Online Appendix

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

American Economic Journal: Microeconomics

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

2024

Appendix A Theoretical Appendix

In A.1, we present an extensions of the BAD to the two 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

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

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.

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

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

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. 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:*

$$\delta \geq \delta_{aa}^{BF} = \frac{\phi}{1+\phi} \tag{3}$$

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

$$\sigma_{cd} = \sigma_{cc} + \left(\sigma_{cc} - \sigma_{dd} - \frac{1}{\delta}\right) g \tag{4}$$

and

$$\sigma_{dc} = \sigma_{dd} - \left(\sigma_{cc} - \sigma_{dd} - \frac{1}{\delta}\right) l \tag{5}$$

(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:

$$V_{aa}^c = (1-\delta)(\sigma_{aa} - (1-\sigma_{aa})l) + \delta(\sigma_{aa}V_{cc} + (1-\sigma_{aa})V_{dc}) \tag{6}$$
$$V_{aa}^d = (1-\delta)(\sigma_{aa}(1+g) + (1-\sigma_{aa})0) + \delta(\sigma_{aa}V_{cd} + (1-\sigma_{aa})V_{dd}) \tag{7}$$

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 (7) from (6) 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 (6) 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 (6) 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 (6) solved for σ_{dd} and σ_{dc} this yields (4) and (5) 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 (4) 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 (5) and rearranging gives $\delta > \delta_{aa}^{BF}$ 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 (3) together with $\sigma_{cc} = 1$ and $\sigma_{dd} = 0$ into (4) and (5) to obtain the structure of the T1BF response defined by g and l. □

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:*

$$\delta \geq \delta_{aa}^{SG} = \frac{g+l}{1+g+l} \tag{8}$$

(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:*

$$\sigma_{dd} = \sigma_{cc} - \frac{g+l}{\delta(1+g+l)} \tag{9}$$

and

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

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

Proof of Proposition 1.1.2. Using (4) and (5) yields (9) and (10). 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 $\delta \geq g/(1+g+l)$. □

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 (3) 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:*

$$\delta \geq \delta_{ss}^{BF} = \frac{(1-\epsilon)\phi - \epsilon\psi}{(1-2\epsilon)(1-2\epsilon+(1-\epsilon)\phi-\epsilon\psi)} \tag{11}$$

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

$$\sigma_{cd} = \sigma_{cc} + \frac{\sigma_{cc} - \sigma_{dd} - \frac{1}{\delta(1-2\epsilon)}}{1-2\epsilon}((1-\epsilon)g - \epsilon l) \tag{12}$$

and

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

(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:

$$\begin{aligned} V_{ss}^c = (1-\delta)(\sigma_{ss} - (1-\sigma_{ss})l) + \delta[&(1-\epsilon)(\sigma_{ss}(1-\epsilon) + (1-\sigma_{ss})\epsilon)V_{cc} \\ &+\epsilon(\sigma_{ss}(1-\epsilon) + (1-\sigma_{ss})\epsilon)V_{cd} \\ &+(1-\epsilon)(\sigma_{ss}\epsilon + (1-\sigma_{ss})(1-\epsilon))V_{dc} \\ &+\epsilon(\sigma_{ss}\epsilon + (1-\sigma_{ss})(1-\epsilon))V_{dd}] \end{aligned} \tag{14}$$

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

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 (15) from (14) 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 (14) 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 (14) 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 (14) 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 (11). 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 (11) together with $\sigma_{cc} = 1$ and $\sigma_{dd} = 0$ into (12) and (13) to obtain the structure of the T1BF response defined by g and l. $\square$

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:*

$$\delta \geq \delta_{ss}^{SG} = \frac{g+l}{(1-2\epsilon)(1+g+l)} \tag{16}$$

(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:*

$$\sigma_{dd} = \sigma_{cc} - \frac{g+l}{\delta(1-2\epsilon)(1+g+l)} \tag{17}$$

and

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

(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 (12) and (13) yields (17) and (18). 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 (16). $\square$

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:*

$$\delta \geq \delta_{as}^{BF} = \frac{\phi}{1-2\epsilon+(1-\epsilon)\phi-\epsilon\psi} \tag{19}$$

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*

$$\sigma_{cd} = \sigma_{cc} + \frac{\sigma_{cc} - \sigma_{dd} - \frac{1}{\delta}}{1 - 2\epsilon - \epsilon(g+l)} g \tag{20}$$

and

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

(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 $V^{a_i}_{a_j s_i}$ 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:

$$\begin{aligned} V^c_{as} &= (1-\delta)(\sigma_{as} - (1-\sigma_{as})l) + \\ \delta\left((1-\epsilon)\sigma_{as}V_{cc} + \epsilon\sigma_{as}V_{cd} + (1-\epsilon)(1-\sigma_{as})V_{dc} + \epsilon(1-\sigma_{as})V_{dd}\right) \end{aligned} \tag{22}$$

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

Subtracting (23) from (22) gives:

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

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 (22) 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 (22) 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 (22) 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 (19).

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 (19) 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 (19) also requires $\delta \geq 1$. This proofs (i). To complete the proof, insert (19) together with $\sigma_{cc} = 1$ and $\sigma_{dd} = 0$ into (20) and (21) 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:*

$$\delta \geq \delta_{as}^{SG} = \frac{g+l}{1-2\epsilon+(1-\epsilon)(g+l)} \tag{24}$$

(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:*

$$\sigma_{dd} = \sigma_{cc} - \frac{g+l}{\delta(1-2\epsilon+(1-\epsilon)(g+l))} \tag{25}$$

and

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

(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 (20) and (21) yields (25) and (26). 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 (24). □

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:

$$V_{pp} \geq V_r \tag{27}$$

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:

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

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:

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

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

With our experimental parameters, the condition $V_{comply}^{punished} \geq V_{deviate}^{punished}$ 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:

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

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:

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

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

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

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

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

With our parameters, $V^{punished}_{comply} \geq V^{punished}_{deviate}$ 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

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	d signal	c signals	diff	p-value	d signal	c signal	diff	p-value
Coordination	0.45	0.29	0.17	0.01	0.48	0.29	0.19	0.01
Deliberation	0.12	0.13	-0.01	0.85	0.08	0.09	-0.01	0.85
Relationship	0.26	0.40	-0.14	0.03	0.24	0.32	-0.08	0.26
Information	0.66	0.34	0.33	0.00	0.64	0.34	0.29	0.00
Trivia	0.38	0.53	-0.15	0.00	0.39	0.54	-0.15	0.04
Report of action	0.41	0.02	0.39	0.00	0.44	0.09	0.35	0.00
Report of C	0.40	0.02	0.38	0.00	0.44	0.09	0.35	0.00
Report of D	-	-	-	-	-	-	-	-
Report of signal	0.56	0.33	0.23	0.00	0.64	0.33	0.31	0.00
Report of c	0.09	0.33	-0.24	0.00	0.01	0.32	-0.31	0.00
Report of d	0.48	-	-	-	0.64	0.00	0.63	0.00

Notes: Frequency of communication categories for subject-round observations with cooperative history of both players up to round t. A participant has a cooperative history if all her previous actions were C and all signals she observed in rounds $< t$ were c. Columns compare the communication in round $t+1$ conditional on the signals received in round t. Frequencies indicate the probability that both raters indicated the category for a text unit. P-values derived from logit models with standard errors clustered on participant and match. Zero frequencies omitted (-).

Appendix C Strategy Estimation

We use the strategy frequency estimation method (Dal Bó and Fréchette, 2011) and its adaptation to behavior strategies (Breitmoser, 2015) to analyze participants' strategies across treatments. The estimation is perfomed with the R package *stratEst* (Dvorak, 2023).

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.[1] 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

[1]See Bland (2020) for a recent adaptation of SFEM which allows for heterogeneity in the trembles.

Table B5: Communication after First Defection Signal – All Supergames

		Public Repeated			Private Repeated		
#	Subcategory	*d* signal	*c* signals	diff	*d* signal	*c* signal	diff
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	*d* signal	*c* signals	diff	*d* signal	*c* signal	diff
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.

Table B7: Frequency and Truthfulness of Private Information Exchange - All Supergames

	Public		Private	
	p(report)	p(true)	p(report)	p(true)
Actions				
Report of action	0.11	0.94	0.14	0.89
Report of C	0.09	0.95	0.14	0.88
Report of D	0.02	0.97	0.01	1.00
Report of C if $\omega_i = d$	0.11	0.83	0.14	0.61
D and report of D if $\omega_i = d$	0.12	1.00	0.03	1.00
C and report of C $\omega_i = d$	0.30	1.00	0.30	1.00
D and report of C if $\omega_i = d$	0.03	0.00	0.08	0.00
Signals				
Report of signal	-	-	0.33	0.95
Report of c	-	-	0.23	0.98
Report of d	-	-	0.10	0.86
Report of d if $\omega_{-i} = d$	-	-	0.33	-

Notes: Frequencies of coding in all participant-round observations after round one for the repeated communication treatments with public monitoring (columns 2 and 3) and private monitoring (columns 4 and 5). 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 B8: Private Information Exchange and Mutual Cooperation - All Supergames

	Public			Private		
	estimate	std. error	p-value	estimate	std. error	p-value
intercept	-0.14	0.23	0.55	-0.76	0.32	0.02
Report of C	0.65	0.36	0.07	2.42	1.16	0.04
Report of d	-	-	-	1.42	0.41	0.00
Report of C $\times$ Report of d	-	-	-	-2.12	1.15	0.06
Trivia	0.76	0.30	0.01	-0.12	0.32	0.70

Notes: Table shows coefficients of logit models with standard errors clustered on participant and match. Report of C is a dummy that indicates if C is reported by the player for whom the signal indicated d in the last round. Report of d is a dummy that indicates whether the defection signal was reported by the player who received the signal. Data of all supergames. A coding is considered valid if both raters indicated the same sub-category for a participant-round observation.

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 Table C1, we perform treatment-wise strategy estimation starting with the candidate set of 24 strategies listed in Tables C2-C5. 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.

SFEM Results

Table C1 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.

Adaptation of Strategies

Tables C2-C5 list the set of 24 strategies used to obtain the strategy estimation results reported in Table C1. Circles in Table C4 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 four strategies are behavior strategies. Two of the behavior strategies are motivated

Table C1: Strategy Frequency Estimation

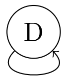
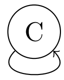
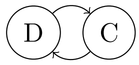
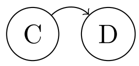
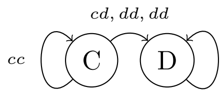
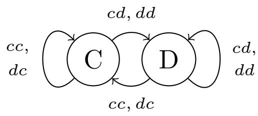
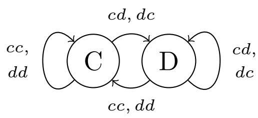
		Perfect			Public			Private		
	lenient/forgiving	No	Pre	Rep	No	Pre	Rep	No	Pre	Rep
ALLD	no	0.42	-	-	0.61	0.02	-	0.50	0.02	-
		(0.07)	-	-	(0.08)	(0.02)	-	(0.07)	(0.02)	-
ALLC	yes	-	-	-	-	-	0.32	-	-	0.27
		-	-	-	-	-	(0.19)	-	-	(0.20)
DC	no	-	-	-	-	-	-	-	-	-
		-	-	-	-	-	-	-	-	-
FC	no	-	-	-	-	0.08	0.01	-	-	-
		-	-	-	-	(0.04)	(0.02)	-	-	-
GRIM	no	0.08	0.23	-	-	0.02	-	0.03	-	-
		(0.06)	(0.16)	-	-	(0.02)	-	(0.04)	-	-
TFT	yes	0.08	-	-	0.03	-	-	-	-	-
		(0.06)	-	-	(0.04)	-	-	-	-	-
PTFT	yes	-	-	0.17	-	-	-	-	-	-
		-	-	(0.18)	-	-	-	-	-	-
T2	yes	-	-	-	-	-	-	-	-	-
		-	-	-	-	-	-	-	-	-
TF2T	yes	-	-	-	0.01	0.01	-	-	0.07	0.07
		-	-	-	(0.02)	(0.04)	-	-	(0.09)	(0.08)
TF3T	yes	-	-	-	-	-	-	-	-	-
		-	-	-	-	-	-	-	-	-
T2FT	yes	-	-	-	-	0.04	-	-	0.05	-
		-	-	-	-	(0.04)	-	-	(0.06)	-
T2F2T	yes	0.04	-	0.40	-	0.15	-	-	0.24	0.09
		(0.03)	-	(0.21)	-	(0.10)	-	-	(0.15)	(0.10)
GRIM2	yes	-	-	0.44	-	0.20	0.21	0.19	0.10	0.15
		-	-	(0.20)	-	(0.10)	(0.15)	(0.07)	(0.14)	(0.09)
GRIM3	yes	-	-	-	0.04	0.02	0.32	-	0.01	0.12
		-	-	-	(0.03)	(0.04)	(0.18)	-	(0.06)	(0.17)
PT2FT	yes	-	-	-	-	-	-	-	-	-
		-	-	-	-	-	-	-	-	-
DTFT	yes	0.12	-	-	-	-	-	-	-	-
		(0.06)	-	-	-	-	-	-	-	-
DTF2T	yes	0.02	-	-	0.07	-	-	-	-	-
		(0.02)	-	-	(0.04)	-	-	-	-	-
DTF3T	yes	-	-	-	-	-	-	-	-	-
		-	-	-	-	-	-	-	-	-
DGRIM2	yes	-	-	-	0.02	-	-	0.01	-	-
		-	-	-	(0.04)	-	-	(0.02)	-	-
DGRIM3	yes	-	-	-	-	-	-	-	-	-
		-	-	-	-	-	-	-	-	-
SGRIM	yes	0.09	-	-	0.09	-	0.04	0.24	0.22	0.05
		(0.08)	-	-	(0.06)	-	(0.04)	(0.09)	(0.11)	(0.05)
M1BF	yes	-	-	-	0.03	0.38	-	-	0.10	0.08
		-	-	-	(0.05)	(0.10)	-	-	(0.09)	(0.09)
T1BF$_{as}$	yes	0.11	0.77	-	0.05	-	0.06	-	0.07	0.13
		(0.07)	(0.28)	-	(0.05)	-	(0.05)	-	(0.08)	(0.09)
RAND	no	0.03	-	-	0.05	0.08	0.05	0.03	0.11	0.03
		(0.03)	-	-	(0.03)	(0.05)	(0.04)	(0.04)	(0.05)	(0.03)
$\sum$ lenient/forgiving		0.46	0.77	1.00	0.34	0.79	0.94	0.45	0.87	0.97
		(0.09)	(0.16)	(0.01)	(0.08)	(0.06)	(0.04)	(0.08)	(0.06)	(0.04)
γ		0.06	0.01	0.01	0.07	0.06	0.03	0.05	0.02	0.04
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Notes: Treatment-wise maximum-likelihood shares of the 24 strategies listed in Tables C2-C5 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. Zero shares are omitted (-). Analytic standard errors in parentheses. Values might not add up as expected due to rounding.

by Backhaus and Breitmoser's (2021) 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 (2021) 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.

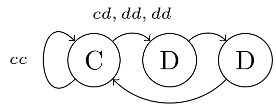
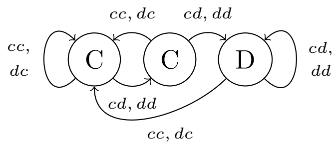
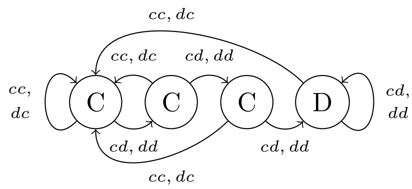
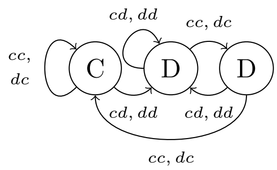
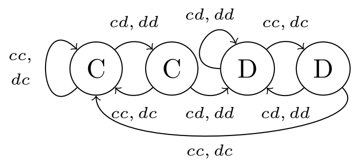
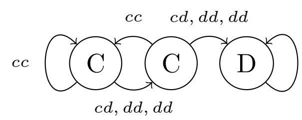
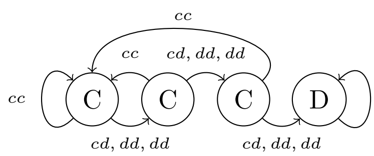
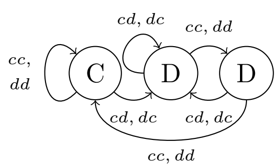
The third behavioral strategy that we include is a M1BF strategy that conditions on the observed actions ($\sigma_\emptyset = 1, \sigma_{cc} = 1, \sigma_{cd} = 0.75, \sigma_{dc} = 0.5, \sigma_{dd} = 0$). The M1BF strategy results for $\delta = 0.8$ when assuming that subjects start with cooperation, cooperate after mutual cooperation, and defect after mutual defection. The fourth behavior strategy that we include is the T1BF strategy that which conditions on the own action and the signal about the action of the partner in the previous round ($\sigma_\emptyset = 1, \sigma_{cc} = 1, \sigma_{cd} = 0.5, \sigma_{dc} = 1, \sigma_{dd} = 0$). The behavior of T1BF$_{as}$ after round one is the unique behavior of all memory-one belief-free equilibrium strategies that can be played under imperfect monitoring (see Appendix A for the derivation of these equilibrium strategies).

Table C2: 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.	

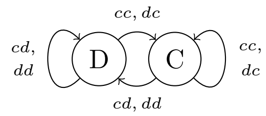
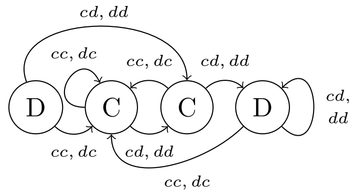
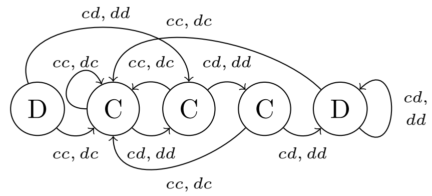
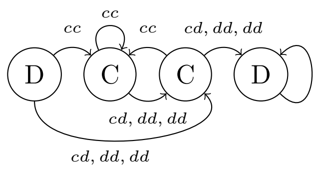
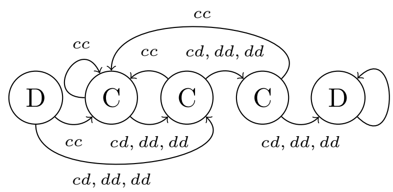
Notes: Circles represent the states of an automaton. The first state from the left is the start state. The labels C and D indicate whether the automaton prescribes cooperation or defection in the state. Arrows represent deterministic state transitions. The labels indicate the information profiles of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

Table C3: Strategies 8-15

Acronym	Description	Automaton
T2	Play C until either player plays D, then play D twice and return to C (regardless of all actions during the punishment rounds).	
TF2T	Play C unless partner played D in both of the last 2 rounds.	
TF3T	Play C unless partner played D in all of the last 3 rounds.	
T2FT	Play C unless partner played D in either of the last 2 rounds (2 rounds of punishment if partner plays D).	
T2F2T	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).	
GRIM2	Play C until 2 consecutive rounds occur in which either player played D, then play D forever.	
GRIM3	Play C until 3 consecutive rounds occur in which either player played D, then play D forever.	
PT2FT	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.	

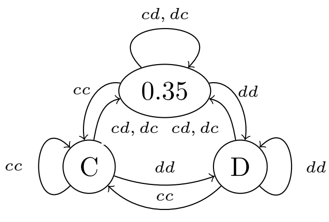
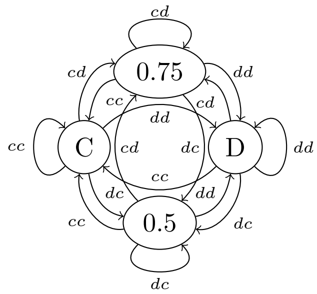
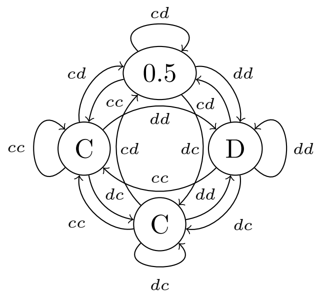
Notes: Circles represent the states of an automaton. The first state from the left is the start state. The labels C and D indicate whether the automaton prescribes cooperation or defection in the state. Arrows represent deterministic state transitions. The labels indicate the information profiles of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

Table C4: 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 C5: Behavior Strategies 21-24

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	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 0.75. If the own action was D and the other player played C, play C with probability 0.5.	
$T1BF_{as}$	Play C if you played C and the signal was c, and D if you played D and the signal was d. If the own action was C and the signal was d, play C with probability 0.5. If the own action was D and the signal was c, play C with probability 1.	
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. Numbers in indicate the probability of cooperation in the current state of the automaton. 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 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.

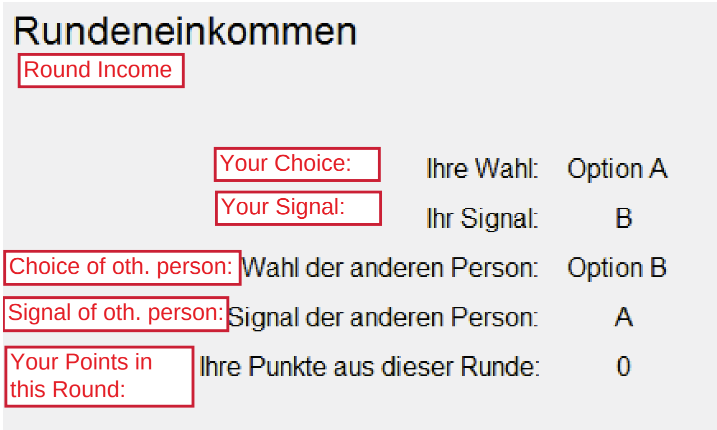
Once you and the other person have chosen an option, chance decides which signals are transmitted and what round earnings result from them with the probabilities given above.

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]



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]

References

Backhaus, T. and Y. Breitmoser (2021): "Inequity aversion and limited foresight in the repeated prisoner's dilemma," *Center for Mathematical Economics Working Papers*, 652.

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

Bland, J. R. (2020): "Heterogeneous trembles and model selection in the strategy frequency estimation method," *Journal of the Economic Science Association*, 6, 113–124.

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.

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

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

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. (2023): "stratEst: a software package for strategy frequency estimation," *Journal of the Economic Science Association*, 9, 337–349.

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.

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.

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