_{dc}	σ_{dd}	share	$\sigma_{\emptyset}$	σ_{cc}	σ_{cd}	σ_{dc}	σ_{dd}	share	$\sigma_{\emptyset}$	σ_{cc}	σ_{cd}	σ_{dc}	σ_{dd}
No	s1	0.55	0.64	0.98	0.32	0.67	0.03	0.73	0.10	0.81	0.39	0.05	0.07	0.45	0.82	0.98	0.51	0.21	0.06
		(0.07)	(0.05)	(0.01)	(0.05)	(0.07)	(0.01)	(0.06)	(0.03)	(0.09)	(0.11)	(0.02)	(0.01)	(0.07)	(0.05)	(0.01)	(0.09)	(0.06)	(0.02)
	s2	0.32	0.00	no obs	0.00	0.00	0.02	0.27	0.74	0.94	0.47	0.94	0.20	0.45	0.03	1	0.22	0.01	0.01
		(0.07)	(0.00)	-	(0.00)	(0.00)	(0.01)	(0.06)	(0.09)	(0.02)	(0.07)	(0.08)	(0.06)	(0.07)	(0.03)	(0.00)	(0.17)	(0.05)	(0.01)
	s3	0.13	0.20	0.36	0.1	0.36	0.23	-	-	-	-	-	-	0.10	0.46	0.00	0.13	0.2	0.22
		(0.05)	(0.12)	(0.6)	(0.12)	(0.22)	(0.08)	-	-	-	-	-	-	(0.05)	(0.23)	(0.00)	(0.19)	(0.12)	(0.09)
	ICL	347.46						359.27						289.69					
	ln$L$	-306.82						-335.27						-251.56					
Pre-Play	s1	1	0.98	0.99	0.56	0.67	0.00	0.59	0.97	0.99	0.78	0.21	0.11	0.8	1	0.99	0.69	0.55	0.00
		(0.00)	(0.01)	(0.00)	(0.21)	(0.14)	(0.00)	(0.07)	(0.02)	(0.01)	(0.07)	(0.07)	(0.08)	(0.06)	(0.00)	(0.01)	(0.04)	(0.11)	(0.00)
	s2	-	-	-	-	-	-	0.28	0.91	0.7	0.53	0.66	0.26	0.16	0.92	0.70	0.30	0.76	0.24
		-	-	-	-	-	-	(0.07)	(0.05)	(0.04)	(0.10)	(0.13)	(0.09)	(0.05)	(0.09)	(0.12)	(0.09)	(0.19)	(0.08)
	s3	-	-	-	-	-	-	0.13	0.71	0.36	-	0.05	0.05	0.04	0.38	1	1	0.00	0.00
		-	-	-	-	-	-	(0.05)	(0.18)	(0.10)	(0.00)	(0.11)	(0.03)	(0.03)	(0.16)	(0.00)	(0.00)	(0.00)	(0.00)
	ICL	82.86						397.78						257.29					
	ln$L$	-70.89						-356.91						-218.81					
Repeated	s1	1	1	0.99	0.57	0.57	0.75	0.82	0.98	0.97	0.89	1	0.9	0.8	0.99	0.99	0.79	0.49	0.19
		(0.00)	(0.00)	(0.00)	(0.29)	(0.35)	(0.15)	(0.06)	(0.01)	(0.01)	(0.04)	(0.00)	(0.11)	(0.07)	(0.01)	(0.01)	(0.04)	(0.10)	(0.08)
	s2	-	-	-	-	-	-	0.18	0.92	0.8	0.54	0.34	0.13	0.2	0.93	0.83	0.54	0.82	0.69
		-	-	-	-	-	-	(0.06)	(0.08)	(0.05)	(0.14)	(0.18)	(0.10)	(0.07)	(0.10)	(0.06)	(0.20)	(0.15)	(0.36)
	s3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
		-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	ICL	66.72						243.21						257.76					
	ln$L$	-54.76						-216.49						-227.03					

Notes: The reported results summarize the behavior in the last three supergames across the nine experimental treatments. The behavior in each treatment is characterized based on memory-one Markov strategies. The number of strategies is selected based on ICL. The reported values indicate the probability of cooperation after the five possible memory-one histories $\emptyset$, cc, cd, dc, and dd. Bootstrapped standard errors in parentheses (10000 repetitions).

Table C7: Inference of Pure Strategies for the Last 3 Supergames

	Perfect			Public			Private		
$(\sigma_{\emptyset},\sigma_{cc},\sigma_{cd},\sigma_{dc},\sigma_{dd})$	No	Pre	Rep	No	Pre	Rep	No	Pre	Rep
(0,1,0,0,0)	0.64	-	-	0.7	-	-	0.61	-	-
	(0.07)			(0.07)			(0.08)		
(1,0,0,0,0)	-	-	-	-	-	0.04	-	-	-
						(0.03)			
(1,1,0,0,0)	-	0.4	-	-	0.38	-	-	0.26	-
		(0.16)			(0.08)			(0.07)	
(1,1,0,1,0)	0.36	-	-	0.17	-	-	-	-	0.14
	(0.07)			(0.06)					(0.06)
(1,1,1,0,0)	-	-	0.37	0.13	-	0.24	0.39	-	0.22
			(0.17)	(0.05)		(0.1)	(0.08)		(0.10)
(1,1,1,1,0)	-	0.6	-	-	0.62	-	-	0.61	-
		(0.16)			(0.08)			(0.08)	
(1,1,0,1,1)	-	-	0.63	-	-	-	-	0.13	-
			(0.17)					(0.06)	
(1,1,1,1,1)	-	-	-	-	-	0.72	-	-	0.63
						(0.1)			(0.11)
γ	0.10	0.01	0.01	0.11	0.13	0.06	0.10	0.07	0.07
ICL	393.27	132.57	122.88	390.36	449.89	296.34	353.65	317.54	321.97
$\ln L$	-362.56	-76.66	-66.33	-349.67	-415.48	-241.52	-323.34	-267.94	-263.38

Notes: The reported results summarize the behavior in the last three supergames across the nine experimental treatments. The behavior in each treatment is characterized based on memory-one Markov strategies with pure strategy parameters. The number of strategies is selected based on ICL. The reported values indicate the probability of cooperation after the five possible memory-one histories $\emptyset$, cc, cd, dc, and dd. Analytic standard errors in parentheses.

Table C8: Classification of Strategies (Camera et al., 2012)

	Perfect			Public			Private		
	No	Pre	Rep	No	Pre	Rep	No	Pre	Rep
ALLD	71	4	0	84	7	1	71	4	1
ALLC	19	152	160	17	101	128	32	116	127
GRIM	39	149	154	14	49	58	34	80	73
TFT	41	151	156	17	66	87	30	89	83
PTFT	22	147	156	10	44	57	26	74	74
T2	22	147	154	6	45	58	27	77	73
TF2T	27	153	160	16	99	125	39	118	122
TF3T	21	152	160	14	98	126	34	116	125
T2F1T	37	150	156	15	70	84	30	87	75
T2F2T	27	153	160	16	102	126	38	118	124
LGRIM2	27	152	161	15	93	119	42	116	121
LGRIM3	22	152	160	14	101	126	33	115	125
PTFT2	23	147	156	7	43	57	27	76	72
FC	24	1	-	12	16	3	14	3	2
DTFT	51	4	-	34	2	2	34	-	3
DTF2T	13	-	-	13	1	2	6	-	1
DTF3T	14	-	-	8	1	2	4	-	1
DLGRIM2	13	-	-	11	1	2	6	-	1
DLGRIM3	17	-	-	10	1	2	5	-	1
DC	5	-	-	2	4	1	3	2	1

Notes: The reproted results characterize subjects' behavior in the last three supergames based on the strategy classification method proposed by Camera et al. (2012). The values reflect the number of supergames in which the behavior of one player is accurately predicted by a strategy. If a supergame has x rounds and y actions are observed that are not predicted by the pure strategy, the strategy predicts the behavior accurately if the frequency of errors y/x is smaller or equal to 5 percent.

rounds of that interaction.

Length of an Interaction

The length of an interaction is determined randomly. After each round there is an 80% chance that there will be at least one more round.

You can imagine this as follows. A 100-sided dice is rolled after each round. If the roll is 20 or less, there is no further round. If the roll is a different number (21-100), the interaction continues. Note that the probability of another round does not depend on the round you are in. The probability of a third round when you are in round 2 is 80%, as is the probability of a tenth round when you are in round 9.

As soon as chance decides after a round that there is no further round in the interaction, the interaction is finished and you are assigned to a new person for the next interaction. After the seventh interaction, the experiment ends.

Interactions and Sequence of Events in a Round

Before each round of interaction, you can chat with the other person on your screen. The chat takes place in an anonymous chat window. In order to protect your anonymity, it is important that you do not provide any information about yourself or your seat number during communication. Otherwise we reserve the right not to pay you any money in the end. The entire chat content is displayed during the interaction and can be read again.

After the first chat the first round begins.

In each round, you select one of two possible options, A or B. The other person also selects one of two possible options, A or B.

There is a 90% probability that the option you have chosen will be correctly communicated to the other person. There is a 10% probability that the option you have not selected will be transmitted. What the other person receives is what we call the other person's signal. The same applies to the other person's option and your signal. For example, if the other person chooses option A, you receive Signal A with 90% probability and with 10% probability you get Signal B. Assuming you choose Option B, the other person receives Signal A with 10% probability and Signal B with 90% probability.

Your round income depends on your selected option and the signal received. Likewise, the payout of the other person depends on their chosen option and the signal they receive.

Once you and the other person have chosen an option, chance decides which signals are transmitted and what round earnings result from them with the probabilities given above.

Figure D1: Round Income [Figure 1 from Instructions]

Ihre Optionen Your options	Ihr Einkommen bei Signal Your income with signal		Erwartetes Einkommen, wenn die andere Person Expected income if the other person	
	A	B	Option A wählt chooses option A	Option B wählt chooses option B
Option A	30	0	27	3
Option B	37	17	35	19

The four cells on the right in Figure 1 show the expected earnings depending on your option choice and the option choice of the other person. For example, if you select option B and the other person selects option A, you receive Signal A with 90% probability and Signal B with 10%. Therefore you will receive 37 points with 90% probability and 17 points with 10% probability, that is, your expected earnings in this case are: 0.9*37+0.1*17=35 points.

Figure D2: Part of Feedback Screen (Example) [Figure 2 from Instructions]

Rundeneinkommen
Round Income

Your Choice:	Ihre Wahl:	Option A
Your Signal:	Ihr Signal:	B
Choice of oth. person:	Wahl der anderen Person:	Option B
Signal of oth. person:	Signal der anderen Person:	A
Your Points in this Round:	Ihre Punkte aus dieser Runde:	0

At the end of the round, you will receive feedback on your chosen option, the signal received, the other person's choice of an option, the signal received by the other person, and your own round earnings (see Figure 2).

All possible following rounds are identical in sequence. The course of the current interaction, that is, the feedback that you received at the end of all previous rounds, is shown in a table in every round.

End and Payoff

As soon as chance ends the seventh interaction, the experiment is over.

At the end of the experiment, the points from all rounds are converted into euros and paid out privately.

The last screen of the last round of the seventh interaction shows you how much you have earned in euros.

Questions?

Take your time to go over the instructions again. If you have any questions, please raise your hand. An experimenter will then come to your place.

If you think you have understood everything well, you can start the quiz on your screen. The quiz is only to ensure that everyone has understood the instructions well. The answers do not affect your payout.

Quiz [on screen]

[The quiz was the same in all nine treatments. The correct answers appeared on the next screen.]

1. How many interactions are there?

   [1,7, it is random]

2. What is the probability that there is a first round of an interaction?

   [20%, 80%, 100%]

3. What is the probability that there will be a second round in an interaction when you are currently in the first?

   [20%, 80%, 100%]

4. What is the probability that there will be a third round in an interaction when you are currently in the second?

   [20%, 80%, 100%]

5. What is the probability that there will be a third round in an interaction when you are currently in the first?

   [64%, 80%, 100%]

6. You choose Option B and the other person cooses Option B.

   (a) What is the probability that you receive Signal A?

   [10%, 90%, 100%]

   (b) What is the probability that the other person receives Signal B?

   [10%, 90%, 100%]

   (c) How high is your payoff in case you receive Signal A?

   [19, 35, 37]

   (d) How high is the expected payoff of the other person?

   [19, 35, 37]