Payoffs, Beliefs, and Cooperation in Infinitely Repeated Games*

Maximilian Andres† Lisa Bruttel‡ Juri Nithammer§

November 12, 2024

Abstract

Disentangling the determinants of cooperation in social dilemmas remains paramount in economics. This paper shows theoretically that a player's belief about their opponent's probability of cooperation determines the effect of payoff changes on cooperation in the infinitely repeated prisoner's dilemma: While increasing the gain from unilateral defection strongly decreases cooperation iff beliefs are high, increasing the loss from unilateral cooperation strongly decreases cooperation iff beliefs are low. Otherwise, the effects of payoff changes are negligible. Two empirical tests support the theory: A meta study based on existing experimental data and a new laboratory experiment, where we vary beliefs exogenously.

JEL-codes: C72, C73, D81, D83
Keywords: infinitely repeated prisoner's dilemma, equilibrium selection, beliefs, payoffs, cooperation

*We thank John Duffy, Sebastian Fehrler, Urs Fischbacher, Friedericke Fromme, Paul Heidhues, Mats Köster, Friederike Mengel, Hans-Theo Normann, Catherine Roux, Vasilisa Werner, seminar participants from Basel, Potsdam, and the Bergen Berlin Behavioral Economics Workshop as well as participants at ESA Lyon 2023, ESA Exeter 2023, EARIE Amsterdam 2024, VfS Berlin 2024, and GfeW Köln 2024 for very helpful comments and suggestions. Anna Nold, Birte Prado-Brand, and Jonas Voigt provided excellent research assistance. The experimental design, number of subjects, point predictions and hypotheses, as well as a pre-analysis plan for the laboratory experiment in Section 5, were preregistered in the AEA RCT registry (`https://doi.org/10.1257/rct.12816-2.0`). This research is funded by the Volkswagen Foundation, which we gratefully acknowledge.

†Universität Potsdam and Berlin School of Economics. maximilian.andres@uni-potsdam.de
‡Universität Potsdam and Berlin School of Economics. lisa.bruttel@uni-potsdam.de
§Universität Potsdam. juri.nithammer@uni-potsdam.de

1 Introduction

How can we intervene in social dilemmas to promote cooperation? Answering this question is key for the social sciences because social dilemmas capture a tension between socially efficient cooperation and individually optimal defection that is at the heart of many strategic interactions.

According to an influential body of literature in economics, the answer to this question is to decrease the gain from unilateral defection or the loss from unilateral cooperation (Blonski et al., 2011; Dal Bó and Fréchette, 2011, 2018; Embrey et al., 2018; Mengel, 2018; Baader et al., 2024; Gächter et al., 2024). Yet, to design effective and efficient interventions to promote cooperation, we also need to disentangle when which payoff changes are decisive for cooperation, and why.

This paper shows that the effects of changes in the payoffs depend systematically on a player's belief about the probability of cooperation by their opponent. If a player's beliefs about their opponent's probability to cooperate is *high*, increasing the gain from unilateral defection has a large negative effect on cooperation, while increasing the loss from unilateral cooperation has only a negligible effect. However, if a player's belief is *low*, increasing the gain has a negligible effect, while increasing the loss has a large effect. Consequently, players with low beliefs react more strongly to changes in the loss than to the changes in the gain, and players with high beliefs react more strongly to the changes in gain than to changes in the loss. Thus, we show that beliefs are even more important than recently shown (see, e.g., Aoyagi et al., 2024; Andres, 2024; Boczoń et al., 2024; Dvorak and Fehrler, 2024; Gill and Rosokha, 2024; Huck et al., 2024; Martinez-Martinez and Normann, 2024) because they determine if, how, and how much changes in the payoffs of a social dilemma affect cooperation. With this, we add a comprehensive understanding to the economic literature of how we can promote or deter cooperation in social dilemmas through changes in the payoff parameters, which many policy interventions aim at.

To show this, we provide a formal proposition, based on the theory of equilibrium selection in the infinitely repeated prisoner's dilemma, where both cooperation and defection might be equilibria (Fudenberg and Maskin, 1986). Building on the seminal contributions of Blonski et al. (2011) and Dal Bó and Fréchette (2011), we introduce a novel critical discount factor. This critical discount factor captures both the incentive structure of the game, i.e. the payoffs from cooperation and defection, and a player's belief about the probability of cooperation by their opponent. Thus, we are able to show that the effects of changes in the payoffs on cooperation depend on players' beliefs.

We provide two empirical tests of our proposition. First, we use the data provided by Dal Bó and Fréchette (2018) and Fudenberg and Karreskog Rehbinder (2024) to conduct a meta study with existing experimental data. Because beliefs were not elicited in the experimental studies collected in the meta data, we construct participants' beliefs from their experiences. For this purpose, we model belief formation as a Markov process that weighs the prior belief with the cooperation experience in each period of a supergame. We find that the negative effect of the gain on cooperation increases in the belief, while the negative effect of the loss decreases in the belief. This finding is robust to variations in the weights in the belief construction. Second, we conduct a new laboratory experiment in which we vary beliefs exogenously by varying the probabilities of being matched to a player that cooperates. This procedure ensures that beliefs are truly exogenous to the participants' underlying preferences and variation in the payoff parameters. In the experiment, we directly elicit participants' critical discount factors across different stage game parameterizations, varying the gain and the loss. We find a large negative effect of the belief on the effect of the loss on the critical discount factor, and a relatively small positive effect of the belief on the effect of the gain. In both the meta study and the lab experiment, we find that participants react more strongly to the loss than to the gain if their belief is low and more strongly to the gain than to the loss if their belief is high. Thus, both empirical approaches provide strong support for the theory.

The remainder of this paper is organized as follows. We begin by introducing the theoretical setup in Section 2, which includes the setup of the stage game, the infinite repetition, and the critical discount factor as a function of the payoffs and the belief. We present our main results in Section 3 by demonstrating how the critical discount factor behaves in the payoffs conditional on the belief. We then continue with our two empirical tests. In Section 4, we present the methodology of constructing beliefs from experiences, validation against previous findings, and results of the meta study. In Section 5, we present our experimental design, hypotheses, and results. Section 6 concludes.

2 The Infinitely Repeated Prisoner's Dilemma

This section first describes the stage game and its parameterization in detail. We will then continue with the infinite repetition of the game and equilibrium selection criteria discussed in the literature so far.

Stage Game The standard prisoner's dilemma $\Gamma(T, R, P, S)$ is a symmetric game of two players $i \in \{X, Y\}$ that face the same choice of action a simultaneously: To cooperate or to defect, $a_i \in \{C, D\}$. The left panel of Table 1 summarizes the payoffs in the game. If both players cooperate, both receive the reward payoff R. If both defect, both receive the punishment payoff P. If one player defects unilaterally, they receive the temptation payoff T. The player who cooperates unilaterally, while the other player defects, receives the sucker's payoff S. In the prisoner's dilemma Γ, two conditions must be met (Rapoport et al., 1965): First, $T > R > P > S$ ensures that mutual cooperation is Pareto-superior to mutual defection ($2R > 2P$), and that there is an incentive to defect because the individual payoff from unilateral defection is greater than from mutual cooperation ($T > R$), and the individual payoff from mutual defection is larger than from unilateral cooperation ($P > S$). Second, $2R > T + S$ ensures that mutual cooperation is also Pareto-superior to the asymmetric outcome of unilateral defection.

Following Stahl II (1991), we normalize the payoffs to reduce the prisoner's dilemma to the game $\Gamma(g, l)$ of the gain from unilateral defection g and the loss from unilateral cooperation $-l$, with $g, l > 0$. The normalization subtracts the punishment payoff P from the original payoff R, S, T or P and then divides by $R - P$. The normalized temptation payoff is $\frac{T-P}{R-P} = 1 + g$, and the normalized sucker's payoff is $\frac{S-P}{R-P} = -l$, while the normalized reward and punishment payoffs are $\frac{R-P}{R-P} = 1$ and $\frac{P-P}{R-P} = 0$, respectively. Thus, in the normalized version of the game, $T > R > P > S$ holds because $1 + g > 1 > 0 > -l$, and $2 \cdot R > T + S$ holds as long as $2 > 1 + g - l \Leftrightarrow l > g - 1$. The right panel of Table 1 presents the resulting payoffs after normalization.

Table 1: Stage Game Payoffs in the Prisoner's Dilemma Γ

	C	D
C	R, R	S, T
D	T, S	P, P

(a) Original

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

(b) Normalized

Infinite Repetition In the infinitely repeated prisoner's dilemma, players discount future payoffs by a factor of δ, where $0 < \delta < 1$. This understanding is equivalent to a setup with an indefinite number of repetitions, where δ is the probability that the game will be played again after a particular period. Therefore, the expected number of periods is $\frac{1}{1-\delta}$ (Roth and Murnighan, 1978).

Following Dal Bó and Fréchette (2018) and the literature therein, we reduce the strategy set and assume that players choose between two strategies at the game's inception: Grim and AlwaysDefect (AD).[1] In AD, players defect in each period, forever. In Grim, players start by cooperating and cooperate as long as the other player has cooperated in the previous period. Once the other player defects, they defect forever. Hence, Grim is the strongest possible and completely unforgiving retaliation for defection by the other player.

Multiple Equilibria and Equilibrium Selection In the infinitely repeated prisoner's dilemma, $\{\text{AD}, \text{AD}\}$ is always a subgame-perfect Nash-equilibrium. In addition,

[1]Experimental evidence shows that the pure strategies Grim, AlwaysDefect and Tit-for-Tat are most common (Dal Bó and Fréchette, 2019; Romero and Rosokha, 2023). Against AlwaysDefect, Tit-for-Tat and Grim behave identically. Thus, the results presented here remain the same, irrespective of whether we compare AlwaysDefect against Grim or AlwaysDefect against Tit-for-Tat. Furthermore, Tit-for-Tat and Grim always cooperate with each other.

$\{\text{Grim}, \text{Grim}\}$ is a subgame-perfect Nash-equilibrium if

$$\delta \geq \delta^{SPE}(g) \quad \text{where} \quad \delta^{SPE}(g) \equiv \frac{g}{1+g} \tag{1}$$

Thus, if players are sufficiently patient or if the continuation probability is sufficiently high, multiple equilibria exist, including cooperative equilibria. $\delta^{SPE}(g)$ is increasing in g, therefore a higher gain parameter g makes cooperation less likely. Furthermore, note that $\delta^{SPE}(g)$ does not depend on the loss parameter l. This implies that the subgame-perfect Nash-equilibrium excludes the possibility that variations in the sucker's payoff S might affect players' strategy choices, too.

With multiple equilibria, the literature has recognized the need for equilibrium selection criteria to determine the outcome of an infinitely repeated prisoner's dilemma. Two equilibrium selection criteria stand out: Pareto-dominance (Friedman, 1971; Fudenberg and Maskin, 1986) and risk-dominance (Harsanyi and Selten, 1988; Blonski et al., 2011; Dal Bó and Fréchette, 2011; Blonski and Spagnolo, 2015; Breitmoser, 2015).

According to Pareto-dominance, if multiple equilibria exist, players choose the Pareto-superior equilibrium. Thus, $\{\text{Grim}, \text{Grim}\}$ is Pareto-dominant as soon as $\{\text{Grim}, \text{Grim}\}$ is an equilibrium, i.e. if

$$\delta \geq \delta^{PD}(g) \quad \text{where} \quad \delta^{PD}(g) \equiv \frac{g}{1+g} \tag{2}$$

Theories of equilibrium selection can also account for *strategic uncertainty*: A cooperative equilibrium is risk-dominant if it is individually optimal to cooperate given that players are maximally uncertain, believing that their opponent randomizes between Grim and AD with a probability of 50% (see Blonski and Spagnolo, 2015). Thus, $\{\text{Grim}, \text{Grim}\}$ is risk-dominant if

$$\delta \geq \delta^{RD}(g,l) \quad \text{where} \quad \delta^{RD}(g,l) \equiv \frac{g+l}{1+g+l} \tag{3}$$

The critical value $\delta^{RD}(g,l)$ is increasing in g and l. Thus, both a higher gain parameter g and a higher loss parameter l make cooperation less likely. Moreover, for any prisoner's dilemma, it holds that $\delta^{RD}(g,l) > \delta^{PD}(g)$, i.e. if mutual cooperation is risk-dominant, it is also Pareto-dominant (Blonski et al., 2011).

In this paper, we consider the belief of a player about the probability $p \in [0,1]$ that the other player chooses a cooperative strategy.[2] With beliefs as a continuous measure of varying degrees of strategic uncertainty, we can derive a continuous measure of equilibrium selection based on the expected payoffs of cooperation and defection. For all players, it is

[2]If they have a belief of $p = 0$, they are certain that their opponent will defect. Then, AlwaysDefect is a dominant strategy for all δ and all $\Gamma(g,l)$. If a player has a belief of $p = 1$, they are perfectly certain that their opponent will play Grim.

individually optimal to play Grim if and only if the expected payoff from Grim is larger than or equal to the expected payoff of AlwaysDefect:

$$p \cdot \left(\frac{1}{1-\delta}\right) + (1-p) \cdot (-l) \geq p \cdot (1+g) \tag{4}$$

where, for each player, $p \cdot \left(\frac{1}{1-\delta}\right)$ is the expected discounted payoff from mutual cooperation, $(1-p) \cdot (-l)$ is the expected payoff from unilateral cooperation, and $p \cdot (1+g)$ is the expected payoff from unilateral defection. Rearranging (4) for δ yields the condition on δ for which mutual cooperation is individually optimal. Thus, the critical discount factor $\delta^*(p, g, l)$ is a continuous function of p, g, and l.[3]

Definition. *Let $\Gamma(g, l)$ be an infinitely repeated prisoner's dilemma of two players with strategy set $\omega \equiv \{Grim, AD\}$, discount factor $\delta \in (0, 1)$, belief $p \in [0, 1]$, and domain $\{g, l \in \mathrm{I\!R} | g, l > 0 \wedge l > g - 1\}$. Then the continuous function $\delta^* : D \mapsto R$ with domain $D \equiv \{p, g, l \in \mathrm{I\!R} | g, l > 0 \wedge l > g - 1 \wedge 0 \leq p \leq 1\}$, range $R \in \mathrm{I\!R}$, image $I = [0, 1]$ and formula*

$$\delta^*(p, g, l) = \frac{p \cdot (g-l) + l}{p \cdot (1+g-l) + l} \tag{5}$$

is called the critical discount factor of $\Gamma(g, l)$ for any given p.

This critical discount factor can be understood as a generalization of equilibrium selection that encapsulates both risk-dominance and Pareto-dominance. This is because plugging in $p = 0.5$ or $p = 1$ into $\delta^*(p, g, l)$ yields $\delta^{RD}(g, l)$ or $\delta^{PD}(g)$, respectively. Therefore, assuming that both players choose their strategy according to the risk-dominance criterion is equivalent to assuming $p = 0.5$. Assuming that both players choose their strategy according to the Pareto-dominance criterion is equivalent to assuming $p = 1$. Thus, equilibrium selection based on risk- or Pareto-dominance is equivalent to focusing on two specific beliefs, $p = 0.5$ or $p = 1$. We will show that these specific beliefs yield two special cases of the effects of the payoff parameters on cooperation.

Following the literature (see, e.g., Athey et al., 2004; Gilo et al., 2006; Bruttel, 2009; Blonski et al., 2011), we interpret the critical discount factor $\delta^*(p, g, l)$ as a measure of the (inverse of the) probability of cooperation. If the critical discount factor increases, cooperation becomes *less* probable. If the critical discount factor decreases, cooperation becomes *more* probable. Thus, policy interventions designed to hinder or foster cooperation can be analyzed by their potential to manipulate the critical discount factor in the desired direction (Blonski et al., 2011).

[3] Rearranging (4) for p would yield the size of the basin of attraction of defection, which captures the range of beliefs for which defection is optimal (Dal Bó and Fréchette, 2011).

3 Interaction of Payoffs, Beliefs, and Cooperation

In this section, we will show that the effect size of changes in the payoff parameters on cooperation depends inherently on a player's belief about the probability of cooperation by their opponent. We will show that if the beliefs are *high*, increasing the gain g from unilateral defection has a large negative effect on cooperation, while increasing the loss l from unilateral cooperation has a negligible effect. However, if beliefs are *low*, increasing g only has a negligible effect, while increasing l has a large negative effect on cooperation.

For ease of presentation, we consider interior $p \in (0,1)$ in the proposition. For $p = 1$, $\delta^*(p,g,l)$ increases in the gain g and is independent of the loss l, whereas for $p = 0$, $\delta^*\big(p,g,l\big)$ is independent of the gain g and the loss l.

Proposition. *Let $\Gamma(g,l)$ be an infinitely repeated prisoner's dilemma of two players with strategy set $\omega \equiv \{Grim, AD\}$, discount factor $\delta \in (0,1)$, belief $p \in (0,1)$, domain $\{g,l \in \mathbb{R} | g,l > 0 \wedge l > g-1\}$ and critical discount factor $\delta^*(p,g,l)$. Then:*

(a) The positive effect of an increase in g on δ^* increases in p.

(b) For any given l, the maximum possible positive effect of an increase in l on δ^* decreases in p.

Proof. Let us first consider how the critical discount factor behaves in the payoff parameters. Deriving $\delta^*\big(p,g,l\big)$ with respect to g and l yields:

$$\frac{\partial \delta^*\big(p,g,l\big)}{\partial g} = \frac{p^2}{\big(l - l \cdot p + p + g \cdot p\big)^2} > 0 \tag{6}$$

$$\frac{\partial \delta^*\big(p,g,l\big)}{\partial l} = \frac{p - p^2}{\big(l - l \cdot p + p + g \cdot p\big)^2} > 0 \tag{7}$$

For $0 < p < 1$, both derivatives are positive, i.e. the critical discount factor increases in both g and l.

Deriving (6) and (7) with respect to p demonstrates how these effects depend on p:

$$\frac{\partial \delta^*\big(p,g,l\big)}{\partial g \partial p} = \frac{2 \cdot l \cdot p}{\big(l - l \cdot p + p + g \cdot p\big)^3} \tag{8}$$

$$\frac{\partial \delta^*\big(p,g,l\big)}{\partial l \partial p} = \frac{l - p \cdot (1 + g + l)}{\big(l - l \cdot p + p + g \cdot p\big)^3} \tag{9}$$

The denominator in both derivatives is positive because $l > l \cdot p$ for $0 < p < 1$. The numerator in Equation (8) is positive because $l, p > 0$. Thus, the positive effect of an increase in the gain g on the critical discount factor $\delta^*\big(p,g,l\big)$ increases in the belief p:

$$\frac{\partial \delta^*\big(p,g,l\big)}{\partial g \partial p} > 0 \tag{10}$$

This establishes part (a) of the proposition.

The numerator in Equation (9) is negative for

$$p > \frac{l}{1+g+l} \tag{11}$$

Thus, the positive effect of an increase in the loss l on the critical discount factor $\delta^*\big(p, g, l\big)$ decreases in the belief p if and only if (11) is true:

$$\frac{\partial \delta^*\big(p, g, l\big)}{\partial l \partial p} < 0 \tag{12}$$

If the condition in (11) is reversed, the positive effect of an increase in the loss l on the critical discount factor $\delta^*\big(p, g, l\big)$ increases in the belief p. Thus, there exists a maximum of Equation (7) at $p' = \frac{l}{1+g+l}$. Rearranging p' for l yields:

$$l(p, g) = \frac{p \cdot (1+g)}{1-p} \tag{13}$$

Taking Equation (13) as the solution to maximizing $\delta^*_l(p, g, l) \equiv \frac{\partial \delta^*(p,g,l)}{\partial l}$ with respect to p, and plugging in yields:

$$\delta^*_l\big(p, g, l = l(p, g)\big) = \frac{p - p^2}{\Big(\frac{p\cdot(1+g)}{1-p} - \frac{p\cdot(1+g)}{1-p} \cdot p + p + g \cdot p\Big)^2} \tag{14}$$

Equation (14) describes the maximum-possible effect of l on the critical discount factor $\delta^*\big(p, g, l\big)$, that is reachable for any given l, as a function of the belief p and the gain g. Deriving Equation (14) with respect to p yields:

$$\frac{\partial \delta^*_l\big(p, g, l = l(p, g)\big)}{\partial p} = -\frac{1}{4 \cdot (1+g)^2 \cdot p^2} < 0 \tag{15}$$

Thus, for any given l, the maximum-possible positive effect of an increase in the loss l on the critical discount factor $\delta^*\big(p, g, l\big)$ decreases in p. This establishes part (b) of the proposition. ■

Let us provide an intuition for the proposition. Consider how the gain g affects the critical discount factor $\delta^*\big(p, g, l\big)$ for varying beliefs p. If the belief p is relatively low, the players find it unlikely that their opponent will cooperate. An increase in the payoff from unilateral defection, i.e., an increase in the gain g will be relatively inconsequential because players do not expect to enter an outcome where they *defect unilaterally.* However, if the belief p is relatively high, players find it likely that their opponent will cooperate. Then, an increase in the gain g will matter substantially because defecting unilaterally is a relatively probable outcome. The players react to the *expected* payoffs of their actions, not the payoffs as such. Thus, the positive effect of an increase in the gain g on the critical discount factor δ^* increases in the belief p.

Consequently, the opposite must be true for the effect of the loss l on the critical discount factor $\delta^*\big(p, g, l\big)$: The effect of the loss l will be negligible if players find it very likely that the other player will cooperate. However, suppose that cooperation by the other player seems unlikely; in that case, players will attach considerable weight to the payoff from unilateral cooperation, i.e. the loss l, because this is a relatively probable outcome if they cooperate in such an environment. Furthermore, defecting oneself also becomes more probable with decreasing p. In the extreme case of $p = 0$, players will certainly defect, and hence will not react to any changes in either payoff. This opposing effect loses relevance with decreasing l because cooperation becomes more likely with decreasing l. Thus, increasing the loss l has large effects in situations where the belief p and the loss l are low to start with.[4]

With the proposition, we can also compare the marginal effects of gain and loss across beliefs: Players with low beliefs react more strongly to the loss than to the gain, while players with high beliefs react more strongly to the gain than to the loss.

Corollary. *Consider the differences in the marginal effects of g and l:*

$$\frac{\partial \delta^*\big(p, g, l\big)}{\partial g} - \frac{\partial \delta^*\big(p, g, l\big)}{\partial l} = \frac{2 \cdot p^2 - p}{\big(l - l \cdot p + p + g \cdot p\big)^2} \tag{16}$$

The difference is positive (negative) iff $p > 0.5$ ($p < 0.5$). Thus, the effect of an increase in g on δ^* is larger (smaller) than the effect of an increase in l on δ^* for $p > 0.5$ ($p < 0.5$). For $p = 0.5$, the effects of g and l are equal.

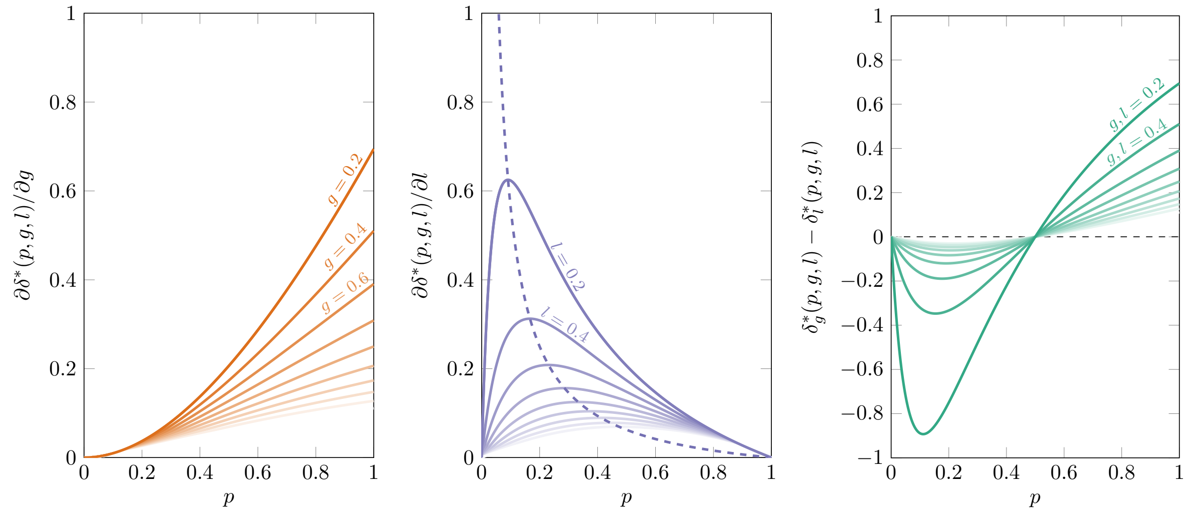
The difference has two local extrema. For $p < 0.5$, the absolute value of the difference is highest at $p' = \frac{l}{g+3 \cdot l+1}$. For $p > 0.5$, the difference is largest for $p \to 1$.

From the corollary, we can see interesting cases of the proposition: Risk-dominance and Pareto-dominance. For $p = 0.5$, i.e. risk-dominance, the critical discount factor $\delta^*\big(p, g, l\big)$ increases in the gain g and the loss l by the same magnitudes. This is by design and an advantage over Pareto-dominance: Blonski et al. (2011) derive risk-dominance from the axiom of "equal weight" of gain and loss, among others. However, it also implies that if players were to have other beliefs than 50/50, the risk-dominance criterion underestimates the effect of the gain and overestimates the effect of the loss for $p > 0.5$, while for $p < 0.5$, it underestimates the effect of the loss and overestimates the effect of the gain. For $p = 1$,

[4] If we include an additional restriction on the domain of a prisoner's dilemma—cooperation always increases efficiency, i.e., $T + S > 2 \cdot P$ or $1 + g > l$, see Friedman and Sinervo (2016)—the positive effect of an increase in the loss l on the critical discount factor $\delta^*(p, g, l)$ always decreases in p for $p > 0.5$. This is because plugging Equation (13) into Equation (7) yields $p > 0.5$, for which $\frac{\partial \delta^*(p,g,l)}{\partial l \partial p}$ is always negative.

i.e. Pareto-dominance (see Section 2), we can see that only the gain has an effect on cooperation, while the loss does not.

Figure 1: The Effects of Gain and Loss, Conditional on the Belief



(a) Effect of the Gain (b) Effect of the Loss (c) Difference in the Effects

Note: The derivative of the critical discount factor $\delta^*(p, g, l)$ with respect to (a) the gain g and (b) the loss l as well as (c) the difference between the derivatives, for different beliefs p. Figure (a) additionally varies g, while holding $l = 1$ constant. Figure (b) additionally varies l, while holding $g = 1$ constant. Figure (c) varies g and l symmetrically. All plots satisfy $l > g - 1$ (mutual cooperation is Pareto-superior to unilateral defection), ensuring that every line represents a true prisoner's dilemma. The dashed line in Figure (b) displays the maximum possible effect of l for any given l in p, see Equation (14).

In Figure 1, we visualize the Proposition and the Corollary. We draw the first derivatives of the critical discount factor $\delta^*(p, g, l)$ with respect to the gain g (Figure 1a, Equation 6) and the loss l (Figure 1b, Equation 7) as a function of the belief p and for different initial values of the gain g and the loss l.

In Figure 1a, the positive effect of the gain g on the critical discount factor $\delta^*(p, g, l)$ clearly increases in the belief p. Additionally, for any given p, there is a decreasing marginal effect of g (the higher g, the lower the lines across p). In Figure 1b, the dashed line indicates the maximum possible effect of l in p. The strictly decreasing trend of this curve illustrates that the positive effect of the loss l on the critical discount factor $\delta^*(p, g, l)$ decreases with increasing p. We can also see the non-monotonicity of the effect of the belief on the effect of the loss: As the belief approaches zero, the effect of the loss on $\delta^*(p, g, l)$ decreases because the direct incentive to defect becomes very large. Figure 1c, finally, displays the difference of the two first derivatives, see Equation (16). For low beliefs, players react more strongly to the loss than to the gain, and for high beliefs, players react more strongly to the gain than to the loss.[5]

[5]In this section, we assume that the belief p is exogenous to the payoff parameters of the game. The exogenous belief p may, for example, be interpreted as the player's belief over the distribution of cooperative players in the population. In Section A.1 in the appendix, we relax this assumption of

4 Empirics I: Meta Study of Lab Experiments

In this section, we present the first part of empirical evidence testing our proposition that the effect of the payoff parameters on on cooperation depends on a player's belief about the probability of cooperation by their opponent. For this test, we use existing experimental data collected by Dal Bó and Fréchette (2018) and extended by Fudenberg and Karreskog Rehbinder (2024).[6] This data set contains data on the stage game parameters g and l, and a given participant's cooperation decision as well as their opponent's in a given period of a given supergame. After removing observations which we cannot use to test our theory,[7] we obtain a total of $205,924$ decisions.

In the theory section, we reduced players' strategy sets to one cooperative strategy (Grim) and one defective strategy (AlwaysDefect), which players choose once at the beginning of the game. In contrast, in the experimental data, participants made decisions about cooperation or defection in each period of a supergame. In line with the literature (Dal Bó and Fréchette, 2018; Fudenberg and Karreskog Rehbinder, 2024), we interpret the decision of a participant in the first period of a supergame as the choice between a cooperative or a defective strategy. The data set contains $54,176$ such decisions from the first period of a supergame.

Since the literature on prisoner's dilemma experiments has only started eliciting beliefs rather recently (Aoyagi et al., 2024; Andres, 2024; Gill and Rosokha, 2024), the data set at hand does not contain information on the players' beliefs. However, for the present research question, we need information about players' beliefs to study how beliefs moderate the effect of the payoff parameters on cooperation. We approach this challenge with the assumption that within a given experimental session, a participant's beliefs about the probability of cooperation by their opponent will be shaped by their experiences in this session. We discuss how we construct beliefs from this assumption in Subsection 4.1, and validate these constructed beliefs in Subsection 4.2. We present and discuss our results in Subsections 4.3 and 4.4, respectively.

---

exogeneity by allowing for beliefs that vary in the parameters g and l.

[6] The meta study by Dal Bó and Fréchette (2018) analyzes data by Andreoni and Miller (1993); Cooper et al. (1996); Dal Bó (2005); Dreber et al. (2008); Aoyagi and Fréchette (2009); Duffy and Ochs (2009); Dal Bó et al. (2010); Dal Bó and Fréchette (2011); Blonski et al. (2011); Fudenberg et al. (2012); Bruttel and Kamecke (2012); Sherstyuk et al. (2013); Kagel and Schley (2013); Fréchette and Yuksel (2017) and Dal Bó and Fréchette (2019). Fudenberg and Karreskog Rehbinder (2024) add data by Arechar et al. (2018); Dal Bó and Fréchette (2019); Proto et al. (2019); Honhon and Hyndman (2020).

[7] The combined data set contains data of $2,612$ participants and $232,298$ decisions. We remove $9,146$ observations from the study by Duffy and Ochs (2009) because the participant IDs for this study are not unique, which makes the belief construction procedure (see Section 4.1) impossible as we cannot uniquely identify past experiences. Further, we restrict the data set to observations where (C, C) is an equilibrium path because in this paper we consider equilibrium selection, which only comes into play once multiple equilibria exist, see Section 2. This additionally removes $17,228$ observations.

4.1 Meta Study: Constructing Beliefs from Experiences

We model beliefs as period-beliefs that the opponent plays action $a_j^{s,t} = C$, denoted by $\mu_i^{s,t}$, following Aoyagi et al. (2024), who elicit period-beliefs on actions. Because we do not observe any beliefs in the data, we construct them assuming that beliefs are updated from period to period through a simple Markov-process. In each period of each supergame, a player updates their prior belief with their experience in the previous period via a linear combination, attaching a given weight λ onto the prior over the experience. For the first period of the first supergame, we assume that the players are maximally uncertain and hold a belief of 0.5 (Blonski et al., 2011; Dal Bó and Fréchette, 2011).

The period-wise updating method uses as much information as possible to construct beliefs and reflects recent experimental evidence showing that beliefs are heavily influenced by experiences in previous periods and supergames (Aoyagi et al., 2024; Gill and Rosokha, 2024). More explicitly, we know that the last periods of previous supergames are especially important for cooperation in the next supergame (Fudenberg and Karreskog Rehbinder, 2024) and that beliefs carry over across supergames, even after rematching (Gill and Rosokha, 2024).

As a proxy for experience, we take an indicator of whether the two players achieved the Pareto-efficient outcome in the previous period. If both players cooperated, the experience is positive. Otherwise, the experience is negative. This follows recent empirical evidence by Aoyagi et al. (2024, p. 14), who find that "beliefs become equally pessimistic regardless of which player has defected." This method incorporates that the own action, too, might affect the other's behavior and, thus, should affect the own belief: If I defected, the other player will be more likely to defect in the next period. Thus, we can define the Markov-Belief piece-wise:

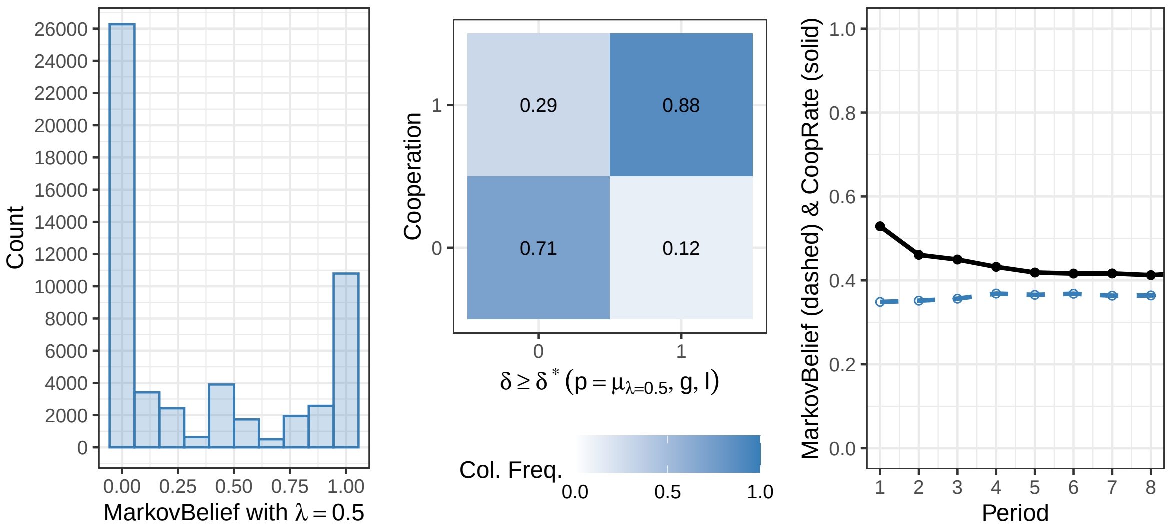
Definition. *The Markov Belief of a player in period t of supergame s is given by:*

$$\mu_i^{s,t} = \begin{cases} 0.5 & \text{if } s = 1, t = 1 \\ \lambda\mu_i^{s,t-1} + (1-\lambda)c_i^{s,t-1} & \text{if } t > 1 \\ \lambda\mu_i^{s-1,T} + (1-\lambda)c_i^{s-1,T} & \text{else, i.e. if } s > 1, t = 1 \end{cases} \tag{17}$$

where $\mu_i^{s,t} \in [0,1]$ is the belief about the probability that the other player plays $a_j^{s,t} = C$, c_i is an indicator for (C,C) in the previous period $t-1$ of the same supergame s or an indicator for (C,C) in the final period T of the previous supergame $s-1$, and $\lambda \in [0,1]$ is the weight on the prior.

This belief construction procedure is straightforward and can be applied very flexibly by varying the weighting parameter λ. With increasing λ, a player places more weight onto their prior over their experience. In the main text, we present the results using

Figure 2: Validation of Markov Beliefs



(a) Distribution of Beliefs (b) Predicting Cooperation (c) Accuracy of Beliefs

Note: Figure 2a presents the distribution of the Markov belief in the first period. Figure 2b presents the confusion matrix of predicted cooperation vs. actual cooperation with column percentages depicted by color intensity. Figure 2c presents average beliefs (dashed) and the average cooperation rate (solid) over the first eight periods, across supergames. All figures display Markov beliefs estimated with $\lambda = 0.5$. In the appendix, we present the three figures for other values of λ, see Figure 6.

$\lambda = 0.5$. In the appendix, we show details on the constructed beliefs and the results for other values of λ, see Sections A.2.1 and A.2.2, respectively.

4.2 Meta Study: Validation of the Belief Construction

Figure 2a depicts the distribution of Markov beliefs in the sample of first-period decisions that we use to test our proposition. Although there is support over the entire domain $[0, 1]$, we see clear peaks at $\mu \approx 0.5$ and $\mu \approx 1$, as well as a large number of observations at $\mu \approx 0$.

As a first validation exercise of the constructed beliefs, we compute how accurately these beliefs predict cooperation choices across all periods and supergames. For this exercise, we assume that a participant should cooperate whenever the actual discount factor in their experimental treatment is greater than the critical discount factor presented in Section 2, $\delta^*(p, g, l)$. For p, we plug in the estimated Markov beliefs. Thus, the prediction is $a_i = C$ if $\delta \geq \delta^*_i(p_i = \mu_i^{s,t}, g, l)$.

Figure 2b presents the confusion matrix of predicted cooperation and actual cooperation decisions in the first periods. The confusion matrix displays correct and incorrect predictions with column percentages. We can see that our predictions are very accurate for both defection and cooperation, correctly predicting defection in 71% of cases and correctly predicting cooperation in 88% of cases. Overall accuracy is then a weighted average of these column-accuracies: Using the critical discount factor presented in this paper, we correctly predict cooperation or defection by a player in 78% of cases, compared to an accuracy of 53% with subgame-perfection/Pareto-dominance and 66% with risk-dominance.

As a further validation exercise, we replicate the comparison of cooperation rates and beliefs from Aoyagi et al. (2024) in Figure 2c.[8] When comparing beliefs and actual cooperation rates over time, we find a remarkably similar picture to Aoyagi et al. (2024), who elicit period-beliefs from the participants directly. Comparing our figure with theirs, we find that here and there, players are overly pessimistic in the beginning, but then the cooperation rate decreases, and the beliefs slightly increase such that the two lines are very close.

The high predictive accuracy and similarity to key features of previous empirical results make us confident that our belief construction procedure yields a good proxy for participants' true beliefs. Let us now turn to analyzing how the payoff parameters affect cooperation depending on the constructed beliefs.

4.3 Meta Study: Results

The meta study data contain various treatments that vary the stage game payoff parameters g and l in the domains $[0.09, 2.57]$ and $[0.11, 8]$, respectively. In this section, we study how participants' decisions vary with these parameters depending on their estimated Markov belief. In order to facilitate the comparison of the resulting graphical illustrations with the theoretical results on $\delta^*(p, g, l)$ depicted in Figure 1, we estimate the effect of the gain and the loss on defection, not cooperation.

We do the analysis in two steps, via sample-splitting and with explicit interactions. First, we estimate the partial effects of the gain and the loss on defection for different levels of the Markov beliefs and compare the size of the partial effects across belief levels. This approach illustrates that the partial effects of the gain and the loss, and the difference between the two effects, do vary in the beliefs as predicted by the theory. In the second step, we add beliefs to the regressions to provide an explicit test of how the effects of gain and loss differ in the Markov beliefs.

For the sample-splitting approach, we estimate the following models via Probit regressions, clustering standard errors at the individual participant level:

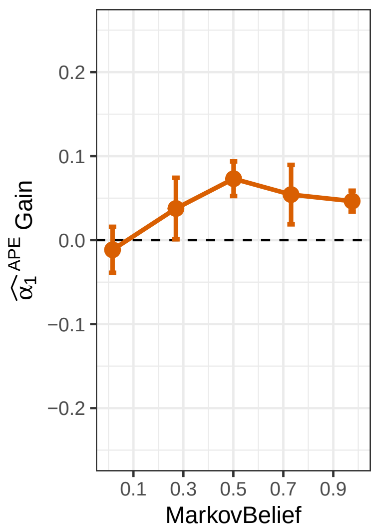
$$M_1^{g,l}: \quad E[D] = \Phi\big(\alpha_0 + \alpha_1 \cdot g + \alpha_2 \cdot l\big) \tag{18}$$

$$M_1^{\Delta}: \quad E[D] = \Phi\big(\gamma_0 + \gamma_1 \cdot g + \gamma_2 \cdot (g + l)\big) \tag{19}$$

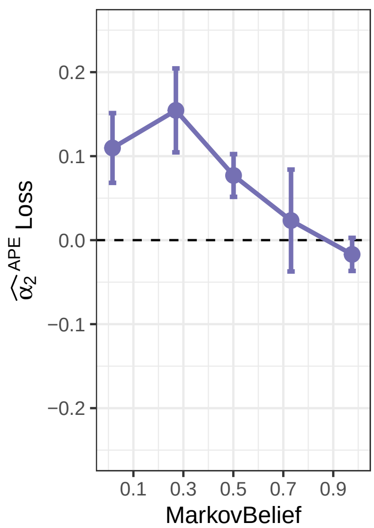
[8] See "Figure 2: Choices and Beliefs by Round" and "Figure 9: Choices and Beliefs by Round (Equivalent to Figure 2)," in Aoyagi et al. (2024). In their Figure 9, Aoyagi et al. (2024) find that differences in the stage game parameters do affect the level of the lines, but not the relative position of cooperation rate and beliefs. This is exactly what we see when we compare their figures with ours.

$M_1^{g,l}$ estimates the partial effects of the gain and the loss on the probability to defect in the first period of a supergame. We include both parameters in one model since they are correlated in the data set ($r = 0.66$) and we aim to identify the partial effects of both parameters, holding the other constant.[9] In M_1^{Δ}, we then estimate the difference between the effects of gain and loss. If gain and loss have the same effect on defection, then γ_1 should be indistinguishable from zero because γ_2 would capture this common effect. Thus, we take γ_1 as a measure of the difference of the effects of gain and loss. To illustrate whether the partial effects differ in the beliefs as predicted by our proposition, we estimate $M_1^{g,l}$ and M_1^{Δ} separately for five intervals in the beliefs, centered on $0.1, 0.3, 0.5, 0.7, 0.9$, respectively. We believe that these groups offer sufficiently fine-grained information while preserving sufficiently large sample sizes even for 0.3 and 0.7, especially given that there is some degree of multicollinearity between the regressors.

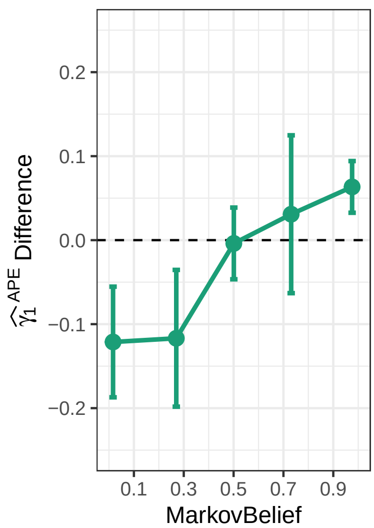
Figure 3: The Effects of Gain and Loss, Conditional on the Markov Belief



(a) Effect of the Gain



(b) Effect of the Loss



(c) Difference in the Effects

Note: All figures report average partial effects on the probability of defection and 95% confidence intervals following the Probit models $M_1^{g,l}$ and M_1^{Δ}. See Table 5 in the appendix for more details. Standard errors are clustered at the individual participant level. The coefficients are displayed at the mean of the Markov belief in the respective bracket $[0, 0.2], (0.2, 0.4], (0.4, 0.6], (0.6, 0.8], (0.8, 1]$, i.e. at $0.02, 0.27, 0.50, 0.73, 0.98$, respectively.

Figure 3 presents the results. In the panels, we report the average partial effects on defection for the gain (3a), the loss (3b), and the difference between the two effects (3c). Table 5 in the appendix reports the detailed estimation results. Across all three panels, we observe strong support for the theory. An increase in the gain from unilateral defection has a positive effect on the probability of defecting, but only for relatively high beliefs.

[9]The surprisingly high correlation of g and l in the data set stems from the fact that many studies vary the reward payoff R across treatments. The reward payoff R and the punishment payoff P affect both the gain and the loss simultaneously, see Section 2.

For low beliefs, there is a null effect, which is precisely estimated for the interval around 0.1, with a relatively large sample size. Increases in the loss, too, affect the probability to defect positively, but this effect is largest when beliefs are low, and disappears completely for high beliefs. Finally, comparing the two coefficients reveals that participants with low beliefs react more strongly to the loss than to the gain, while participants with high beliefs react more strongly to the gain than to the loss. Participants with medium beliefs react similarly to both parameters, as predicted by the theory.[10]

To provide an explicit test of how the partial effects of gain, loss, and the difference between both change in the belief, we estimate fully interacted models:

$$\begin{aligned} M_2^{g,l}: \quad E[D] &= \Phi\big(\beta_0 + \beta_1 \cdot g + \beta_2 \cdot l + \beta_3 \cdot \mu \\ &\qquad + \beta_4 \cdot g{\times}\mu + \beta_5 \cdot l{\times}\mu\big) \qquad (20) \\ M_2^{\Delta}: \quad E[D] &= \Phi\big(\tau_0 + \tau_1 \cdot g + \tau_2 \cdot (g{+}l) + \tau_3 \cdot \mathbb{1}(\mu{>}0.5) \\ &\qquad + \tau_4 \cdot g{\times}\mathbb{1}(\mu{>}0.5) + \tau_5 \cdot (g{+}l){\times}\mathbb{1}(\mu{>}0.5)\big) \qquad (21) \end{aligned}$$

$M_2^{g,l}$ estimates the partial effects of gain and loss and their interaction with the Markov belief μ on the probability of defecting. In M_2^{Δ}, we estimate the difference between the coefficients, testing whether they are different for low beliefs ($\mu \leq 0.5$) or for high beliefs ($\mu > 0.5$) by including a dummy for high beliefs. If gain and loss have the same effect on defection for low beliefs, then τ_1 should be indistinguishable from zero. In the same way, if gain and loss have the same effect on defection for high beliefs, then τ_4 should be indistinguishable from zero. Thus, we take τ_1 and τ_4 as a measure of the difference in the effects of gain and loss.

Table 2 presents the average partial effects of $M_2^{g,l}$ and M_2^{Δ}. For beliefs $\to 0$, the gain does not have an effect on the probability of defecting ($\widehat{\beta_1}^{APE}$), while the loss does have a substantial positive effect ($\widehat{\beta_2}^{APE}$). Further, the effect of the gain increases in the belief ($\widehat{\beta_4}^{APE}$), while the effect of the loss decreases in the belief ($\widehat{\beta_5}^{APE}$). Note that $\widehat{\beta_2}^{APE}$ and $\widehat{\beta_5}^{APE}$ cancel out, such that the model predicts a null effect for the loss for beliefs $\to 1$.

The estimated differences in the effects of gain and loss also follow the theoretical predictions. For small beliefs, the gain has a smaller effect than the loss ($\widehat{\tau_1}^{APE}$), while for high beliefs, the gain has a larger effect than the loss ($\widehat{\tau_4}^{APE}$). Finally, the belief itself has a large negative effect on the probability of defecting ($\widehat{\beta_3}^{APE}$ and $\widehat{\tau_3}^{APE}$).

Thus, we find the following overall result, in line with the theoretical proposition: The effect of the gain increases in the belief, and the effect of the loss decreases in the belief. Therefore, participants with low beliefs react more strongly to the loss than to the gain, and participants with high beliefs react more strongly to gain than to the loss. In fact, participants with beliefs close to zero do not react to the gain at all, and participants with beliefs close to one do not react to the loss.

[10] In the appendix, we present the same figures for different values of the weighting parameter λ with $\lambda \in \{0.3, 0.4, 0.5, 0.6, 0.7\}$, see Figure 7. The results are similar.

Table 2: The Effects of Gain and Loss, Conditional on the Markov Belief

$M_2^{g,l}$			M_2^{Δ}		
$\widehat{\beta}_1^{APE}$	Gain	-0.0150	$\widehat{\tau}_1^{APE}$	Difference	-0.0970
		(0.0110)			(0.0277)
		[0.1744]			[0.0005]
$\widehat{\beta}_2^{APE}$	Loss	0.0911	$\widehat{\tau}_2^{APE}$	Gain + Loss	0.1016
		(0.0169)			(0.0173)
		[0.0000]			[0.0000]
$\widehat{\beta}_3^{APE}$	MarkovBelief	-0.6629	$\widehat{\tau}_3^{APE}$	MarkovBelief > 0.5	-0.5660
		(0.0187)			(0.0124)
		[0.0000]			[0.0000]
$\widehat{\beta}_4^{APE}$	MarkovBelief $\times$ Gain	0.1526	$\widehat{\tau}_4^{APE}$	(MarkovBelief > 0.5) $\times$ Difference	0.1195
		(0.0208)			(0.0274)
		[0.0000]			[0.0000]
$\widehat{\beta}_5^{APE}$	MarkovBelief $\times$ Loss	-0.0917	$\widehat{\tau}_5^{APE}$	(MarkovBelief > 0.5) $\times$ (Gain + Loss)	-0.0237
		(0.0335)			(0.0172)
		[0.0062]			[0.1693]

Note: Average partial effects of Probit regressions with defection as the dependent variable. Standard errors, clustered at the individual participant level, in parentheses, p-values in brackets. $N = 54,176$.

4.4 Meta Study: Discussion

So far, we have shown how the Markov beliefs, constructed from past experiences, affect the partial effects of the payoff parameters on cooperation in the predicted way. However, while our validation provides strong evidence that the estimated Markov beliefs correctly reflect the true beliefs, this remains an assumption to some extent. Let us discuss two further caveats.

For a causal interpretation of the effect, we would need to assume that these constructed beliefs are exogenous. Markov beliefs may fall short of this assumption. Experiences are not exogenous after the first period of a given supergame because the opponent's behavior will be influenced by a player's own behavior, making experience an endogenous variable.[11] Such an endogeneity of beliefs would bias the estimation of the partial effect of the constructed belief on defection, and potentially also the the effect of the belief on the partial effects of the payoff parameters we are interested in.

In addition, while our theory assumes that players choose between strategies for the whole supergame, the data in standard infinitely repeated prisoners' dilemma experiments

[11] We model explicitly that players should account for this in their belief formation: It is consistent to assume that a player's own past behavior affects their opponent's future behavior. In the construction of beliefs, we adjust for this by formalizing experience as an indicator of mutual cooperation between the two players. However, the endogeneity remains.

record actions taken every period, which might or might not follow standard strategies. In the meta study, we followed the common approach to interpret the first-period action as an indicator for the strategy choices, but this is another assumption. The next section presents a laboratory experiment that meets both assumptions, i.e. exogenous beliefs and strategy choices.

5 Empirics II: Laboratory Experiment

In this section, we present a laboratory experiment designed to closely match the set-up of the theory. The experiment enforces choices of strategies rather than actions and imposes beliefs exogenously. By restricting participants to choose among the two strategies Grim and AlwaysDefect, we make sure that the strategy space in the experiment exactly matches the theoretical setup, which allows us to conduct a straightforward test of the theory. We implement the exogenous variation of beliefs by informing participants truthfully that we will match them with a partner only after they have chosen their strategy in the infinitely repeated game and by telling them ex ante the probability that we will match them with a partner who plays Grim or AlwaysDefect, respectively.[12]

5.1 Experiment: Design & Procedures

In this section, we lay out an experimental design for testing the proposition in this paper. We will discuss the design first and then how we explained it to the participants.

Experimental Design All participants choose strategies in three different prisoner's dilemma games that vary the stage game payoff parameters within subjects (see Table 3). The baseline game, named Base, has parameters $g = l = 0.2$. The HighGain game has parameters $g = 1,\ l = 0.2$. The HighLoss game has parameters $g = 0.2,\ l = 1$. In the experiment, the stage game payoff tables are shown to the participants in standard notation, using T, S, P, R, indicating the amount of points they earn. The order in which the participants face HighGain and HighLoss is randomized between sessions, but the participants always start with Base.

In the experiment, participants choose between the two strategies Grim and Always-Defect for infinitely repeated versions of all three stage games. These strategies are then implemented, and participants receive the discounted payoffs. Participants choose their strategy for nine discount factors $\delta \in \{0.1, ..., 0.9\}$, i.e. for nine different infinitely repeated prisoner's dilemmas, so that for each participant, we can derive three critical discount factors from their choices, one for each stage game parameterization. This allows us to study

[12]For our test, it is important that we impose beliefs rather than eliciting them. Elicited beliefs may hide an underlying endogenous relationship to unobserved factors, such as preferences for cooperation or social preferences (Costa-Gomes et al., 2014), or they may be affected by the payoff parameters (Aoyagi et al., 2024; Andres, 2024). Our procedure ensures that beliefs are exogenous and that variations in beliefs and payoff parameters are orthogonal.

Table 3: Row Player's Stage Game Payoffs in the Three Games

	C	D		C	D		C	D
C	75	45	C	75	45	C	75	**25**
D	80	50	D	**100**	50	D	80	50
	C	D		C	D		C	D
C	1	-0.2	C	1	-0.2	C	1	$-\mathbf{1}$
D	$1+0.2$	0	D	$1+\mathbf{1}$	0	D	$1+0.2$	0
	Base			HighGain			HighLoss	

Note: T, S, P, R-notation in the first row and g, l-notation in the second row.

variations in the critical discount factors depending on the stage game payoffs and the beliefs.

We compare three treatments in which beliefs vary exogenously between subjects. At the beginning of a session, each participant is informed about the probability with which they will be matched to a player that plays Grim in a given decision. This probability remains constant throughout the experiment and participants are reminded of it on each decision screen.[13] The probability is randomly assigned, drawn individually from a discrete uniform distribution over $\{0.1, 0.5, 0.9\}$, resulting in the three between-subject treatments LowBelief, MedBelief and HighBelief.

Filtering Rule If participants' preferences satisfy the independence axiom, it must hold that if a participant prefers Grim over AlwaysDefect for a specific discount factor $\bar{\delta}$, then they must prefer Grim for all $\delta \geq \bar{\delta}$ and AlwaysDefect for all $\delta < \bar{\delta}$. Only then can a unique critical discount factor be computed for this participant. The actual decisions of the participants in the experiment could violate the independence axiom in two ways. First, they might switch more than once between strategies. Second, they may switch into the "wrong" direction, i.e., choose Grim for low discount factors and AlwaysDefect for high ones. We deal with this issue in the following preregistered way.

We enforce that participants switch at most once between the strategies because with more than one switch, we would obtain more than one critical discount factor for the same stage game parameterization by the same individual, which we could not interpret in a meaningful way. However, we allow participants to switch in the "wrong" direction. Such switches provide a clear-cut filtering rule for lack of comprehension as it is impossible to calculate a critical discount factor for such a participant. Therefore, we exclude participants who choose Grim for low discount factors and switch to AlwaysDefect for higher

[13] Translated from German into English, this reads "With a probability of [p]% you will be assigned to another person who chooses plan A [Grim]. This also means that with a probability of [$1-p$]% you will be assigned to another person who chooses plan B [AlwaysDefect]."

ones in at least one of the three games from all analyses because our outcome variable is undefined for these participants. The sample size N that we specify in the following refers to valid observations in the sense that the critical discount factors can be computed for these participants in the three games. Of the total of 312 participants in our lab, 253 (81%) followed the independence axiom and entered our final sample. Of these, 92 were randomly placed in the LowBelief treatment, 70 were in the MedBelief treatment and 91 were in the HighBelief treatment.

Procedures The sessions were held at the Potsdam Laboratory for Economic Experiments (PLEx). The subject pool consists mainly of students enrolled at the University of Potsdam. The sessions lasted about 60 minutes, including the welcome and payoff procedures. Participants were paid in cash at an exchange rate of 1:18 plus a show-up fee of eight euros. They earned about 17.97 euros on average. Payments were administered privately. Participants were asked to sign an informed consent sheet, detailing data protection issues and including a very broad description of the experiment, before each session began.

Instructions All participants faced the same instructions, regardless of the treatment and stage game order. The experimental instructions describe the underlying game and the concept of strategies to the participants. Actions are denoted as "X" (C) and "Y" (D), and strategies as "A" (Grim) and "B" (AlwaysDefect). Participants were told that they play "many repetitions" of the stage games and thus decide between strategies, instead of single actions, denoted as "plans" that would enact the respective actions in the many repetitions. We explain what these strategies mean and what each of the four possible joint strategies entails. We also explain how the stage game payoffs decrease over the many repetitions, depending on the discount factor, which we call the "residual factor." On each decision screen, participants had access to a payoff calculator for the infinitely repeated game in addition to the stage game payoff table. On request of the participant, this payoff calculator displayed the discounted payoffs for the four possible joint strategies and each possible discount factor.

In the beginning of the experiment, participants went through an extensive guided tour of the experiment on their computer screen. During this tour, they had to try out the payoff calculator in order to answer some comprehension questions, and we confirmed that they recognize the matching probability. In addition, they learned how to enter decisions in the price list format in which the strategies are presented.

In the instructions, we also explain the matching procedure in detail. We told participants that only one of their 27 decisions (three games $\times$ nine discount factors) would be payoff-relevant, and that they would be matched to another participant only after the decision and based on the other participant's decision. With a given probability, they would be matched with a participant who has chosen plan A or plan B in this specific supergame for this specific discount factor. If no such participant could be found, then the computer would step in. (We found a successful match in 98.7% of the cases, i.e. for 308 out of 312 participants.) We also reminded participants that this procedure is the same for all participants in the lab and that therefore it is very well possible that the other participant will not be matched to them in return. This should minimize the impact of social preferences on the participants' decisions, which is desirable in order to more closely match the theoretical setting. See Appendix A.3.1 for an English translation of the instructions.

5.2 Experiment: Hypotheses

For each participant, we elicit their critical discount factor for each of the three possible combinations of stage game parameters. We define the critical discount factor as the discount factor for which a given participant first plays Grim. If a participant chooses Grim for $\delta = 0.6$ and everything above, their critical discount factor is $\tilde{\delta^*} = 0.6$. If they always choose Grim, their critical discount factor is $\tilde{\delta^*} = 0.1$. If they always choose AlwaysDefect, their critical discount factor is $\tilde{\delta^*} = 1$.

Based on our model, we can calculate point predictions for the critical discount factors in each of the three treatments. The first three rows in Table 4 present these predictions for the three different combinations of stage game parameter combinations. The fourth and fifth rows in Table 4 show the differences in the critical discount factors between games. These differences are the theoretical point predictions for our two main outcome variables, i.e. the differences in the critical discount factors between the Base game and the HighGain game, and between the Base game and the HighLoss game, respectively:

$$\Delta\tilde{\delta^*}(p)|_{\Delta g} = \tilde{\delta^*}(p, g = 1, l = 0.2) - \tilde{\delta^*}(p, g = 0.2, l = 0.2) \tag{22}$$

$$\Delta\tilde{\delta^*}(p)|_{\Delta l} = \tilde{\delta^*}(p, g = 0.2, l = 1) - \tilde{\delta^*}(p, g = 0.2, l = 0.2) \tag{23}$$

From the theoretical point predictions, we form qualitative hypotheses based on the two comparisons in which the theory predicts the largest differences. We report, analyze, and discuss the remaining data, too. To test whether the belief moderates the effect of the gain on the critical discount factor, we take the differences between the critical discount factors in the Base game and in the HighGain game and compare these differences between beliefs (H_{Gain}). For a test of the moderating effect of the belief on the effect of the loss on the critical discount factor, we take the differences between the critical discount factors in the Base game and in the HighLoss game and compare these differences between the beliefs (H_{Loss}). Taking the differences in the point predictions for the effect of the gain and the loss allows us to compare the effects on an individual level (see row 6 in Table 4). This is possible because we are comparing symmetric changes in the stage

Table 4: Theoretical Point Predictions and Hypotheses

	LowBelief	MedBelief	HighBelief
(1) Base: $\tilde{\delta}^*(p, g = 0.2, l = 0.2)$	0.67	0.29	0.18
(2) HighGain: $\tilde{\delta}^*(p, g = 1, l = 0.2)$	0.74	0.55	0.51
(3) HighLoss: $\tilde{\delta}^*(p, g = 0.2, l = 1)$	0.90	0.55	0.24
(4) Effect of the gain: $\Delta\tilde{\delta}^*(p)\|_{\Delta g}$	0.07	0.26	0.32
(5) Effect of the loss: $\Delta\tilde{\delta}^*(p)\|_{\Delta l}$	0.24	0.26	0.06
(6) Difference: $\Delta\tilde{\delta}^*(p)\|_{\Delta g} - \Delta\tilde{\delta}^*(p)\|_{\Delta l}$	−0.17	0.00	0.26

H_{Gain}: The positive effect of increases in the gain on the critical discount factor increases from LowBelief to MedBelief.
H_{Loss}: The positive effect of increases in the loss on the critical discount factor decreases from MedBelief to HighBelief.

Note: The table displays the theoretical point predictions for the critical discount factors across belief treatments and stage game parameterizations. Values in the first three rows are calculated by plugging the respective values of g, l, and p into the critical discount factor $\delta^*(p, g, l)$. The predictions for the effects of the gain and the loss are calculated by taking the differences in rows (2) − (1) and (3) − (1), respectively. The prediction for the differences in the effects is calculated by taking (4) − (5).

game payoffs. Here, the theory predicts that participants (a) in LowBelief react more strongly to the loss than to the gain, (b) in MedBelief react equally to both, and (c) in HighBelief react more strongly to the gain than to the loss.

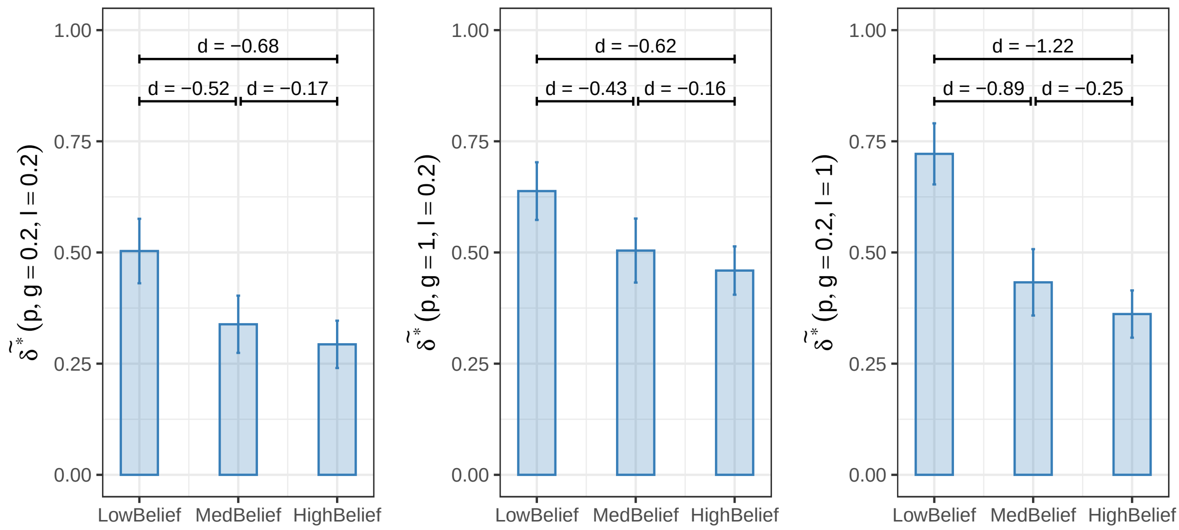
5.3 Experiment: Results

We begin the analysis of the experimental data by showing how the belief treatments affect participants' critical discount factors across stage game parameterizations. Based on this, we show how changes in the critical discount factors, caused by changes in gain and loss, differ in the belief treatments.

Beliefs and Cooperation In the experiment, we elicit participants' critical discount factors in different infinitely repeated prisoner's dilemmas that differ by the stage game parameterizations. Figure 4 reports the average critical discount factors in all treatments, separated into subfigures by the stage game parameterization.

Across the board, increasing the belief has a large negative effect on the reported critical discount factors, and, thus, a large positive effect on cooperation. If participants have a higher belief, they report lower critical discount factors, i.e., they cooperate more. Depending on the stage game parameterization, this effect can be very large, up to an effect size of $d = -1.22$ (LowBelief vs. HighBelief in HighLoss). Furthermore, there is a decreasing marginal return, as predicted by the theory. The difference in critical discount

Figure 4: Belief-Treatments and Cooperation



(a) Base (b) HighGain (c) HighLoss

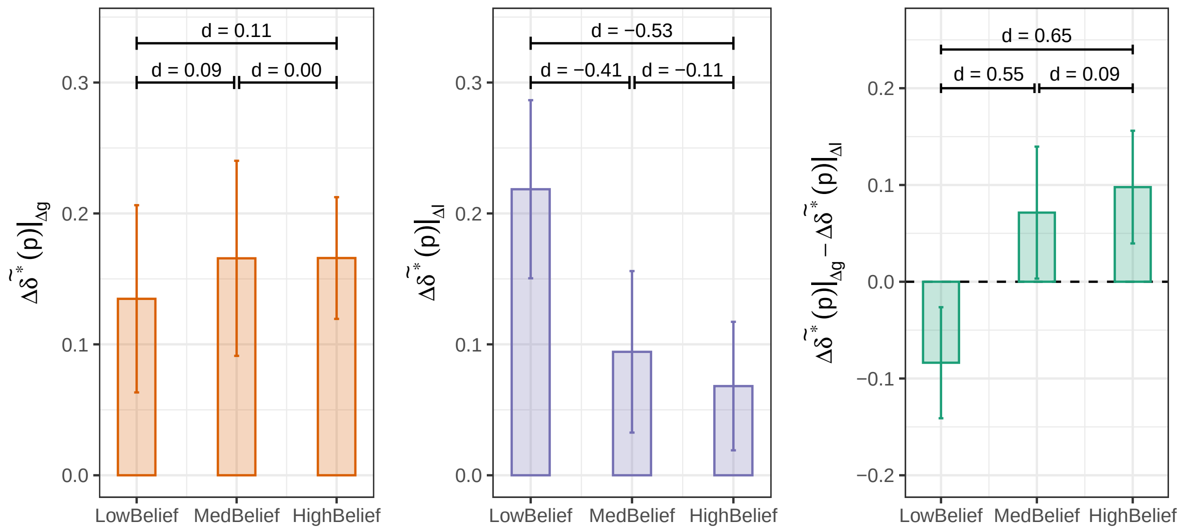
Note: All figures report means, 95% confidence intervals, effect size as Cohen's d. See Table 6 in the appendix for corresponding means, standard deviations, and p-values of the treatment differences.

factors between LowBelief and MedBelief is larger than between MedBelief and HighBelief. To our knowledge, the positive effect of beliefs on cooperation has not been shown before in the context of the infinitely repeated prisoner's dilemma using *exogenous* variation in beliefs.

Payoffs, Beliefs, and Cooperation Following the proposition, we hypothesized that the effect of the gain on the critical discount factor increases in the belief, and that the effect of the loss decreases in the belief. Thus, in general, we expected participants with low beliefs to react more strongly to the loss than to the gain and participants with high beliefs to react more strongly to the gain than to the loss. Figure 5 shows the average effects of (5a) the gain on the critical discount factor, (5b) the loss on the critical discount factor, and (5c) the differences between the two effects across belief treatments.

We see that the effect of the gain increases in the belief as predicted, but only between LowBelief and MedBelief (H_{Gain}) while there is no difference between MedBelief and HighBelief. Further, we find a large negative effect of the belief on the effect of the loss. This effect is stronger between LowBelief and MedBelief than between MedBelief and HighBelief (H_{Loss}). If we compare the effects of the gain and the loss, there is a clear picture, see Figure 5c. Participants in LowBelief react more strongly to the loss than to the gain, participants in HighBelief react more strongly to the gain than to the loss ($d = 0.65$ for LowBelief vs. HighBelief). Already in MedBelief,

Figure 5: The Effects of Gain and Loss, Conditional on the Belief-Treatments



(a) Effect of the Gain (b) Effect of the Loss (c) Difference in the Effects

Note: All figures report means, 95% confidence intervals, effect size as Cohen's d. See Table 7 in the appendix for corresponding means, standard deviations, and p-values of the treatment differences.

participants seem to react more strongly to the gain than to the loss.[14,15]

Thus, the data from the laboratory experiment are largely consistent with our theoretical proposition. They show that the effect of payoff parameters on cooperation as measured with the elicited critical discount factors is moderated by beliefs in the predicted way. However, for the HighBelief case, the effect of the gain is substantially smaller than predicted. In the next subsection, we discuss a potential counteracting effect that offers an explanation.

5.4 Experiment: Discussion

In the experiment, we found that the effect of the gain on the critical discount factor is smaller than expected in the treatment HighBelief, that is, when participants know that their counterpart will choose a cooperative strategy in 90% of the cases. This coincides with a closely related observation in the meta study presented in the previous section, where the increase in defection rates with the estimated belief about the counterpart's cooperation does not extend to beliefs higher than 50% (recall Figure 3a).

[14]When comparing the mean difference in the effects of the gain and the loss to zero within each treatment, we find an effect-size of $d = -0.30$ ($p < 0.01$, two-sided t-test) in LowBelief, an effect-size of $d = 0.25$ ($p = 0.04$, two-sided t-test) in MedBelief, and an effect-size of $d = 0.35$ ($p < 0.01$, two-sided t-test) in HighBelief.

[15]In the Online Appendix, we show results of further preregistered analyses, i.e. OLS estimates with the belief in linear and squared form as well as control variables from a post-experimental questionnaire, a manipulation check and descriptive text-mining based analysis of open-form written statements by the participants. The results presented there are in line with the results presented in the main text and provide additional support for the theory.

We propose that this is due to social norms counteracting the moderating effect of the belief on the effect of the gain if players hold a high belief about the other's willingness to cooperate. It lies in the very definition of a social norm, that it is "jointly recognized" by the players (Krupka and Weber, 2013, p. 498), which implies that a cooperation norm becomes relevant for relatively high beliefs only. More explicitly, we know from previous literature that the willingness to comply with a social norm depends on players' beliefs about others' norm compliance (Krupka et al., 2017; McBride and Ridinger, 2021). Here, the cooperation norm induces players to cooperate when they expect the other player to cooperate, no matter how attractive a deviation would be in terms of their own payoffs. This counteracts the effect of the belief on the effect of the gain, because reacting to the gain by defecting against a cooperating player would violate the cooperation norm.

6 Conclusion

This paper has shown that beliefs determine the effect size of changes in the incentive structure on cooperation in social dilemmas. With this, we contribute a more comprehensive understanding of the determinants of cooperation in social dilemmas to the literature: Equilibrium selection is context-dependent, and players' beliefs need to be considered beyond their direct effect on cooperation.

To show our results theoretically, we derived a novel critical discount factor that allows for varying beliefs. We show that the negative effect of the gain from unilateral defection on cooperation increases in the belief that the other player cooperates, whereas the negative effect of the loss from unilateral cooperation decreases in the belief. Thus, players with a low belief react more strongly to the loss than to the gain, whereas players with a high belief react more strongly to the gain than to the loss.

We provided two empirical tests of the theoretical proposition: First, we conducted a meta study based on Dal Bó and Fréchette (2018) and Fudenberg and Karreskog Rehbinder (2024). Using this data, we constructed players' beliefs from players' experiences in previous interactions and compared the effects of variations in the payoff parameters across these constructed beliefs. Second, we conducted a laboratory experiment that was designed to closely resemble the theory. In this experiment, we varied the beliefs exogenously and independently of the variations in the payoff parameters. Both empirical tests have shown clearly that, indeed, players with a low belief react more strongly to the loss than to the gain, and that players with a high belief react more strongly to the gain than to the loss. These results are in line with the theoretical proposition.

Our results imply that policy interventions that aim to change the incentive structure of a social dilemma should take into account the context of the strategic interaction. In a rather trustworthy context, players find that their opponent is likely to cooperate, and thus changes in the gain from unilateral defection are more effective than changes in the loss from unilateral cooperation. However, in a rather untrustworthy context, changes in the gain from unilateral defection are much less effective than changes in the loss from unilateral cooperation. This should be considered when designing interventions into social dilemmas.

A Appendix

A.1 Payoff-Dependent Beliefs

In the main text, we have followed the previous theoretical literature in assuming that beliefs are exogenously given. Here, we show how changes in the payoff parameters affect the critical discount factor if beliefs are a function of the payoff parameters g and l. We consider beliefs "dependent" if players form their beliefs only after they have seen the game's payoff matrix and then form their beliefs depending on the payoffs. To our knowledge, this has not been considered theoretically in the literature, although very recent experimental literature finds evidence for this. Players have lower beliefs for payoff parameters that are less conducive to cooperation (Aoyagi et al., 2024; Andres, 2024). Based on this empirical result, we define the belief p as a continuous function of the payoff parameters g and l, without assuming a specific functional form.

Definition. *Let* $\Gamma(g, l)$ *be an infinitely repeated prisoner's dilemma of two players with strategy set* $\omega \equiv \{Grim, AD\}$*, discount factor* $\delta \in (0, 1)$ *and domain* $\{g, l \in \mathbb{R} | g, l > 0 \wedge l > g - 1\}$*. Assume that players form their beliefs at the inception of the game after observing the payoff matrix. Furthermore, assume that players' beliefs decrease in* g *and* l*. Then the continuous function* $p : D \mapsto R$ *with domain* $D \equiv \{g, l \in \mathbb{R} | g, l > 0 \wedge l > g - 1\}$*, range* $R \in \mathbb{R}$*, image* $I = (0, 1)$*, and first derivatives* $\frac{\partial p(g,l)}{\partial g} < 0$ *and* $\frac{\partial p(g,l)}{\partial l} < 0$ *is called the dependent belief* $p(g, l)$*.*[16]

This definition has two implications. First, it is now meaningless to ask how the effects of the payoff parameters on cooperation depend on the belief because the belief itself depends on the payoff parameters. Second, increases in the payoff parameters have a greater negative effect on cooperation if beliefs are dependent on g and l than if they were exogenous. This is because two effects add up, a direct effect and an indirect effect. Increases in the payoff parameters have a *direct* negative effect on cooperation because they make defection more tempting or cooperation more risky, as in the proposition. But they also have an *indirect* effect: The belief decreases in both payoff parameters, while cooperation increases in the belief. Thus, an increase in either payoff parameter leads to a decrease in the belief, which leads to an additional decrease in cooperation.

Let us show this formally. Let players have dependent beliefs $p(g, l) \in (0, 1)$. For comparison, assume that there exists an exogenous p^{exo} such that $p(g, l) = p^{exo}$. Deriving

[16]This implies further that the second derivatives must be positive because the limits must be $\lim_{g\to\infty} p(g, l) = \lim_{l\to\infty} p(g, l) = 0$ as the belief cannot fall below zero. Thus, there must be a decreasing marginal effect of the parameters on the belief if $p(g, l)$ is continuous and twice differentiable.

$\delta^*\big(p(g,l),g,l\big)$ with respect to g and l yields:

$$\frac{\partial\delta^*\big(p(g,l),g,l\big)}{\partial g} = \frac{\big(p(g,l)\big)^2 - l\cdot\frac{\partial p(g,l)}{\partial g}}{\Big(l - l\cdot p(g,l) + p(g,l) + g\cdot p(g,l)\Big)^2} \tag{24}$$

$$\frac{\partial\delta^*\big(p(g,l),g,l\big)}{\partial l} = \frac{p(g,l) - \big(p(g,l)\big)^2 - l\cdot\frac{\partial p(g,l)}{\partial l}}{\Big(l - l\cdot p(g,l) + p(g,l) + g\cdot p(g,l)\Big)^2} \tag{25}$$

In both derivatives, the denominators are clearly positive because they are squared. By assumption, namely by $p(g,l)\in(0,1)$, $l>0$ and $\frac{\partial p(g,l)}{\partial g}, \frac{\partial p(g,l)}{\partial l} < 0$, the nominators in both derivatives are positive. Thus, the critical discount factor $\delta^*\big(p(g,l),g,l\big)$ increases in both g and l, also if the belief is dependent.

Because $p(g,l)\in(0,1)$, $l>0$ and $\frac{\partial p(g,l)}{\partial g}, \frac{\partial p(g,l)}{\partial l} < 0$, the terms $-l\cdot\frac{\partial p(g,l)}{\partial g}$ and $-l\cdot\frac{\partial p(g,l)}{\partial l}$ in Equations (24) and (25), respectively, are positive. Thus,

$$\frac{\partial\delta^*\big(p(g,l),g,l\big)}{\partial g} > \frac{\partial\delta^*\big(p^{exo},g,l\big)}{\partial g} > 0 \tag{26}$$

$$\frac{\partial\delta^*\big(p(g,l),g,l\big)}{\partial l} > \frac{\partial\delta^*\big(p^{exo},g,l\big)}{\partial l} > 0 \tag{27}$$

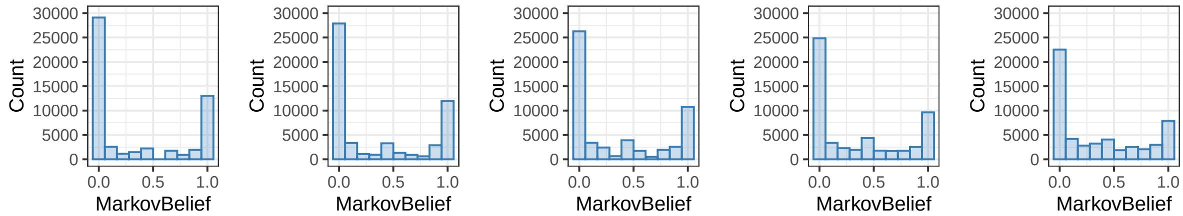
Thus, with dependent beliefs $p(g,l)$, the derivatives of $\delta^*\big(p(g,l),g,l\big)$ with respect to g and l are positive, and larger than the derivatives of $\delta^*\big(p^{exo},g,l\big)$ with respect to g and l, respectively. This difference increases in the size of the first derivatives of $p(g,l)$ with respect to g and l, respectively.

Thus, if beliefs are a function of the payoff parameters, it is still possible to affect cooperation in the infinitely repeated prisoner's dilemma via changes in the payoffs. The effect of changes in the payoff parameters will even be greater than if the beliefs were exogenous.

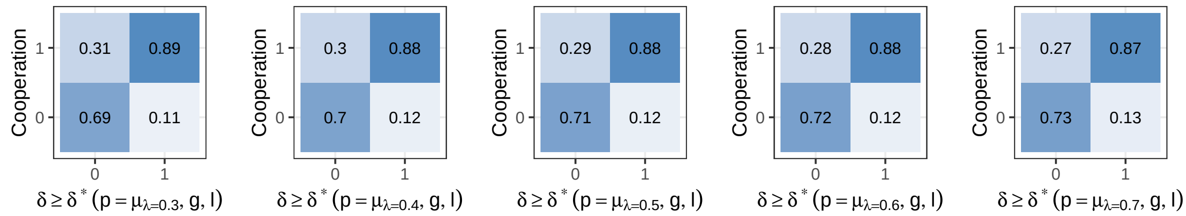
A.2 Additional Details to the Meta Study

A.2.1 Validation of the Belief Construction

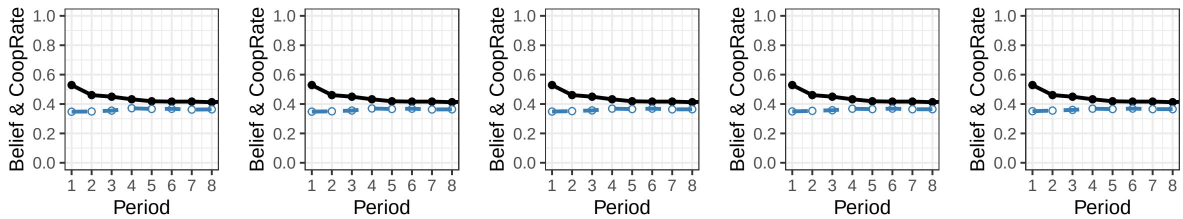
Figure 6: Validation of Markov Beliefs for Different λ



(a) Distribution of Beliefs



(b) Predicting Cooperation



(c) Accuracy of Beliefs

Note: Figure 6a presents the distribution of the Markov beliefs in the first period. Figure 6b presents the confusion matrix of predicted cooperation vs. actual cooperation with column percentages depicted by color intensity. Figure 6c presents average beliefs (dashed) and the average cooperation rate (solid) over the first eighth periods, across supergames. All figures display Markov beliefs estimated with $\lambda \in \{0.3, 0.4, 0.5, 0.6, 0.7\}$, from left to right.

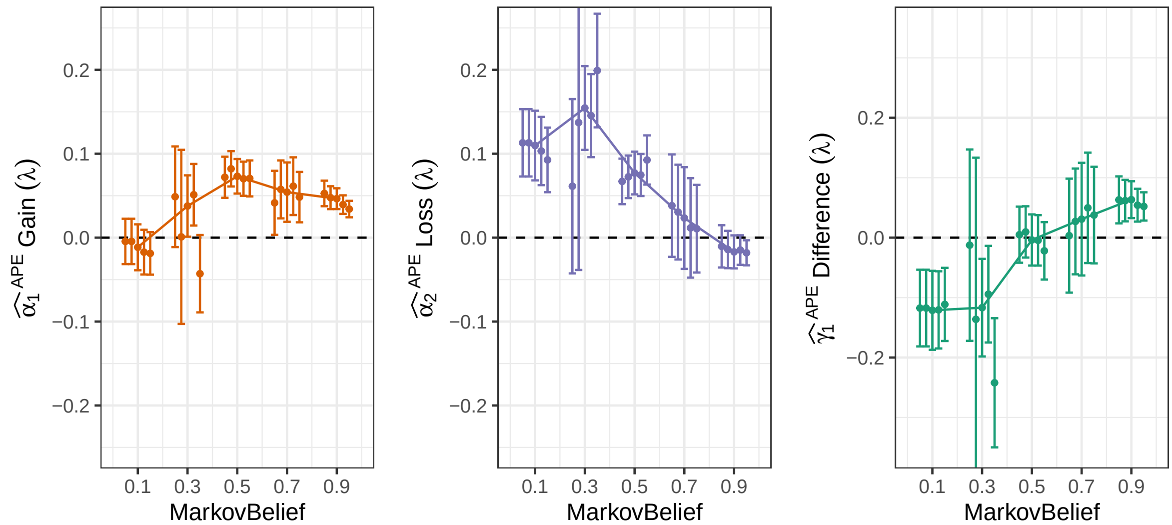
A.2.2 Results

Table 5: Estimated Average Partial Effects in Figure 3

			$M_1^{g,l}$		
$\widehat{\alpha_1}^{APE}$ Gain	-0.0115	0.0377	0.0731	0.0543	0.0465
	(0.0140)	(0.0187)	(0.0105)	(0.0180)	(0.0063)
	[0.4111]	[0.0434]	[0.0000]	[0.0026]	[0.0000]
$\widehat{\alpha_2}^{APE}$ Loss	0.1098	0.1545	0.0770	0.0234	-0.0170
	(0.0212)	(0.0255)	(0.0130)	(0.0309)	(0.0100)
	[0.0000]	[0.0000]	[0.0000]	[0.4498]	[0.0915]
			M_1^{Δ}		
$\widehat{\gamma_1}^{APE}$ Difference	-0.1212	-0.1168	-0.0039	0.0309	0.0634
	(0.0336)	(0.0415)	(0.0218)	(0.0479)	(0.0157)
	[0.0003]	[0.0049]	[0.8584]	[0.5196]	[0.0001]
$\widehat{\gamma_1}^{APE}$ Gain + Loss	0.1098	0.1545	0.0770	0.0234	-0.0170
	(0.0212)	(0.0255)	(0.0130)	(0.0309)	(0.0100)
	[0.0000]	[0.0000]	[0.0000]	[0.4498]	[0.0915]
N	30,091	2,122	6,114	3,919	11,930
Belief Range	$[0, 0.2]$	$(0.2, 0.4]$	$(0.4, 0.6]$	$(0.6, 0.8]$	$(0.8, 1]$

Note: Average partial effects of Probit regressions with defection as the dependent variable. Standard errors, clustered at the individual participant level, in parentheses with p-values in brackets.

Figure 7: The Effects of Gain and Loss, Conditional on the Markov Belief for Different λ



(a) Effect of the Gain (b) Effect of the Loss (c) Difference in the Effects

Note: All figures report average partial effects on the probability of defection and 95% confidence intervals following the Probit models $M_1^{g,l}$ and M_1^{Δ}, for different λ. Point estimates for, from left to right, $\lambda \in \{0.3, 0.4, 0.5, 0.6, 0.7\}$. Standard errors are clustered at the individual participant level. The coefficients are displayed around the center of the respective bracket: $[0, 0.2], (0.2, 0.4], (0.4, 0.6], (0.6, 0.8], (0.8, 1]$.

A.3 Additional Details to the the Experiment

A.3.1 Instructions

Instructions

These instructions are identical for all participants.

Welcome to this experiment!

You will receive a monetary compensation for participating in this experiment. The amount you receive depends on your decisions and the decisions of other participants. It is therefore important that you read the instructions on the following pages carefully. Please take enough time to do this – the decision-making environment in this experiment requires a detailed explanation, which we will guide you through step by step.

For the entire duration of the experiment, you are not allowed to communicate with other participants. We therefore ask you not to talk to each other. Please also refrain from using your cell phones. Violation of these rules will result in exclusion from the experiment and payment.

If you have a question, please give us a hand signal. We will then come to you and answer your question personally.

During the experiment, we do not talk of euros, but of points. Your total income will initially be calculated in points. Your score will be converted into euros at the end of the experiment, using the following conversion rate:

18 points = 1 euro

At the end of today's experiment, you will receive the points you have scored from the experiment converted into euros in cash. In addition, you will receive 8 euros today for being on time for the experiment. The payment procedure is organized in such a way that the other participants will not see the amount you receive.

The experiment consists of two parts:

- In the first part, you can familiarize yourself with the experiment. This part has no influence on your payout. We call this part the test phase.
- In the second part, the actual experiment, you will make various decisions. These decisions will determine which payout you receive.

We will then ask you to complete a short questionnaire. You will then receive your payout in cash.

The decision situation

The decisions you will make in the experiment concern different versions of a specific decision situation. We would like to present these to you first.

You and another person make a decision between actions X and Y at the same time. The other person will later be assigned to you by the computer program. Neither you nor the other person will know anything about the identity of the other person.

Your payout depends on what you choose and what the other person chooses. The payouts are shown in the following table:

Your action	Action of the other person	Your payout	Payout of the other person
X	X	**75**	75
Y	X	**80**	45
X	Y	**45**	80
Y	Y	**50**	50

The payouts in the decision situations you face in the experiment will partially differ from those in this table. However, you and the other person will always receive the same payout if you both choose the same action, and different payouts if you choose different actions.

Many repetitions

When you and the other person make the decision described above in the experiment, you don't just do it once, but many times in succession. After each decision, you find out what the other person has done. The payouts that we have described above become smaller and smaller from round to round. This happens very evenly: After each round, all four possible payouts are multiplied by a number that we call the residual factor.

The residual factor is a number between 10% and 90%. The higher the residual factor, the more is left over after each round. The smaller the residual factor, the faster the payouts shrink. With a residual factor of 90%, the four payouts from the table above shrink in each round to 90% of the value in the previous round. With a residual factor of 50%, they are only half as large in each round as in the previous round.

On the next page we show you an overview of how this shrinkage looks for the value 50, depending on the residual factor. This is shown for the first 30 rounds. In the table you can see that the payouts become very small at some point.

Round (1-15)

Residual-Factor	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
90%	50	45	40.5	36.45	32.81	29.52	26.57	23.91	21.52	19.37	17.43	15.69	14.12	12.71	11.44
80%	50	40	32	25.6	20.48	16.38	13.11	10.49	8.39	6.71	5.37	4.29	3.44	2.75	2.2
70%	50	35	24.5	17.15	12.01	8.4	5.88	4.12	2.88	2.02	1.41	0.99	0.69	0.48	0.34
60%	50	30	18	10.8	6.48	3.89	2.33	1.4	0.84	0.5	0.3	0.18	0.11	0.07	0.04
50%	50	25	12.5	6.25	3.13	1.56	0.78	0.39	0.2	0.1	0.05	0.02	0.01	0.01	0
40%	50	20	8	3.2	1.28	0.51	0.2	0.08	0.03	0.01	0.01	0	0	0	0
30%	50	15	4.5	1.35	0.41	0.12	0.04	0.01	0	0	0	0	0	0	0
20%	50	10	2	0.4	0.08	0.02	0	0	0	0	0	0	0	0	0
10%	50	5	0.5	0.05	0.01	0	0	0	0	0	0	0	0	0	0

Round (16-…)

Residual-Factor	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**…**
90%	20.59	18.53	16.68	15.01	13.51	12.16	10.94	9.85	8.86	7.98	7.18	6.46	5.81	5.23	4.71	…
80%	3.52	2.81	2.25	1.8	1.44	1.15	0.92	0.74	0.59	0.47	0.38	0.3	0.24	0.19	0.15	…
70%	0.47	0.33	0.23	0.16	0.11	0.08	0.06	0.04	0.03	0.02	0.01	0.01	0.01	0	0	…
60%	0.05	0.03	0.02	0.01	0.01	0	0	0	0	0	0	0	0	0	0	…
50%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	…
40%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	…
30%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	…
20%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	…
10%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	…

[In the instructions, this table filled an entire page in landscape.]

Deciding on plans, not actions

The decisions between action X and action Y in the many repetitions are not all made individually in the experiment. Instead, at the beginning of a series of repetitions, you choose a plan once, as you would like to decide in the many rounds. You have two plans to choose from: Plan A and Plan B.

Plan A stipulates that you choose action X in the first round. If the other person also chooses action X in the first round, you stick with action X in the subsequent rounds. If the other person chooses action Y in the first round, you choose action Y from the second round onwards.

Plan B is that you always choose action Y, regardless of what the other person does. You start with action Y and choose action Y in every round.

As you and the other person both choose between plan A and plan B, there are only four ways in which the decisions can be made in the many repetitions:

Your plan	Plan of the other person	Progression of the actions in the many repetitions
A	A	In the first round and in all subsequent rounds, you both select action X.
B	A	In the first round, you choose action Y and the other person chooses action X. From the second round onwards, you both choose action Y for all subsequent rounds.
A	B	In the first round, you choose action X and the other person chooses action Y. From the second round onwards, you both choose action Y for all subsequent rounds.
A	B	In the first round and in all subsequent rounds, you both select action Y.

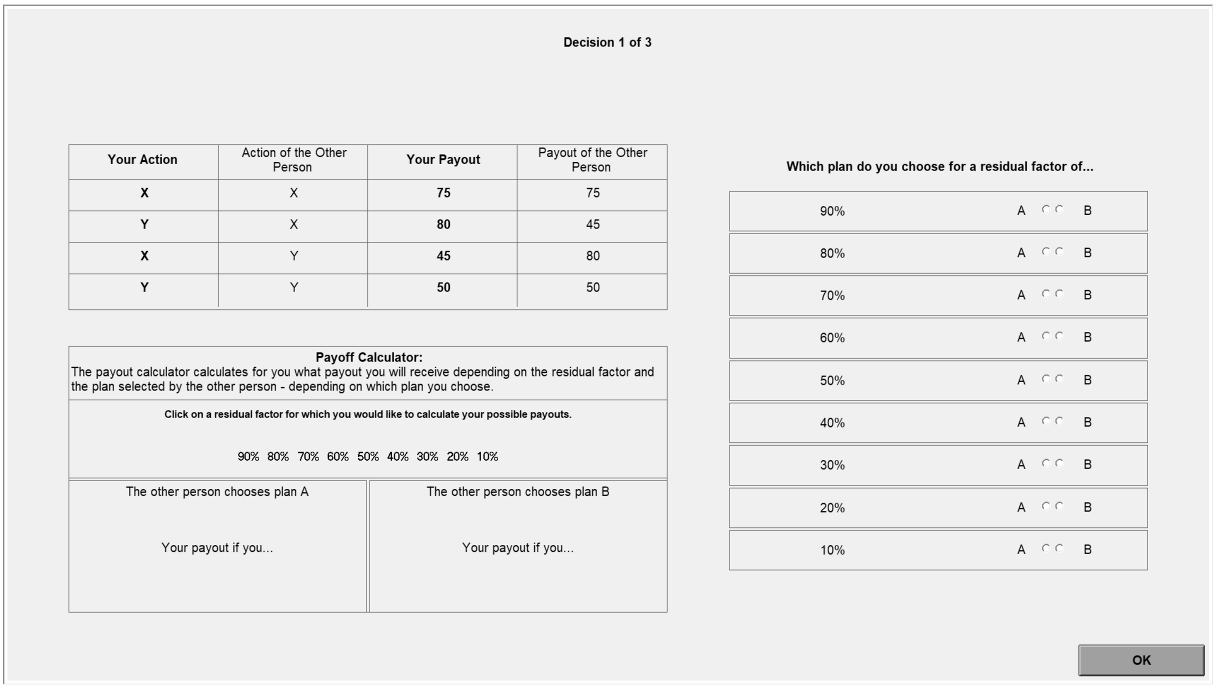
In the experiment, you never decide between actions X and Y in individual rounds, but only ever decide between plan A and plan B for the entire duration of all repetitions. The computer program uses this to recognize how you and the other person's decisions will play out and calculates your payout directly from this.

To do this, the computer program assumes mathematically that there are an infinite number of repetitions of the decision situation. As you can see in the table, however, the payouts per round shrink to very small values after relatively few rounds, so that a large part of your total payout is determined in the first repetitions.

You make the decision between plan A and plan B several times in the experiment – for a total of three different decision situations, which differ in the the payouts, and in each case for the residual factors 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20% and 10%.

For each of the remaining factors, you specify whether you choose plan A or plan B. You may only switch between the plans once per decision situation. This means: If, for example, you want to choose plan A for the residual factor 40% and plan B for the residual factor 30%, you must also choose plan A for the residual factors 90% to 50%; and you must also choose plan B for the residual factors 20% to 10%. You can also select the same plan for all residual factors.

Here we show you a screenshot of the first decision situation on the computer.



In the experiment, you will have a payout calculator at your disposal that will show you how high the total payout will be over all rounds – depending on the residual factor and which plan you and the other person choose.

Payout of a single decision

In the experiment, you make the decision between plan A and plan B for three different versions of the decision situation. For each of these three versions, you decide for each of the nine different residual factors between the two plans. This means that you make a total of $3 \cdot 9 = 27$ decisions in the experiment. The payout you receive for your participation in the experiment is determined by exactly one of these 27 decisions. Which one this is is decided by a random mechanism of the computer, which gives all 27 decisions the same probability of being drawn as a decision relevant to the payout. Your payout corresponds to your points from this decision converted into euros. You will receive this payout in addition to the 8 euros you receive for showing up on time for the experiment.

This means that each of your decisions can be the one that determines your payout! You will only be told which one at the end of the experiment.

What do you learn about the other person?

The other person, whose decision together with your own decision determines your payout, will only be assigned to you after you have made your payout-relevant decision. You initially make all your decisions between Plan A and Plan B on your own. The allocation is based on the decisions made by the other people.

At the beginning of the experiment, you are told how likely it is that you will be assigned to a person who has chosen plan A in the decision relevant to your payout, or a person who has chosen plan B. This probability is then always the same for all decisions and does not change.

You therefore know from the outset for all decisions with what probability the other person who is ultimately assigned to you for the calculation of your payout will have chosen plan A or plan B. This probability does not depend on your own decision.

The payouts for the other people are calculated in exactly the same way. This means that for all other people, too, one of the 27 decisions is initially selected by a random computer mechanism. After the decision, each of the other people – just like for you – is assigned a second person whose decision for plan A or plan B determines the first person's payout. It is therefore not unlikely that the person who is assigned to you for the calculation of your payout will not be assigned to you for the calculation of this person's payout, but to another person from today's experiment.

During this procedure, it may happen that there is no person who has opted for the plan to be assigned to you. In such a case, the computer takes over the role of the other person and selects the corresponding plan so that your payout is also calculated normally.

Test phase

At the beginning of the experiment, all participants can familiarize themselves with the display of the decision on the computer screen. In the test phase, you will be shown the screen for the first of the three versions of the decision situation.

In the test phase, quiz questions are asked on the screen. The experiment only begins when all participants have answered all the quiz questions correctly. Your answers to the quiz questions have no consequences for your payout at the end of the experiment!

If you have a question, please give us a hand signal. We will then come to you and answer your question personally.

A.3.2 Additional Tables

Table 6: Effect of the Belief on the Critical Discount Factor

	LowBelief	MedBelief	HighBelief
Base	0.50 (0.35)	0.34 (0.27)	0.29 (0.26)
	LowBelief vs. MedBelief	$p < 0.01$	$d = -0.52$
	MedBelief vs. HighBelief	$p = 0.14$	$d = -0.17$
	LowBelief vs. HighBelief	$p < 0.01$	$d = -0.68$
HighGain	0.64 (0.31)	0.50 (0.30)	0.46 (0.26)
	LowBelief vs. MedBelief	$p < 0.01$	$d = -0.43$
	MedBelief vs. HighBelief	$p = 0.16$	$d = -0.16$
	LowBelief vs. HighBelief	$p < 0.01$	$d = -0.62$
HighLoss	0.72 (0.33)	0.43 (0.31)	0.36 (0.26)
	LowBelief vs. MedBelief	$p < 0.01$	$d = -0.89$
	MedBelief vs. HighBelief	$p = 0.06$	$d = -0.25$
	LowBelief vs. HighBelief	$p < 0.01$	$d = -1.22$

Note: The table shows means of the elicited critical discount factors across stage game parameterizations (rows) treatments (columns), with standard deviations in parentheses. The remaining rows report p-values of one-sided t-tests for the respective treatment comparisons and the corresponding effect-size as Cohen's d.

Table 7: Effect of the Belief on the Effect of the Gain and the Loss

	LowBelief	MedBelief	HighBelief
Effect of the gain:	0.14 (0.35)	0.17 (0.31)	0.17 (0.22)
	LowBelief vs. MedBelief	$p = 0.28$	$d = 0.09$
	MedBelief vs. HighBelief	$p = 0.50$	$d = 0.00$
	LowBelief vs. HighBelief	$p = 0.23$	$d = 0.11$
Effect of the loss:	0.22 (0.33)	0.09 (0.26)	0.07 (0.24)
	LowBelief vs. MedBelief	$p < 0.01$	$d = -0.41$
	MedBelief vs. HighBelief	$p = 0.25$	$d = -0.11$
	LowBelief vs. HighBelief	$p < 0.01$	$d = -0.53$
Differences:	-0.08 (0.28)	0.07 (0.29)	0.10 (0.28)
	LowBelief vs. MedBelief	$p < 0.01$	$d = 0.55$
	MedBelief vs. HighBelief	$p = 0.28$	$d = 0.09$
	LowBelief vs. HighBelief	$p < 0.01$	$d = 0.65$

Note: The table shows means of the effects on the elicited critical discount factors across stage game parameterizations (rows) treatments (columns), with standard deviations in parentheses. The remaining rows report p-values of one-sided t-tests for the respective treatment comparisons and the corresponding effect-size as Cohen's d.

References

Andreoni, J. and J. H. Miller (1993). Rational cooperation in the finitely repeated prisoner's dilemma: Experimental evidence. *The Economic Journal 103*(418), 570–585.

Andres, M. (2024). Equilibrium selection in infinitely repeated games with communication. *Available at SSRN 4759938*.

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

Aoyagi, M., G. R. Fréchette, and S. Yuksel (2024). Beliefs in repeated games: An experiment. *American Economic Review (forthcoming).*

Arechar, A. A., M. Kouchaki, and D. G. Rand (2018). Examining spillovers between long and short repeated prisoner's dilemma games played in the laboratory. *Games 9*(1), 5.

Athey, S., K. Bagwell, and C. Sanchirico (2004). Collusion and price rigidity. *The Review of Economic Studies 71*(2), 317–349.

Baader, M., S. Gächter, K. Lee, and M. Sefton (2024). Social preferences and the variability of conditional cooperation. *CeDEx Discussion Paper Series, No. 2024-04*.

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*(3), 164–192.

Blonski, M. and G. Spagnolo (2015). Prisoners' other dilemma. *International Journal of Game Theory 44*(1), 61–81.

Boczoń, M., E. Vespa, T. Weidman, and A. J. Wilson (2024). Testing models of strategic uncertainty: Equilibrium selection in repeated games. *Journal of the European Economic Association*, jvae042.

Breitmoser, Y. (2015). Cooperation, but no reciprocity: Individual strategies in the repeated prisoner's dilemma. *American Economic Review 105*(9), 2882–2910.

Bruttel, L. (2009). The critical discount factor as a measure for cartel stability? *Journal of Economics 96*, 113–136.

Bruttel, L. and U. Kamecke (2012). Infinity in the lab: How do people play repeated games? *Theory and Decision 72*(2), 205–219.

Cooper, R., D. V. DeJong, R. Forsythe, and T. W. Ross (1996). Cooperation without reputation: Experimental evidence from prisoner's dilemma games. *Games and Economic Behavior 12*(2), 187–218.

Costa-Gomes, M. A., S. Huck, and G. Weizsäcker (2014). Beliefs and actions in the trust game: Creating instrumental variables to estimate the causal effect. *Games and Economic Behavior 88*, 298–309.

Dal Bó, P. and G. R. Fréchette (2019). Strategy choice in the infinitely repeated prisoner's dilemma. *American Economic Review 109*(11), 3929–3952.

Dal Bó, P. (2005). Cooperation under the shadow of the future: Experimental evidence from infinitely repeated games. *American Economic Review 95*(5), 1591–1604.

Dal Bó, P., A. Foster, and L. Putterman (2010). Institutions and behavior: Experimental evidence on the effects of democracy. *American Economic Review 100*(5), 2205–2229.

Dal Bó, P. and G. R. Fréchette (2011). The evolution of cooperation in infinitely repeated games: Experimental evidence. *American Economic Review 101*(1), 411–429.

Dal Bó, P. and G. R. Fréchette (2018). On the determinants of cooperation in infinitely repeated games: A survey. *Journal of Economic Literature 56*(1), 60–114.

Dal Bó, P. and G. R. Fréchette (2019). Strategy choice in the infinitely repeated prisoner's dilemma. *American Economic Review 109*(11), 3929–3052.

Dreber, A., D. G. Rand, D. Fudenberg, and M. A. Nowak (2008). Winners don't punish. *Nature 452*(7185), 348–351.

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

Dvorak, F. and S. Fehrler (2024). Negotiating cooperation under uncertainty: Communication in noisy, indefinitely repeated interactions. *American Economic Journal: Microeconomics 16*(3), 232–258.

Embrey, M., G. R. Fréchette, and S. Yuksel (2018). Cooperation in the finitely repeated prisoner's dilemma. *The Quarterly Journal of Economics 133*(1), 509–551.

Friedman, D. and B. Sinervo (2016). *Evolutionary games in natural, social, and virtual worlds*. Oxford University Press.

Friedman, J. W. (1971). A non-cooperative equilibrium for supergames. *The Review of Economic Studies 38*(1), 1–12.

Fréchette, G. R. and S. Yuksel (2017). Infinitely repeated games in the laboratory: Four perspectives on discounting and random termination. *Experimental Economics 20*(2), 279–308.

Fudenberg, D. and G. Karreskog Rehbinder (2024). Predicting cooperation with learning models. *American Economic Journal: Microeconomics 16*(1), 1–32.

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

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*(2), 720–749.

Gächter, S., K. Lee, M. Sefton, and T. O. Weber (2024). The role of payoff parameters for cooperation in the one-shot prisoner's dilemma. *European Economic Review 166*, 104753.

Gill, D. and Y. Rosokha (2024). Beliefs, learning, and personality in the indefinitely repeated prisoner's dilemma. *American Economic Journal: Microeconomics 16*(3), 259–283.

Gilo, D., Y. Moshe, and Y. Spiegel (2006). Partial cross ownership and tacit collusion. *RAND Journal of Economics 37*(1), 81–99.

Harsanyi, J. C. and R. Selten (1988). *A general theory of equilibrium selection in games.* The MIT Press.

Honhon, D. and K. Hyndman (2020). Flexibility and reputation in repeated prisoner's dilemma games. *Management Science 66*(11), 4998–5014.

Huck, S., H.-T. Normann, and F. Petros (2024). Concerns about rising prices may raise prices. *WZB Discussion Paper, No. SP II 2024-203*.

Kagel, J. H. and D. R. Schley (2013). How economic rewards affect cooperation reconsidered. *Economics Letters 121*(1), 124–127.

Krupka, E. L., S. Leider, and M. Jiang (2017). A meeting of the minds: Informal agreements and social norms. *Management Science 63*(6), 1708–1729.

Krupka, E. L. and R. A. Weber (2013). Identifying social norms using coordination games: Why does dictator game sharing vary? *Journal of the European Economic Association 11*(3), 495–524.

Martinez-Martinez, I. and H.-T. Normann (2024). Cooperation in multiplayer dilemmas. *Available at SSRN 4294660*.

McBride, M. and G. Ridinger (2021). Beliefs also make social-norm preferences social. *Journal of Economic Behavior & Organization 191*, 765–784.

Mengel, F. (2018). Risk and temptation: A meta-study on prisoner's dilemma games. *The Economic Journal 128*(616), 3182–3209.

Proto, E., A. Rustichini, and A. Sofianos (2019). Intelligence, personality, and gains from cooperation in repeated interactions. *Journal of Political Economy 127*(3), 1351–1390.

Rapoport, A., A. M. Chammah, and C. J. Orwant (1965). *Prisoner's dilemma: A Study in Conflict and Cooperation.* The University of Michigan Press.

Romero, J. and Y. Rosokha (2023). Mixed strategies in the indefinitely repeated prisoner's dilemma. *Econometrica 91*(6), 2295–2331.

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

Sherstyuk, K., N. Tarui, and T. Saijo (2013). Payment schemes in infinite-horizon experimental games. *Experimental Economics 16*(1), 125–153.

Stahl II, D. O. (1991). The graph of prisoners' dilemma supergame payoffs as a function of the discount factor. *Games and Economic Behavior 3*(3), 368–384.