The Importance of Being Even: Restitution and Cooperation*

Maria Bigoni[1], Marco Casari[1], Andrea Salvanti[2], Andrzej Skrzypacz[3], and Giancarlo Spagnolo[4]

[1]University of Bologna
[2]University Pompeu Fabra and BSE
[3]Stanford GSB
[4]University of Rome Tor Vergata & EIEF

April 27, 2024

Abstract

We study – empirically and theoretically – how restitution helps restore cooperation. After a breach, restitution strategies "propose" returning to cooperation by cooperating against defection, and condition actions on the balance between the cooperation given and received. We reanalyze experimental data from three classes of repeated games and find compelling empirical support for restitution strategies in general and for a strategy we named Payback in particular. Considering restitution strategies enables to resolve discrepancies between theory and experiments emerging from prior literature - such as the prevalent use of non-equilibrium strategies like Tit-for-tat - and questions the predominance of memory-one strategies.

Keywords: Asymmetric strategies, laboratory experiments, social dilemmas, indefinitely repeated games.

JEL: C72, C73, C91, D82

*This paper replaces, greatly extends and generalizes a previous note circulated under the title "It's Payback Time: New Insights on Cooperation in the Repeated Prisoners' Dilemma". We are grateful to the editor and the referees of a previous submission for inspiration and useful suggestions, and to Dirk Engelmann, Karl Sigmund, Sigrid Suetens, Robert Sugden, Matthijs van Veelen, and participants at the LEG 2022 workshop, ASFEE 2023, BSE Summer Forum 2023, SMYE 2023, and seminars in Lausanne, Bologna, Humboldt University Berlin, Milan, Konstanz, Amsterdam and Pompeu Fabra for constructive comments. Spagnolo acknowledges funding from the Swedish Research Council (Project 2020-02251).

1 Introduction

Cooperation can be fragile and easily disrupted. In situations where trust has been breached, whether intentionally or unintentionally, rebuilding mutual confidence is essential. Here we focus on strategies for restoring cooperation in social dilemmas after a defection and argue that *Restitution* is a particularly intuitive and effective approach when communication is absent. Restitution involves "proposing" a return to cooperation by restoring gains lost from the breach. This gesture, which is materially costly for the offender and rewarding for the offended party, works as a credible message. Furthermore, restitution can be seen as fair, as it closes the payoff gap created by the deviation and makes subjects even.

While the potential role of restitution in rebuilding cooperation has been discussed in theoretical studies, the experimental literature on repeated games appears to have overlooked it. As far as our knowledge extends, until now it has primarily focused on tit-for-tat and other forgiving or lenient symmetric strategies as effective means of restoring cooperation after a breach (Fudenberg et al., 2012). Moreover, recent experimental studies have increasingly centered on memory-one strategies, which, by definition, exclude forms of restitution as paths to restoring cooperation.[1]

In this study, we investigate the presence of restitution strategies in experiments on three broad classes of games, the repeated prisoner's dilemmas with finite and indefinite repetition and perfect monitoring, and the indefinitely repeated Prisoner's Dilemma with imperfect monitoring. We then focus on a specific, simple restitution strategy – which we term Payback – and we discuss its theoretical properties and empirical performance. We find robust support for restitution strategies in all types of repeated games we consider. We also argue that these strategies have theoretical advantages with respect to the simple memory-one strategies that the experimental literature mainly focused on, which justify their relatively widespread adoption.

Restitution strategies imply asymmetric play in the "punishment" phase following a defection, before resorting again to cooperation. Appealing theoretical features of these strategies have been highlighted in a number of contexts, including renegotiation-proofness in the repeated Prisoner's Dilemma (Van Damme, 1989); repeated games with monetary payments, like relational contracts (Levin, 2003); and tacit collusion in repeated auctions (Skrzypacz and Hopenhayn, 2004). In the realm of pricing games, the superiority of asymmetric strategies over strongly symmetric ones has been established by Athey and Bagwell (2001) and Harrington and Skrzypacz (2007). Evolutionary game theory has also considered asymmet-

[1] See, for example, Breitmoser (2015), Dal Bó and Fréchette (2018), Romero and Rosokha (2018), Dal Bó and Fréchette (2019), Romero and Rosokha (2019), Normann and Sternberg (2023), and Fudenberg and Karreskog (2024).

ric strategies. Sugden (1986) showed that while no pure strategies are evolutionary stable within the Prisoner's Dilemma framework without noise, mistakes in actions enable an asymmetric punishment strategy termed "Sophisticated Tit-for-tat," a restitution strategy, to be evolutionary stable.[2] Asymmetric strategies are not purely theoretical constructs: Harrington (2006), for example, describes several instances of their use by real-life cartels. Despite their theoretical merits, the experimental literature appears to have overlooked restitution strategies and their role in rebuilding cooperation.

To address this gap, we make use of data collected from previous experiments. We rely on three meta-datasets that comprise distinct classes of repeated games. One meta-dataset covers indefinitely repeated prisoner's dilemmas with perfect monitoring (Dal Bó and Fréchette, 2018); another meta-dataset includes finitely repeated Prisoner's Dilemma games (Embrey et al., 2017); the third meta-dataset – which we constructed *ex novo* – incorporates indefinitely repeated games with imperfect monitoring. In all the three databases, we observe that the likelihood of cooperation depends not only on the outcome of the previous round of interaction but also on the round before that. Consider for instance participants' behavior that follows a unilateral defection by the opponent. We observe higher cooperation rates in histories where, before that, the opponent cooperated *and* the participant defected, as compared to any of the other possible histories.

This finding aligns with the adoption of restitution strategies and is incompatible with the strategies that emerged as most prominent in the previous experimental literature – including Grim Trigger (Grim), Tit-for-Tat (TFT), t-period punishment strategies, and their more lenient and forgiving variations. It also suggests that subjects adopt strategies with memory longer than one. Our analyses indicate that "restitution" may occur even after more than two periods: subjects cooperate more frequently after a unilateral defection of the opponent when, based on past outcomes, they were 'in debt' with the opponent at the time when the opponent defected. This suggests that focusing exclusively on memory-one strategies in the analysis of experimental repeated games – as it is often done in the recent literature – may be problematic as it makes it impossible to detect some important features of subjects' behavior.

Restitution strategies tend to balance the number of cooperative actions by the subject and her opponent – which leads subjects to end up about *even* – hence the adoption of this type of strategies may be driven by a heuristic related to fairness and inequity-aversion (Fehr and Schmidt, 1999; Charness and Rabin, 2002), rather than, or besides, strategic considerations. This might explain why restitution strategies in finitely repeated games initially emerge

[2] In this literature, the strategy was also referred to as "Contrite Tit-for-tat". See Boyd (1989), Wu and Axelrod (1995), Boerlijst et al. (1997), and Graser and van Veelen (2024).

at frequencies comparable to that observed in indefinitely repeated games, even though they cannot be equilibrium strategies in the former setting. To shed some light on these issues we analyze the effects of experience comparing behavior in the first and last four supergames in each session. We find that after subjects gain experience, restitution strategies fade away in finite horizon games, while they persist or increase in infinitely repeated games, where they can be part of an equilibrium. This suggests that strategic considerations are the strongest driver of the adoption of restitution strategies, in the long run.

To further deepen our understanding of the role played by restitution strategies, we need to focus on a specific, well-defined one. We concentrate on Contrite (or Sophisticated) TFT, as it was already praised by the evolutionary game theory literature for its resilience in environments with noise, and we re-name it "Payback" to better emphasize the difference with its main "rival" TFT.[3] This strategy prescribes cooperation in the initial period. If a unilateral defection occurs, the player who suffered the deviation should retaliate by defecting, while the original defector should cooperate. Once this repayment phase takes place, both players revert to mutual cooperation.

We study the theoretical and empirical properties of Payback, compared to Grim and TFT, in infinitely repeated games, where cooperation can be sustained in equilibrium. Under perfect monitoring we show that, contrary to TFT, Payback can form a subgame perfect equilibrium. We also prove that, when gains from unilateral deviations are equal to losses from playing cooperation while the opponent defect, Payback is an optimal penal code in the sense of Abreu (1988): it sustains cooperation as a subgame perfect equilibrium for the same set of parameters as Grim, which uses the strongest punishment strategy to discipline deviations. At the same time, it is not subject to renegotiation concerns (in line with Van Damme, 1989).

Games with imperfect monitoring are of particular interest in this perspective, as in these settings the differences between Payback and other common strategies such as Grim and TFT emerge more prominently. Under perfect monitoring, the way off-path punishments are structured is not particularly relevant for the value that the strategy achieves on the equilibrium path (provided they are strong enough). On the contrary, with imperfect monitoring deviations happen on the equilibrium path and symmetric punishments become very costly. As an example, consider two players who wish to coordinate on mutual cooperation, and both start playing C in period 1. With probability E, the action of either of them is changed to D. With Grim, they would immediately switch to mutual defection until the end of the supergame. With TFT, a sequence of asymmetric CD-DC outcomes would follow.[4]

[3] This strategy is among the simplest restitution strategies according to the complexity measures proposed by Oprea (2020), based on the number of states and transitions in the automata representation.

[4] Under standard assumptions on payoffs this is less profitable than remaining on the cooperative path.

The players would be able to revert to mutual cooperation only when a second random error materializes and changes the intended action of one of them from D to C. With Payback, instead, they would be able to revert to mutual cooperation after a single period of asymmetric punishment, unless a second random error takes place. In this sense, Payback is an error-correcting strategy. This intuitive consideration illustrates why as soon as we allow for a positive probability of implementation errors the value produced by two players using Payback is higher than the value produced by using TFT and Grim.

We show that, under imperfect monitoring, Payback has several theoretical advantages: (i) it can be part of a public perfect equilibrium; (ii) it can be an equilibrium for a larger set of parameters compared to Grim; (iii) it achieves a higher value on the equilibrium path and (iv) it is less exposed to strategic risk than Grim or TFT. The latter theoretical result has relevant implications for our understanding of the determinants of cooperation. While many papers provide strong evidence that strategic risk is one of the main predictors of the emergence and sustainability of cooperation in games with perfect monitoring, Fudenberg et al. (2012, p. 733) observed that their data "do not show the strong support for risk dominance *of TFT* as the key determinant of the level of cooperation in games with noise that was seen in studies of games without noise." We show, instead, that risk dominance *of Payback* does emerge as a strong predictor of cooperation rates, also in their set-up with noise. Cooperation rates are almost 20 percentage points higher in the treatments in which Payback is risk-dominant, as compared to the treatments in which it is not.

The theoretical advantages of Payback over TFT and Grim led us to wonder whether this strategy is in fact adopted by subjects. To answer this question, we perform a maximum likelihood estimation of strategies with and without including Payback among the set of strategies. The results show that, in indefinitely repeated games with perfect monitoring, Payback is played by a share of subjects comparable to the one attributed to Grim and TFT; with imperfect monitoring instead the estimated share of Payback play is higher than the one of TFT and Grim, and comparable to that of other more lenient and forgiving strategies. Notably, the introduction of Payback reduces the share assigned to TFT both with perfect and with imperfect monitoring, which suggests that Payback represents a good empirical alternative to TFT in classifying subjects' choices and might explain the sustained rate of cooperation observed in treatments where TFT (and sometimes Grim) are not equilibrium strategies. This can offer a new perspective on a number of other experiments that have reported an important role for TFT in categorizing behavior (for example Dal Bó and Fréchette, 2011; Bigoni et al., 2013; Embrey et al., 2016; Romero and Rosokha, 2018; Dal Bó and Fréchette, 2019; Romero and Rosokha, 2019; Dvorak and Fehrler, 2023; Romero and Rosokha, 2023).

The remainder of the paper is organized as follows. Using meta-data from previous experiments involving repeated Prisoner's Dilemma games, Section 2 provides evidence for the use of a broad class of restitution strategies. Section 3 introduces Payback, which is one of the simplest restitution strategies, and derives theoretical results characterizing its equilibrium and strategic-risk properties. Section 4 presents results from a maximum likelihood estimation of strategies, providing empirical support for the use of Payback. Section 5 concludes.

2 Restitution strategies in three classes of games

In this section, we present our general empirical results obtained using data from previous experiments involving repeated Prisoner's Dilemma games. To ensure comparability across different studies, we have standardized the payoffs as shown in Table 1. Here, the payoff from deviating from the cooperative action profile (C,C), known as the temptation payoff, is represented as $1 + g$, where parameter g represents the additional gain from defection. Conversely, the loss of payoff from cooperating while the opponent defects, referred to as the sucker's payoff, is denoted as $-l$. Both g and l take strictly positive values.[5]

		Player 2: C	Player 2: D
Player 1	C	$1, 1$	$-l, 1+g$
	D	$1+g, -l$	$0, 0$

Table 1: Standardized stage game payoffs for the Prisoner's Dilemma.

We will utilize data from three different meta-datasets, which we refer to as Standard, Noise, and Finite, as detailed in Table 2. Each of these datasets features the Prisoner's Dilemma as the stage game and is structured into sequences of play involving fixed pairs of participants, which we term 'supergames'. However, they differ in the time-horizon of the supergame (which is either indefinite or finite), and in the level of information provided. Each dataset represents a distinct class of games.

Standard includes data from previous experiments on indefinitely repeated games with perfect monitoring: the stage game is repeated with constant continuation probability δ, and subjects observe each other's action without any noise. The Finite dataset consists of finitely repeated games with perfect monitoring; hence the stage game is repeated with probability one for a finite number of periods, which is common knowledge. The Noise

[5] In line with the standard practice in the literature, we only consider games where the condition $g - l < 1$ holds. This condition implies that alternating between cooperation and defection is not more profitable than mutual cooperation.

Repeated interaction:	**Indefinite**	**Indefinite**	**Finite**
Monitoring:	**Perfect**	**Imperfect**	**Perfect**
Source:	Dal Bo & Frechette (2018)	This paper	Embrey et al.(2018)
Number of sessions:	103	27	27
Number of subjects:	1734	598	552
Number of observations:	116,644	60,334	65,720
Dataset:	STANDARD	NOISE	FINITE

Notes: Since the analyses we conduct are not compatible with supergames lasting less than 3 periods, we drop those observations from all the datasets. The NOISE dataset includes data from Fudenberg et al. (2012), Arechar et al. (2017) (only the treatments without communication), and Aoyagi et al. (2019) (only the treatments strategically equivalent to the setting of Fudenberg et al., 2012 and Arechar et al., 2017).

Table 2: Summary information on the meta-dataset of repeated PD experiments.

dataset encompasses indefinitely repeated games with imperfect monitoring, studied in Fudenberg et al. (2012), Arechar et al. (2017), and Aoyagi et al. (2019). In each period there is a probability E that the realized action or a payoff-relevant signal of the chosen action deviates from the player's intended action. This deviation occurs independently across the two players. When a random deviation occurs, players only observe the profile of the realized public signals, which is payoff-relevant. To take noise into account, the standardization of payoffs reported in Table 1 is done on the expected stage-game payoffs.

2.1 Evidence of restitution strategies

In this section we focus on the frequency of cooperation conditionally on the outcomes from previous periods. To eliminate any effect of reputation and strategic teaching behavior, we limit our analysis to the decisions made in the third period of each supergame. To identify restitution strategies, we examine instances where the opponent defected unilaterally in period 2. We then analyze how the player's action in period 3 changes depending on the outcome of period 1.

To elucidate the implications of restitution strategies, we adopt the concept of 'standing' used in evolutionary game theory (Boerlijst et al., 1997). This term refers to a state variable that tracks the balance between cooperation and defection in a player's history. For instance, a sequence of outcomes DD, CD results in a 'credit' standing for the player. This is because the player cooperated once during the supergame, while the opponent never did. In this case, a restitution strategy would prescribe defection. On the other hand, a sequence of outcomes DC, CD results in an 'even' standing, as both players cooperated an equal number of times. Consequently, a restitution strategy prescribes cooperation. This latter scenario is particularly relevant in our analyses, because it serves as a litmus test: in it, restitution strategies diverge from other common strategies such as Tit-for-Tat, Grim-Trigger, or t-period punishment. While restitution strategies align with these alternatives in most other cases, this particular instance allows us to tell them apart.

Table 3 presents the empirical frequency of cooperation in period 3, following CD outcome in period 2. This is done for all possible outcomes in period 1 (CD, CD and CD, CC). The number of observations for each history is also provided in parentheses.

Outcome in period 1	Standing at start of period 2	STANDARD	NOISE Init.intention C	NOISE Init.intention D	FINITE
DC	debt	0.566	0.859	0.475	0.522
		(746)	(99)	(198)	(458)
DD	even	0.282	0.521	0.186	0.242
		(440)	(73)	(167)	(132)
CD	credit	0.297	0.240	0.179	0.246
		(617)	(425)	(28)	(284)
CC	even	0.239	0.632	0.100	0.217
		(268)	(353)	(10)	(138)
Total		0.383	0.472	0.325	0.367
		(2071)	(950)	(403)	(1012)

Notes: Number of observations in parentheses. In the NOISE dataset, the initial intention is the action chosen by the subject in period 1 before the random realization of the noise.

Table 3: Cooperation frequency in period 3 following a CD outcome in period 2.

Finding A. *Consistent with restitution strategies, in the third period of a supergame subjects are more likely to cooperate following a unilateral defection by the opponent, if they themselves have unilaterally defected in the first period.*

Support for Finding A emerges from Table 3 which shows that cooperation rates in the CD, DC history are approximately 57%, compared to the 24%-30% observed in the other three cases of the STANDARD dataset. A similar pattern is observed in the FINITE dataset, with a cooperation rate of 52% versus 22%-25% in other cases. In the NOISE dataset, we categorize observations based on the intended action in period 1. This provides insight into the subject's strategy, even though the opponent may not have observed it. The pattern in the NOISE dataset mirrors those in the other datasets under both conditions. In conclusion, Finding A is consistent across all three classes of games. It suggests that a sizable proportion of subjects deviate from the behavior prescribed by Grim-trigger or by Tit-for-tat, which would be defection in all four cases of Table 3.

Finding A is substantiated by the results of a series of panel regressions, as shown in Table 4. This table provides estimates derived from a linear probability model that incorporates random effects at the subject level and fixed effects at both the treatment and supergame levels. The treatment-level effects account for variability arising from game-specific parameters, while the supergame-level effects control for the influence of experience. The large and highly significant coefficient for the CD outcome in period 1, observed across all regressions, strongly indicates a widespread adoption of restitution strategies to reestablish cooperation in repeated social dilemmas.

	Standard	Noise		Finite
Outcome in period 1		initial C	initial D	
DC (debt)	0.332***	0.205***	0.369***	0.319***
	(0.036)	(0.045)	(0.117)	(0.054)
DD (even)	0.048	-0.167***	0.141	0.030
	(0.038)	(0.064)	(0.108)	(0.062)
CD (credit)	0.038	-0.421***	0.064	-0.022
	(0.037)	(0.035)	(0.124)	(0.053)
Intercept	0.339***	0.508***	-0.002	0.131
	(0.088)	(0.080)	(0.148)	(0.128)
N. of observations	2071	950	403	1012
N. of subjects	949	394	244	348
R^2-within	0.150	0.323	0.110	0.188
R^2-between	0.140	0.167	0.180	0.115
R^2-overall	0.143	0.229	0.151	0.139

Notes: Linear probability model, with one observation per subject, per period. The model incorporates random effects at the subject level and fixed effects at both the treatment and supergame levels. The default case is outcome CC in $t-1$. Standard errors robust for heteroskedasticity. $^{***}p < .01$, $^{**}p < .05$, $^{*}p < .1$.

Table 4 Cooperation in period 3 when the outcome was CD in period 2.

Anecdotal evidence for the use of restitution strategies also emerges from the chats among the participants of Fudenberg et al. (2012) (included in our Noise dataset). The following descriptions, for instance, align with restitution strategies and contrasts with other strategies such as Grim or Tit-for-Tat:[6]

> "Keep the other person honest. Compensate for errors"

> "[...] If I had accidentally picked B, I would pick A and make it up to the other person. then if they had chosen B I would assume they were checking me and continue to choose A"

[6] Many other chats reported in the Online Appendix of Fudenberg et al., 2012 are consistent with restitution strategies but also with other strategies.

> "If the other person gave me 4 points and lost 2 points in the previous round I would try to do the same [...] I would choose B after a round where I played A and the other played B, because in the previous round I gave up points to the other person so I expect them to do the same for me in this round"

> "If it became clear we were going tit for tat with each other I'd try and break out of the cycle by playing A even if his or her last move was B"

2.2 Beyond the first three periods

Our analyses so far have relied on the first three periods of each supergame to avoid reputational or strategic teaching effects. Here, we extend our analyses to the following periods to gain further insights.

Table 5 reports the results of regressions in which the dependent variable is the subject's action in period t, conditional on the outcome in period $t-1$ being CD. The independent variable is whether the subject was "in debt" at the beginning of period $t-1$, when the opponent's unilateral defection took place. The results indicate a positive correlation between the state of being "in debt" and cooperation in periods following a unilateral deviation by the opponent, in all the three classes of experiments we consider. This result indicates that subjects employ strategies with memory longer than one. Thus, the analysis of data from periods beyond the third one confirms Finding A, showing that subjects react to a unilateral defection by the opponent in a different way, depending on their standing.

One might wonder whether the standing of being "in debt" simply captures the outcome in $t-2$, or instead subjects also react to "debts" that were generated earlier in time. In fact, restitution strategies predict higher cooperation with a "debt" standing, even considering histories before $t-2$. The question is of particular importance in light of the recent literature on the repeated Prisoner's Dilemma, which predominantly emphasizes the use of memory-1 strategies to explain subjects' behavior. While we have previously demonstrated the existence of memory-2 strategies in our Finding A, here we propose a method to ascertain the empirical relevance of restitution strategies that employ even longer memory spans.

To this aim, we replicate the regression presented in Table 5, restricting the sample to interactions where the outcome in $t-2$ was DC. If subjects only reacted to debts that emerged in $t-2$, then the coefficient of the variable Debt_{t-1} should no longer be significant. The results of this analysis are presented in Table 6. They reveal that even if the imbalance in cooperation emerged more than two periods before, the probability of cooperation is still higher when subjects are 'in debt', compared to when they are not. The effect, however, is only significant in Standard, and in Noise when the initial intention was cooperative. This

	Standard	Noise		Finite
		Initial C	Initial D	
Debt $_{t-1}$	0.247***	0.390***	0.123***	0.374***
	(0.023)	(0.031)	(0.041)	(0.055)
Period	0.004**	−0.000	0.003	−0.027***
	(0.001)	(0.002)	(0.002)	(0.004)
Debt $t_{-1}\times$ Period	−0.009***	−0.020***	−0.012**	−0.034***
	(0.003)	(0.003)	(0.006)	(0.009)
Intercept	−0.038	0.251***	0.112***	0.354***
	(0.025)	(0.039)	(0.039)	(0.075)
N. of observations	6161	5683	2851	3827
N. of subjects	1283	499	378	438
R^2-within	0.075	0.062	0.020	0.132
R^2-between	0.052	0.097	0.141	0.006
R^2-overall	0.061	0.064	0.052	0.089

Notes: Linear probability model, with one observation per subject, per period. The model incorporates random effects at the subject level and fixed effects at both the treatment and supergame levels. Standard errors robust for heteroskedasticity. $^{***}p < .01$, $^{**}p < .05$, $^{*}p < .1$.

Table 5: Cooperation in period t when the outcome was CD in $t-1$ (all periods after 2).

	Standard	Noise		Finite
		Initial C	Initial D	
Debt $_{t-1}$	0.107**	0.340***	0.060	0.196
	(0.049)	(0.065)	(0.092)	(0.137)
Period	−0.000	−0.003	0.006	−0.024*
	(0.004)	(0.006)	(0.007)	(0.014)
Debt $t_{-1}\times$ Period	0.000	−0.007	0.000	−0.006
	(0.005)	(0.007)	(0.010)	(0.023)
Intercept	0.000	0.251***	0.071	0.456**
	(0.085)	(0.078)	(0.098)	(0.179)
N. of observations	1103	1078	540	499
N. of subjects	565	388	245	195
R^2-within	0.064	0.131	0.083	0.172
R^2-between	0.077	0.137	0.093	0.070
R^2-overall	0.081	0.126	0.082	0.110

Notes: Linear probability model, with one observation per subject, per period. The model incorporates random effects at the subject level and fixed effects at both the treatment and supergame levels. Standard errors robust for heteroskedasticity. $^{***}p < .01$, $^{**}p < .05$, $^{*}p < .1$.

Table 6: Cooperation in period t when the outcome was CD in $t-1$ and DC in $t-2$ (all periods after 3).

is broadly in line with the use of restitution strategies with memory longer than two periods as estimated in the previous subsection. These results are consistent across all classes of games, and starkly contrast with the predictions of both Grim trigger and Tit-for-tat strategies, which prescribe defection after a CD outcome, regardless of the history of play before period $t-1$.

Finding B. *When considering all periods in a supergame, empirical evidence consistently aligns with restitution strategies. Choices are strongly affected by the outcomes of previous periods, documenting the relevance of strategies that employ a memory span of two or more periods.*

The evidence presented so far clearly points to a widespread use of restitution strategies to rebuild cooperation in repeated social dilemmas. In particular, it suggests that a deviation by the opponent is deemed more acceptable and often leads to increased cooperation, if it makes players more "even". Notably, extending our analysis beyond the third period has enabled us to identify the use of restitution strategies with a memory span longer than two periods, which has further implications. Finding B diverges from the prevalent approach adopted in the recent literature on repeated prisoner's dilemma experiments, as it implies that focusing solely on memory-one strategies when analyzing experimental data on repeated games could overlook significant aspects of subjects' behavior, in this case how asymmetries in past play affect restitution and cooperation.

The reason why Findings A and B may have previously gone unnoticed is that their discovery necessitates an examination of specific histories of miscoordination where restitution and other strategies have conflicting predictions.[7] These specific histories could easily be obscured in an aggregate analysis of all data. Yet, focusing on them is essential because for most other recorded histories, the predictions of restitution strategies align with those of other commonly considered strategies, such as Tit-for-Tat. Based on these other scenarios, even if all subjects were playing restitution strategies, an observer could erroneously conclude that they are all playing Tit-for-Tat (or vice versa). Therefore, focusing on the aforementioned specific subset serves as a crucial litmus test for determining whether subjects employ restitution strategies. Notice, however, that even though we only focus on a relatively small subsample of the datasets, this restriction does not imply that our conclusions are based on the behavior of a minority of subjects. The sub-sample of observations we focus on in this section covers 77% of the total number of subjects included in the three meta-datasets. Section 4 will present additional empirical evidence, this time based on unrestricted meta-datasets.

Two important remarks are in order with respect to the Noise and Finite class of games. First, Findings A and B are particularly noteworthy, as they emerge also in the Noise

[7] These instances only constitute about 8% of the combined observations from the three datasets.

dataset, which is characterized by imperfect monitoring. In this context, subjects generally exhibit more leniency than under perfect monitoring (Fudenberg et al., 2012), and this might imply a looser link between a subjects' standing and their cooperative choices. Contrarily, we present systematic evidence consistent with restitution strategies, particularly when the intended action in the first period was C: a scenario which is the most compatible with the implementation of a restitution strategy. Second, while one might anticipate empirical evidence of restitution strategies in the Standard dataset, their presence in the Finite dataset is more difficult to interpret, as canonical theory would not predict their occurrence in finitely repeated games. Van Damme (1989) proved that restitution strategies can be equilibrium strategies in infinitely repeated games with perfect monitoring. We corroborate this in Section 3, extending the prediction to games with imperfect monitoring. However, these strategies cannot be equilibrium strategies in finitely repeated games (e.g. Embrey et al., 2017). In the next subsection we delve deeper into this unexpected finding by examining the impact of experience on the adoption of restitution strategies.

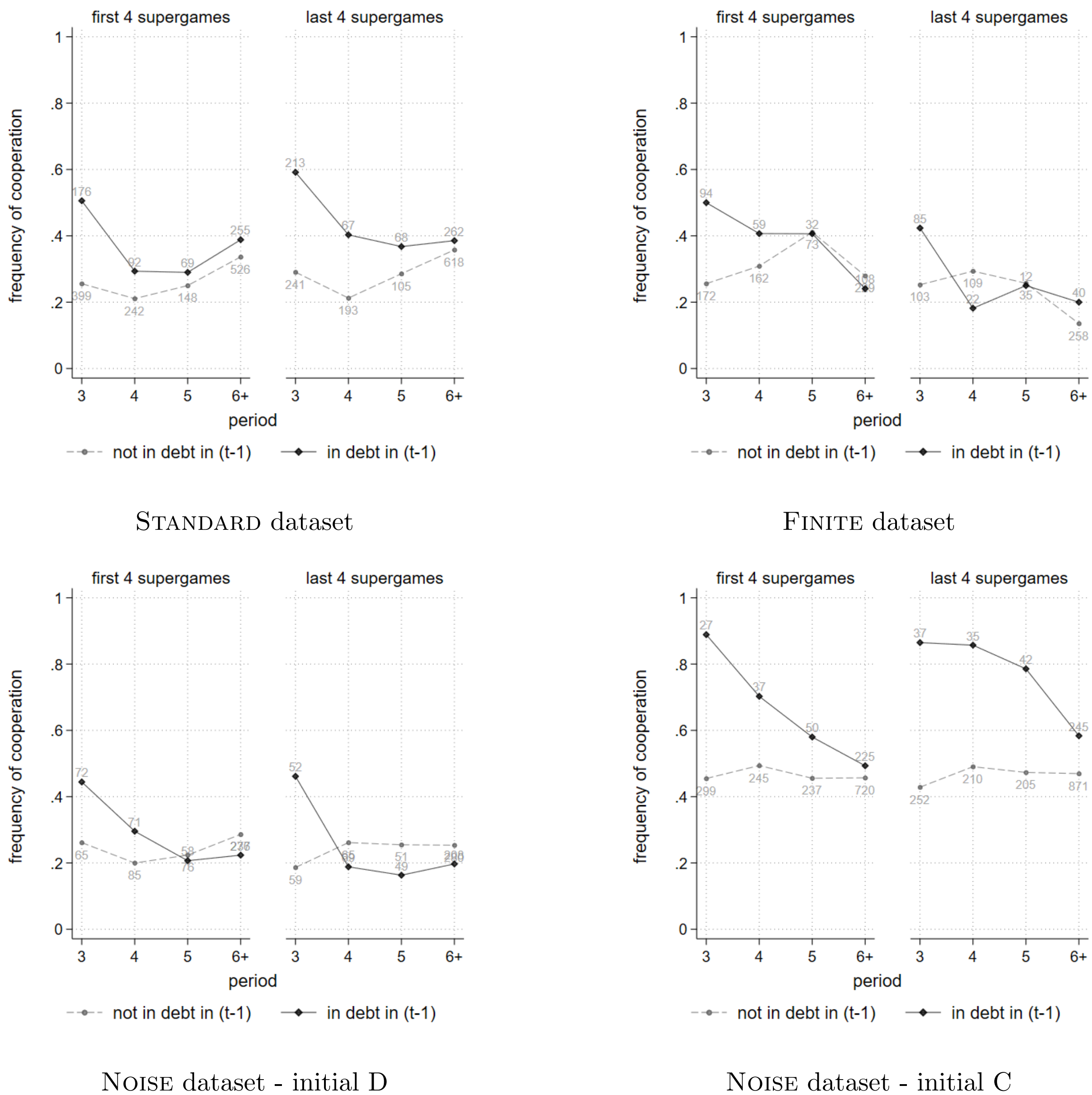
2.3 The Role of Experience

In our three datasets, subjects participated in multiple supergames during each experimental session. To examine the strategy adoption while subjects gain experience, we contrast the empirical patterns that emerged in the earliest supergames with those from the latest supergames within each session.

We adopt the same approach as in Subsection 2.2, and contrast the first and last four supergames of each session. Figure 1 illustrates the frequency of cooperation following a unilateral defection by the opponent (CD), with Tables 7 and 8 providing additional analyses. As in previous subsections, we partition observations based on the subject's standing – whether 'in debt' or not – in the period when the opponent unilaterally defected.

Let us point our attention to the first four supergames, represented in Figure 1. The solid line typically lies above the dashed line, suggesting a higher frequency of cooperation among subjects who were 'in debt' when their opponent unilaterally defected, compared to those who were 'even' or 'in credit'. This observation is corroborated by the results in Table 7 and aligns broadly with Findings A and B, which indicated the adoption of a restitution strategy by a sizable fraction of subjects.

Let us now turn to the last four supergames. The results differ depending on the class of games. In finitely repeated games, with experience the effect diminishes in size and becomes less significant with experience (Finite dataset). Conversely, in indefinitely repeated games, the effect not only endures with experience, but it also grows in magnitude (Standard dataset). When evaluating the evidence under imperfect monitoring (Noise dataset), it



Standard dataset

Finite dataset

Noise dataset - initial D

Noise dataset - initial C

Figure 1: Inexperienced vs. experienced behavior:
cooperation rates conditional on standing (all periods after 2).

	STANDARD	NOISE		FINITE
		Initial C	Initial D	
Debt $_{t-1}$	0.170***	0.321***	0.125*	0.358***
	(0.035)	(0.049)	(0.069)	(0.089)
Period	0.005**	0.002	0.006	−0.023***
	(0.002)	(0.003)	(0.005)	(0.008)
Debt $t_{-1}\times$ Period	−0.006*	−0.023***	−0.015	−0.052***
	(0.004)	(0.005)	(0.009)	(0.015)
Intercept	0.346	0.260***	0.144**	0.321***
	(0.360)	(0.052)	(0.057)	(0.099)
N. of observations	1907	1840	940	929
N. of subjects	804	430	302	320
R^2-within	0.048	0.039	0.013	0.101
R^2-between	0.063	0.088	0.076	0.032
R^2-overall	0.063	0.051	0.045	0.066

Notes: Linear probability model, with one observation per subject, per period. The model incorporates random effects at the subject level and fixed effects at both the treatment and supergame levels. Standard errors robust for heteroskedasticity. $^{***}p < .01$, $^{**}p < .05$, $^{*}p < .1$.

Table 7: Cooperation in period t when the outcome was CD in $t-1$ (all periods after 2, first 4 supergames).

is important to remember that restitution strategies prescribe cooperation as the action in period 1. With an initial intention to cooperate, the effect persists and increases in size. Instead, if the initial intention is to defect, the effect – already smaller for inexperienced subjects – disappears with experience.[8]

Finding C. *With experience, the use of restitution strategies declines in finitely repeated experiments, while it becomes more prominent in indefinitely repeated games.*

A possible interpretation of Finding C relies on both behavioral and theoretical arguments. While both aspects may be pertinent, our results can shed some light on their respective roles and their relative importance in our settings. Restitution strategies may have an intuitive appeal due to fairness considerations, which could account for their systematic adoption by inexperienced subjects. These strategies align well with insights from the literature on inequity aversion (Fehr and Schmidt, 1999; Charness and Rabin, 2002), according to which fairness-related heuristics may prevail over strategic considerations of income maximization. For experienced subjects, instead, the primary determinant of strategy

[8] Figure 1 and Tables 7 and 8 also provide a more detailed view of the effect's magnitude within a supergame. Across all game classes, the gap between the solid and the dashed lines reaches its peak in period 3, then diminishes. This pattern further refines Findings A and B.

	Standard	Noise		Finite
		Initial C	Initial D	
Debt $_{t-1}$	0.250***	0.416***	0.038	0.209
	(0.036)	(0.053)	(0.086)	(0.128)
Period	0.004**	0.001	−0.002	−0.030***
	(0.002)	(0.002)	(0.005)	(0.010)
Debt $t_{-1}\times$ Period	−0.006**	−0.021***	−0.003	−0.017
	(0.003)	(0.006)	(0.012)	(0.021)
Intercept	0.238*	0.250	0.051	0.319***
	(0.126)	(0.204)	(0.433)	(0.111)
N. of observations	1767	1897	833	664
N. of subjects	661	406	252	288
R^2-within	0.065	0.068	0.027	0.142
R^2-between	0.141	0.147	0.155	0.084
R^2-overall	0.099	0.091	0.121	0.094

Notes: Linear probability model, with one observation per subject, per period. The model incorporates random effects at the subject level and fixed effects at both the treatment and supergame levels. Standard errors robust for heteroskedasticity. $^{***}p < .01$, $^{**}p < .05$, $^{*}p < .1$.

Table 8: Cooperation in period t when the outcome was CD in $t-1$ (all periods after 2, last four supergames).

adoption may be whether restitution strategies constitute equilibrium strategies for income-maximizing subjects. Restitution strategies are not equilibrium behavior in finitely repeated games; by contrast, in Section 3 we show that restitution strategies can be consistent with equilibrium behavior in indefinitely repeated games (both with and without noise).[9]

Given the empirical insights from Findings A, B and C, it may seem surprising that restitution strategies have not been considered so far in the experimental literature on repeated social dilemmas. In the Discussion, we elucidate why the toolbox employed for strategy estimation in this branch of literature made the identification of such strategies rather challenging, if not impossible.

3 Payback and the theory of infinitely repeated games

To deepen our understanding of the general findings reported earlier, in this section, we consider a specific restitution strategy - Payback - that has been studied in the evolutionary

[9] The persistence of restitution strategies only in games where they can be part of an equilibrium also links our results to the literature on cooperation and personality traits, which suggests that social preferences and individual inclinations tend to have a significant, but typically transitory effect on cooperation rates (Proto et al., 2021), which are outweighed by strategic concerns (Dreber et al., 2014).

literature under the names of Sophisticated-TFT and Contrite-TFT (Sugden, 1986; Boyd, 1989; Wu and Axelrod, 1995; Boerlijst et al., 1997; Graser and van Veelen, 2024). We have selected Payback because it belongs to the class of the simplest restitution strategies when measured in terms of complexity by the number of states and transitions.[10] After describing Payback, we characterize its theoretical properties in infinitely repeated games, both under perfect and imperfect monitoring, comparing them with those of Grim and TFT.[11] The imperfect monitoring case is particularly important because, there, breaches of cooperation are unavoidable, which better highlights the differences between Payback and strategies not based on restitution.

3.1 The Payback strategy

Figure 2 describes the Payback automaton, which consists of three states represented by circles and a set of transition rules represented by arrows. The upper-case letter within each circle indicates the action to be taken by the agent in that state: we denote cooperation by "C," and defection by "D." The pairs of lower-case letters next to the arrows represent the action pairs that trigger the corresponding transition: "cc" denotes mutual cooperation, "dd" mutual defection, "cd" a unilateral defection by the opponent, and "dc" a unilateral defection by the player. As we discuss later, in the case of a game with perfect monitoring, the actions determining transitions are the actions chosen by the players. In the case of imperfect monitoring, these are actions publicly observed by the agents, which may differ from the actions intended by the players.

Payback prescribes to start at the top-left circle, which we term the "Cooperation state": hence, the initial action is cooperation (C). The automaton remains in the "Cooperation state" if the observed outcome is "cc" or "dd", while if the outcome is "cd" it triggers a transition to the "Punishment state," in the bottom-left circle: thus, it prescribes to defect (D) until observing cooperation (c) from the opponent. Finally, if the outcome in the cooperation state is "dc", Payback triggers a transition to the "Restitution state" (in the upper-right circle), which prescribes cooperation (C).

The Payback strategy differs notably from the Grim strategy because it punishes deviations asymmetrically: when player 1 defects (intentionally or not), his continuation payoff goes down. At the same time, the continuation payoff of player 2 goes up. In the continuation game, player 2 is paid back for the bad outcome of this period. In contrast, Grim (and any strongly symmetric strategy, like the ones described in Green and Porter, 1984) penalizes both players when one of them defects.

[10] See Oprea (2020) and Salant (2011).

[11] We do not discuss finitely repeated games because Payback can never be an equilibrium there.

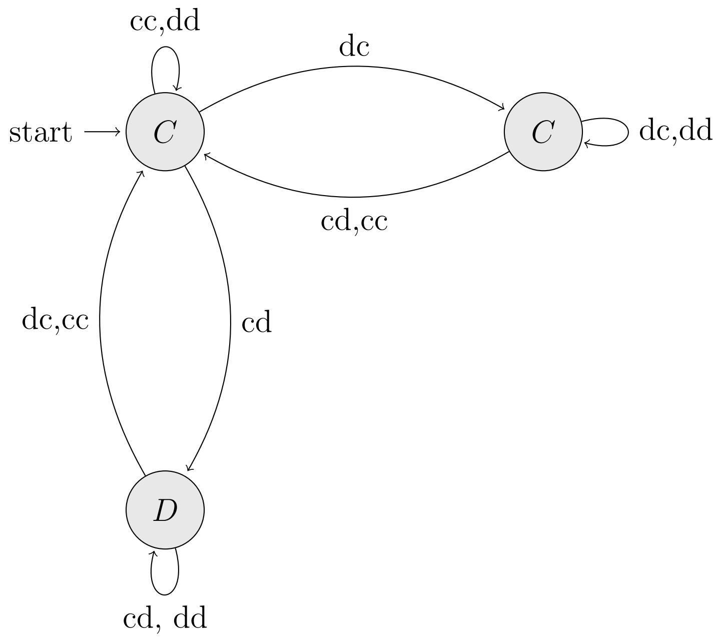


Figure 2. Automaton of "Payback strategy"

At first sight, Payback may look very similar to TFT, but the difference comes from Payback depending on more than what happened in the last period. In particular, there are two instances that set Payback apart from TFT. First, if the observed outcome is cd, the prescription of Payback depends on the previous state: cooperation in the Restitution state and defection in the other states. Second, if the agent observes dd, Payback prescribes defection only if the agent is in the Punishment state. Otherwise, it prescribes cooperation. Instead, in both instances, TFT (and Grim) prescribe defection independent of the previous history.

3.2 Perfect monitoring

Under perfect monitoring, Grim has the lowest possible critical discount factor (δ) among all strategies that sustain cooperation in every period as a subgame perfect equilibrium (SPE) outcome. The reason is that it is an optimal penal code in the sense of Abreu (1988), as it uses the harshest possible punishment for deviations (in PD, the static Nash payoffs are the min-max payoffs). Hence, while on the cooperative path, the value of cooperation is the same, the critical discount factor for Payback has to be weakly higher than for Grim.

Payback prescribes that a deviation on the equilibrium path, with a gain g on top of the equilibrium payoff, is compensated by repayment in the subsequent period, with an implied loss of l.

The larger is g (in comparison to l), the more tempting it is to deviate unilaterally in the cooperative state. Also, if l is too large, there would be an incentive to deviate during the repayment phase. Depending on the parameters, either of these constraints may bind. We show in Appendix A that the relevant constraint for Payback to be a SPE is:

$$\delta \geq \frac{\max\{l, g\}}{1+l}.$$

Specifically, if $g > l$, the binding incentive constraint is in the Cooperation state, while if $l > g$, the binding constraint is in the Restitution state. Following this reasoning, we prove in Appendix A the following proposition:[12]

Proposition 1. *Consider an infinitely repeated PD with perfect monitoring and normalized payoffs as in Table 1. (i) There exists a threshold $\delta^* < 1$ such that for all discount factors $\delta \geq \delta^*$ Payback is a subgame perfect equilibrium strategy. (ii) If $g = l$, this threshold is the same for Grim and Payback; if $g \neq l$, the threshold is strictly lower for Grim than for Payback.*

Note that if $g = l$, then Grim and Payback are SPE for the same set of parameters: the anticipation of the repayment phase offsets gains from a one-sided deviation to the same extent as the commitment to an indefinite punishment. Payback's punishment phase is much more efficient, however, as players can keep cooperating after just one round of punishment (Restitution) in which the defector compensates the opponent for the deviation.[13]

Finally, a standard result in the literature is that if $g - l < 1$, as imposed in our game, TFT is not sustainable as an SPE, for any discount factor. The reason is that after a unilateral deviation, TFT calls for an infinite sequence of mis-coordinated actions (cd-dc) that the players would have incentives to deviate from.

3.3 Imperfect monitoring

We now turn to analyzing Perfect Public Equilibria (PPE) of the repeated game with imperfect monitoring. In this subsection, we assume $g = l$ for two reasons. First, in order to theoretically characterize the implications of imperfect monitoring, we want to compare Grim and Payback under a common perfect monitoring benchmark, which by Proposition 1 is

12 Recall that we have assumed above that $g - l < 1$.

13 While this does not matter for equilibrium values under perfect monitoring because deviations do not happen on the equilibrium path, it matters in terms of robustness of the strategy to renegotiation. The threat of never cooperating again after a player unilaterally deviates is so inefficient that may be implausible if players can renegotiate the continuation equilibrium. The punishment phase in Payback is robust to such concerns because it is much more efficient and asymmetric, so carrying out the repayment phase is more profitable for the party that suffered the defection than renegotiating away the punishment phase and immediately returning to mutual cooperation (Van Damme, 1989).

implied by $g = l$. Second, the main related experimental results under imperfect monitoring (Fudenberg et al., 2012; Arechar et al., 2017; Aoyagi et al., 2019) assume $g = l$.[14]

The specific game with imperfect monitoring we study has been studied in Fudenberg et al. (2012) (FRD, henceforth) and Arechar et al. (2017). We assume that in each period, players privately choose intended actions. Then, each intended action is changed with probability $E < \frac{1}{2}$, independently across the two players. Players publicly observe the implemented action profile, which is payoff relevant.[15]

In Appendix A, we show that in this class of games, Payback is sustainable in a PPE for a larger range of discount factors than Grim:

Proposition 2. *In the infinitely repeated PD game with imperfect monitoring described above, (i) if* $\frac{1+g}{g} > \frac{E}{(1-2E)(1-E)}$*, then there exist threshold levels for the discount factor,* $\delta_{PPE,Payback} < 1$*,* $\delta_{PPE,Grim} < 1$*, above which Payback and Grim are perfect public equilibrium strategies; and (ii) for all* $g = l > 0$*, and* $E \in (0, \frac{1}{2})$*, the threshold for Payback is strictly lower than for Grim:* $\delta_{PPE,Payback} < \delta_{PPE,Grim}$.

The intuition for the ranking of the threshold discount factors for Payback and Grim is as follows. Under Grim, cooperation can be sustained in equilibrium for sufficiently high discount factors, but the punishment sustaining cooperation is much more inefficient than for Payback. When players move from mutual cooperation to mutual defection during the Grim punishment phase, the sum of payoffs drops approximately from 2 to 0 (ignoring noise). In contrast, when a player unilaterally defects and then moves to the Restitution state, accepting the opponent's punishment, the sum of payoffs drops approximately only to 1, while providing even stronger punishment per period (the deviating player loses $1 + l$ instead of 1). So the false-positive punishments caused by noise are just much more efficient (in terms of keeping a high sum of continuation payoffs while satisfying the incentive compatibility constraints) in Payback than in Grim. This improves the expected payoffs from cooperation on the equilibrium path and relaxes the Payback's IC constraints for a PPE.

While Proposition 2 reveals that Payback is a PPE for a wider range of parameters than Grim, the next proposition shows that, when both are PPEs, Payback also offers a higher expected payoff than Grim:

Proposition 3. *In the infinitely repeated PD game with imperfect monitoring described above, for all* $g = l > 0$*,* $\delta \in (0, 1)$ *and* $E \in (0, \frac{1}{2})$ *such that both Grim and Payback are*

[14] The reason is that whenever $g = l$ an equivalent representation of the Prisoner's Dilemma exists, in which cooperation and defection take the "benefit/cost" form, b/c in Fudenberg et al. (2012) notation. Specifically, to cooperate, players have to pay a cost c to give a benefit b to the other player, while defection gives 0 to each party.

[15] As mentioned earlier, the environment is also strategically equivalent to that in the imperfect public monitoring treatment by Aoyagi et al. (2019), which we also include in our analysis of the metadata.

a PPE, the expected payoff in the Payback PPE ($V_{Payback}$) is strictly higher than the expected payoff in the Grim PPE (V_{Grim}).

Remark 1. *It can be shown that in our imperfect monitoring game, TFT cannot be sustained in a PPE.*[16] *Moreover, in the proof of Proposition 3 we also show that the expected payoff in the Payback PPE ($V_{Payback}$) is higher than the expected payoff from both players following TFT (V_{TFT}) for any $\delta \in (0, 1)$. The main reason for this ranking is Payback's ability to keep the value of cooperation high even after accidental deviations. Payback can induce a return to cooperation just one period after the accidental deviation (if no other errors occur), while with TFT, a sequence of asymmetric $cd-dc$ outcomes would follow, reducing the continuation value. With TFT, Players would be able to revert to mutual cooperation only after a second random error would change the intended action from d to c.*

Strategic Risk. Several recent experiments on the infinitely repeated Prisoner's Dilemma with perfect monitoring have shown that the "strategic risk" of cooperating, i.e. the expected cost of cooperating when the opponent defects, is a strong predictor of cooperation[17]. Blonski and Spagnolo (2015) and Blonski et al. (2011) showed (theoretically) that a simple measure of such risk can be obtained by restricting attention to the cooperative strategy and the uniquely safe strategy AllD, and then comparing the two by applying Harsanyi and Selten (1988)'s concept of risk dominance to the resulting 2x2 automaton game.[18] Dal Bó and Fréchette (2011, 2018) noted that the above comparison can alternatively be performed using the basin of attraction, an easy-to-compute indicator from the theory of dynamic systems. The riskiness of a cooperative strategy is then measured by the size of the basin of attraction of AllD (SizeBAD).[19] In the Appendix A, we present the calculations for SizeBAD when AllD plays against Payback, Grim, and TFT, and prove the following proposition:

Proposition 4. *In the infinitely repeated PD game with imperfect monitoring described above, for all $g = l > 0$ and $E \in (0, \frac{1}{2})$ strategic risk (SizeBAD) is lower for Payback than for Grim and TFT.*

[16] This result also follows from Graser and van Veelen (2024), who show that, in the presence of errors, Nash equilibrium strategies that can be represented via finite automata must exhibit a structure that they label "self-mirroring", which immediately rules out Tit-for-Tat.

[17] See, among others, Blonski et al. (2011), Dal Bó and Fréchette (2011), Calford and Oprea (2017), Dal Bó and Fréchette (2018), Ghidoni and Suetens (2022), Boczo et al. (2023), and Martinez-Martinez and Normann (2024).

[18] See Blonski and Spagnolo (2015) for details. The authors also derive a cutoff level for the discount factor, δ_{RD}, below which all cooperation equilibria are risk-dominated by AllD, and cooperation should not be expected. Blonski et al. (2011) derived the same cutoff axiomatically.

[19] SizeBAD is then the probability that a player must assign to the opponent playing the cooperative strategy in order to be indifferent between playing that strategy rather than AllD. Dal Bó and Fréchette (2018) find that the two measures of strategic risk are empirically indistinguishable in their metadata. Note also that SizeBAD=1/2 when δ=δ_{RD}.

Under perfect monitoring, without the possibility of a tremble, the value of cooperation of TFT, Grim and Payback is the same, while they all punish indefinitely after the first deviation if matched with AllD. Therefore, the strategic risk is the same for all the cooperative strategies considered. Under imperfect monitoring this is not true anymore. On the one hand, Payback loses more against AllD compared to Grim, which in turn condemns AllD to lower profits in a pairwise comparison. On the other hand, due to its forgiving nature, the value of cooperation sustained by Payback in equilibrium is much higher than the value sustained by Grim (Proposition 3). Ultimately, the asymmetric punishment phase of Payback solves the trade-off more efficiently, containing the losses against AllD while keeping the value of cooperation high. Conversely, Grim's lower exposure to losses against AllD does not outweigh the low efficiency implied by its stricter punishment phase. A similar reasoning, mostly driven by efficiency considerations, holds when we compare TFT with Payback.

4 Empirical support for Payback

In this section we perform a maximum likelihood estimation of strategies taking Payback into account. We then analyze the ability of strategic risk to predict cooperation with imperfect monitoring.

4.1 Retrieving strategies from observed actions

Differently from Section 2, the empirical evidence reported here is based on all the observations in the three datasets of Table 2.

Table 9 reports Maximum Likelihood Estimates of strategies for the three classes of repeated games we are studying, respectively with and without Payback among candidates. Estimates are obtained using the methodology originally developed by Dal Bó and Fréchette (2011), and – in line with the typical procedure adopted in this literature – they are based on the last four interactions in each session to allow subjects to gain experience. Besides Payback, the set of candidate strategies includes all those considered in Fudenberg et al. (2012). For finitely repeated games, we follow Aoyagi et al. (2024) and consider a different set of strategies, including also threshold strategies, which conditionally cooperate until a threshold round before switching to AllD.

Column 1 presents the results of our strategy estimation for infinitely repeated games with perfect monitoring, without considering Payback. Column 2 presents the same estimates including Payback in the set of candidate strategies. Results in Column 2 indicate that, when Payback is included in the set of strategies, its estimated prevalence (10%) is

Strategies	Standard (1)	Standard (2)	Noise (3)	Noise (4)	Finite (5)	Finite (6)
Payback	/	0.10***	/	0.09***	/	0.02*
TFT	0.24***	0.16***	0.07***	0.02***	0.08***	0.06***
ALLC	0.02***	0.02***	0.05***	0.05***	0.02***	0.02***
TF2T	0.03***	0.02*	0.11***	0.08***	0	0
TF3T	0.01**	0.01***	0.05***	0.05***	/	/
2TFT	0.03***	0.02**	0.07***	0.07***	0	0
2TF2T	0.03***	0.02***	0.10***	0.10***	/	/
Grim	0.14***	0.14***	0.06***	0.06***	0.05***	0.05***
Grim-last2	0.01**	0.02***	0.06***	0.06***	0.03***	0.03***
Grim-last3	0.02***	0.02***	0.10***	0.09***	/	/
ALLD	0.30***	0.30***	0.28***	0.28***	0.35***	0.35***
D-TFT	0.17***	0.17***	0.05***	0.05***	0.10***	0.10***
Threshold $t-3$	/	/	/	/	0.12***	0.12***
Threshold $t-2$	/	/	/	/	0.13***	0.13***
Threshold $t-1$	/	/	/	/	0.13***	0.13***
Gamma	0.37	0.37	0.48	0.48	0.38	0.38
Number of observations	50218		11696		13144	
Number of subjects	1734		598		552	

Notes: Estimates based only on the last four interactions in each session. p-values are based on bootstrapped standard errors.

Table 9: Maximum Likelihood Estimates of strategies.

comparable in magnitude to that of Grim (14%) and TFT (16%). The comparison between Columns 1 and 2 further highlights that the addition of Payback predominantly affects the estimated prevalence of TFT, which decreases by 8 percentage points. This suggests that the SFEM approach struggles to differentiate between these two strategies, which predict identical behavior in most empirical occurrences.

Columns 3 and 4 make the same comparison for infinitely repeated games with imperfect monitoring. Column 4 shows that, also in this class of supergames, Payback captures a substantial and significant fraction of the observed behavior (9%). Comparing Column 3 with Column 4 highlights again how the introduction of Payback reduces the estimated prevalence of TFT, which drops from 7% to 2%.[20] The higher prevalence of Payback as compared to TFT in this setting is not surprising, since with imperfect monitoring TFT and Payback are easier to distinguish from one another; in addition TFT is not an equilibrium strategy, while as shown in the previous section Payback is, and is also much more resilient

[20] By contrast, the introduction of Payback only marginally reduces the estimated prevalence of delayed-trigger strategies, hence it does not question the importance of being "slow to anger", pointed out by Fudenberg et al. (2012).

to strategic risk than TFT.

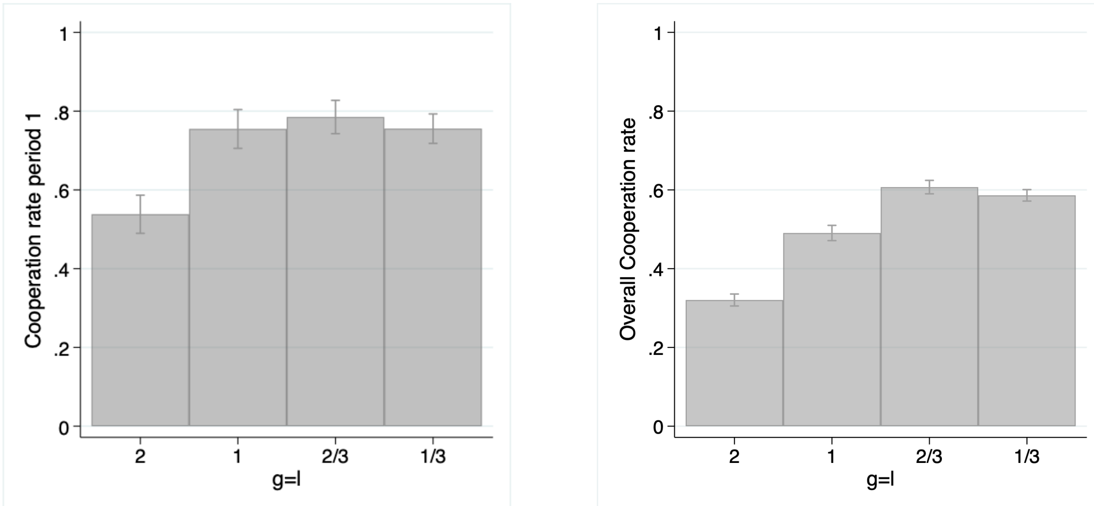
The estimation of strategies through maximum likelihood should be viewed in conjunction with the results discussed in Section 2, rather than in opposition to them. Several reasons underpin this approach. First, here we explicitly focus on a single restitution strategy, namely Payback. However, if this specific strategy fails to emerge, it remains plausible that our Findings A-C could be driven by alternative restitution strategies. Second, unlike in the analyses in Section 2, our strategy estimation framework here assumes that individual subjects consistently employ the same strategy across multiple supergames. Third, we leverage data from all subjects and their complete history of play, rather than restricting our analysis to specific segments that allow for the differentiation between restitution and other strategies.

These methodological considerations are pivotal when interpreting the relatively modest estimated prevalence of the Payback strategy, as observed in Table 9. Notably, the action predicted by Payback diverges from that predicted by other conditionally cooperative strategies only in highly specific instances. Consequently, the methodology employed in this subsection may not be optimally suited to disentangle these subtle differences.

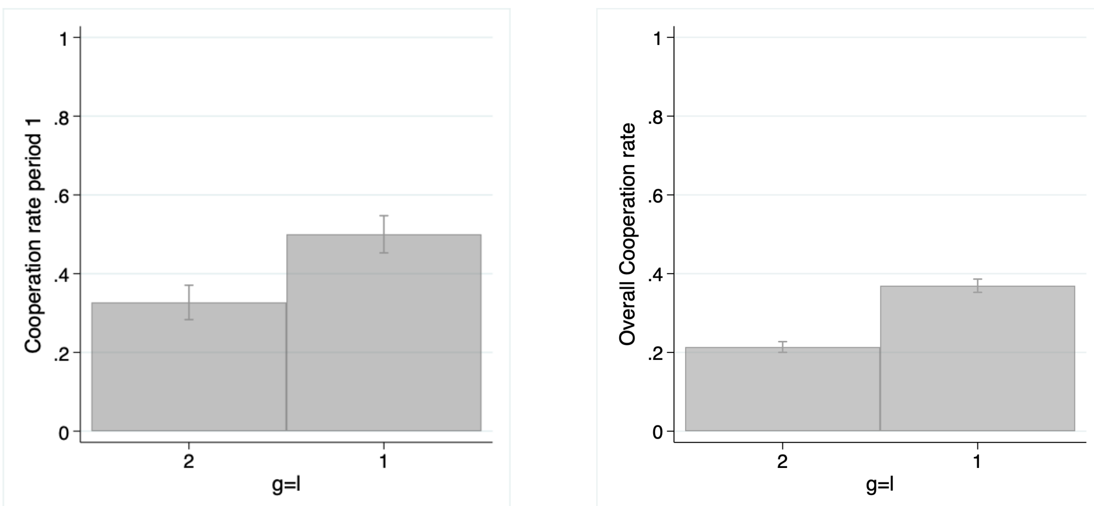
4.2 Strategic risk and imperfect monitoring

Besides their main results on the importance of lenient and forgiving strategies in repeated games with noise, Fudenberg et al. (2012) highlight a "puzzling" finding. They investigate whether risk dominance (in the sense of Blonski and Spagnolo, 2015) is a good predictor of cooperation also with imperfect monitoring, and conclude that this is not the case. However, they measure risk dominance of cooperation taking TFT as the reference strategy.[21] They observe that TFT becomes risk dominant only in treatments with $g = l = 2/3$, but the largest increase in cooperation occurs between treatments with $g = l = 2$ and $g = l = 1$. Therefore, they argue that *"the risk-dominance criterion has at best limited predictive power regarding cooperation in games with noise"* (Fudenberg et al., 2012, p. 742).

[21] Similarly Dvorak and Fehrler (2023) use Grim to find the threshold that makes a cooperative equilibrium risk-dominant.



FRD dataset



Arechar et al. (2017) dataset

Notes: In all treatments, $E = \frac{1}{8}$. Results based on the last 4 supergames of each session.

Figure 3: Cooperation rates by treatment under imperfect monitoring.

We show here that this conclusion is reversed if we consider risk dominance of Payback, rather than TFT. In Figure 3 we present the empirical cooperation rates observed by Fudenberg et al. (2012) and by Arechar et al. (2017), across treatments that only differ for the payoff parameters. Their results reveal that the initial and aggregate cooperation rates increase significantly from the treatments where $g = l = 2$ – in which Payback is not risk-dominant – to the treatments in which $g = l = 1$, where it is. Panel 3 of Table 10 below reports the minimum discount factors for the strategies of interest to be risk dominant in FRD's and Arechar et al. (2017)'s treatments. As illustrated by FRD, for their continuation probability (87.5%) TFT becomes risk dominant only in the treatments with $g = l = 2/3$, which is misaligned with the increase in cooperation between treatments with $g = l = 2$ and $g = l = 1$. Panel 4 shows that the same conclusion can be drawn using SizeBAD to measure strategic risk. The same two panels show that, instead, Payback is not risk dominant for $g = l = 2$ ($\delta_{RD} = 0.95 \geq 0.875$) and becomes risk dominant for $g = l = 1$, that is, precisely in the treatments for which we observe a major increase in cooperation.

Treatment:	g=l=2 E=1/8	g=l=1 E=1/8	g=l=2/3 E=1/8	g=l=1/3 E=1/8	g=l=1/3 E=1/16	g=l=1/3 E=0
Strategies						
Vcoop						
Payback	6.5	12.9	19.4	38.8	42.9	48
Grim	3.3	6.5	9.8	19.6	25.7	48
TFT	5.1	10.2	15.3	30.6	35.2	48
δ_{SPE}						
Payback	0.85	0.67	0.55	0.36	0.30	0.25
Grim	0.91	0.70	0.57	0.37	0.30	0.25
TFT	/	/	/	/	/	/
δ_{RD}						
Payback	0.96	0.83	0.73	0.54	0.46	0.4
Grim	>1	0.91	0.80	0.57	0.47	0.4
TFT	>1	0.89	0.76	0.53	0.46	0.4
Size-BAD						
Payback	0.81	0.40	0.27	0.13	0.08	0.05
Grim	>1	0.57	0.38	0.19	0.11	0.05
TFT	>1	0.52	0.35	0.17	0.10	0.05
Cooperation rates						
First Period (FRD)	54%	75%	79%	76%	87%	83%
First Period (Arechar et al.)	33%	50%	/	/	/	/
Overall (FRD)	32%	49%	61%	59%	82%	78%
Overall (Arechar et al.)	21%	37%	/	/	/	/

Notes: The table includes data from Fudenberg et al. (2012) and Arechar et al. (2017). The value of cooperation is obtained from Tables 4 and 6 in FRD, while the critical thresholds derive from our calculations. The cooperation rates are for the last 4 supergames.

Table 10. Theoretical benchmarks and observed cooperation under imperfect monitoring.

The results in the previous subsection already showed that, when we analyze strategies with the SFEM methodology, the estimated prevalence of TFT falls dramatically after accounting for Payback. In Table 11 we report estimates of strategies for all the treatments with noise, separately. The estimated frequency of TFT is 6% when $g = l = 2$ and 0 when $g = l = 1$. On the contrary, the prevalence of Payback surges from 4% when $g = l = 2$ to 13% when $g = l = 1$, and stays constant for treatment with an higher $g = l$. Importantly, the treatment in which Payback's frequency surges ($g = l = 1$) is precisely the one in which Payback becomes sustainable as a risk-dominant equilibrium. This result can only be theoretically justified by the effect of strategic-risk on the choice of the strategy, since Payback can be sustained in a PPE also in the treatment with $g = l = 2$.

	g=l=2		g=l=1		g=l=2/3		g=l=1/3	
Strategies	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Payback	/	0.04**	/	0.13***	/	0.14***	/	0.14***
TFT	0.09***	0.06***	0.08***	0	0.07*	0.02	0.10*	0
ALLC	0.01	0.01	0.02*	0.02	0	0	0.05*	0.10**
TF2T	0.04	0.03**	0.08**	0.03	0.14*	0.07	0.15**	0.10**
TF3T	0.01	0.01	0.01	0.01	0.08	0.06	0.08	0.08**
2TFT	0.08***	0.08***	0.07***	0.07***	0.02	0.02	0.05**	0.05
2TF2T	0.03	0.03	0.10***	0.10***	0.15**	0.17**	0.15**	0.12*
Grim	0.11***	0.11***	0.05**	0.05**	0.12**	0.12***	0.03	0.03
Grim-last2	0.05***	0.05*	0.10***	0.10***	0.02	0.02	0.08**	0.09***
Grim-last3	0.06***	0.06**	0.15***	0.15***	0.25***	0.22***	0.08*	0.07*
ALLD	0.41***	0.42***	0.29***	0.29***	0.14***	0.14***	0.23***	0.22***
D-TFT	0.11***	0.10***	0.05***	0.05**	0.01	0.02	0	0
Gamma	0.45	0.45	0.47	0.46	0.49	0.48	0.45	0.44
N. obs	5244		5022		2214		3076	
Number of subjects	148		130		64		90	

Notes: Bootstrapped standard errors used to calculate p-values. Pooled data from Arechar et al. (2017) and Fudenberg et al. (2012), last 4 supergames.

Table 11. Maximum Likelihood Estimates of strategies: imperfect monitoring and strategic risk.

To sum up, Payback (i) is much more frequently played than TFT when it becomes risk dominant, and (ii) it becomes risk dominant and is played by a larger share of subjects precisely under the parameter configurations in which the cooperation rate increases. This evidence suggests that measures of strategic risk are good predictors of the empirical frequency of cooperation also in games with imperfect monitoring.

5 Discussion and Conclusions

In this paper, we study restitution strategies and their role in restoring cooperation after a defection, with a focus on both their empirical adoption and their theoretical properties. We rely on meta-datasets comprising many previous experimental studies, covering three classes of games: indefinitely repeated Prisoner Dilemmas with perfect and imperfect monitoring, and finitely repeated PD. We find that, in general, restitution strategies are frequently adopted in all three classes of games. With experience, however, they tend to fade away in finitely repeated games, where they cannot be equilibrium strategies. In contrast, their relevance not only persists but even increases in the other classes of repeated games, where these strategies can indeed be equilibria. This suggests that, initially, restitution strategies may be adopted for their "behavioral" appeal, since they resonate with inequity aversion. Eventually, however, the main driver of their empirical success is related to strategic reasons.

The empirical relevance of restitution strategies – which by definition have a memory of two periods or longer – suggests that restricting the focus on memory-one strategies, as common in the recent literature, may lead to overlook important features of subjects' behavior.

To characterize the theoretical properties of restitution strategies, we focus on Payback, which is one of the simplest strategies within this category. We show that Payback can support cooperation as an equilibrium outcome in settings with perfect and imperfect monitoring where Tit-for-tat, and sometimes even Grim trigger, cannot. In addition, we prove that, under some circumstances, Payback can be more profitable and less exposed to strategic risk than Grim trigger. The observed presence of Payback in our three meta-datasets also sheds light on two puzzles that have emerged from previous literature. First, the high estimated frequency of Tit-for-Tat in environments where it is not a subgame-perfect equilibrium or public-perfect equilibrium; and second, the apparent lack of predictive power of strategic risk in games with imperfect monitoring, in contrast to what occurs in games with perfect monitoring. On one hand, accounting for Payback significantly reduces the estimated prevalence of Tit-for-Tat – primarily because restitution strategies have often been misclassified as Tit-for-Tat. On the other hand, the risk dominance of Payback demonstrates robust predictive power even in games with imperfect monitoring.

Why did previous experimental papers fail to identify restitution strategies? The reason cannot lie in the data we use, which have not been collected ad hoc and have already been analyzed by others to empirically estimate strategies. It lies instead in the approaches proposed, which are unsuitable to capture restitution strategies. The issue may concern either the empirical methodology, the experimental design, or both.

First, the most widespread empirical methodology may overlook restitution strategies. When retrieving strategies from observed actions, researchers have to define ex-ante a set of strategies, and then estimate their likely presence in the experimental data as in Dal Bó and Fréchette (2011) and many subsequent studies. None of the studies following this methodology have included restitution strategies in their set. Detecting restitution strategies with this methodology is anyway challenging, as they differ from other conditionally cooperative strategies only in a relatively small subset of histories, especially in games with perfect monitoring. Consequently, if one does not specifically examine those instances – as we did in this paper –

disentangling strategies becomes nearly impossible, as differences may appear minor and can be easily attributed to random errors. A similar issue emerges with the approach proposed by Breitmoser (2015), which considers stochastic strategies but cannot identify a restitution strategy like Payback. This approach focuses on memory-one strategies, conditioning only on the combination of actions of the two players in the previous period. It would therefore classify Payback within the class of "semi-grim" strategies, because in a period that follows the opponent's unilateral defection Payback may prescribe either defection or cooperation, depending on the subject's "standing".

Second, the design of some experiments may rule out the possibility for participants to employ restitution strategies. The methodologies developed by Dal Bó and Fréchette (2019) and Romero and Rosokha (2018, 2019, 2023) to directly elicit subjects' strategies are clever and effective but are affected by this type of limitation. The former requires subjects to choose from a pre-specified list of options, which did not include Payback or any other restitution strategy. The latter allows subjects to create their own strategy conditioning on past outcomes, but without considering the subjects' standing, which excludes the possibility of constructing a strategy such as Payback.

As Dal Bó and Fréchette (2019) have already pointed out, uncovering the strategies that subjects employ is crucial for several reasons. For instance, identifying what strategies are used to support cooperation can provide a more effective test of theory than relying solely on observable outcomes. Furthermore, it allows scholars to focus on a limited set of empirically relevant strategies for a dialogue with theory, for example in evaluating which strategies are likely to survive evolutionary pressure. Lastly, a better understanding of strategies can aid in designing setups and institutions that promote the emergence and persistence of cooperation (or foster competition, in case cooperation in the game implies tacit collusion against third parties).

An implication of our results is to reinstate trust in the theoretical tools that are generally used to predict cooperation, such as Public Perfect Equilibrium and risk dominance, that were questioned by the overestimation of TFT in experiments. We bridge the gap between theoretical models – extensively using asymmetric restitution strategies and showing that they are evolutionary stable – and the experimental literature that attributed them a minor role so far. More broadly, this study sheds light on the factors contributing to the sustainability of cooperation in the long run. While experimental super-games are relatively short, real-world interactions often last so long that defections – whether intentional or unintentional – eventually take place. Identifying how people restore cooperation after a breach of trust is essential for our grasp of how cooperation endures over time.

Appendix

A Theoretical analysis of the Payback strategy

A.1 Perfect Monitoring

Proof of Proposition 1

We start by deriving the critical discount factors for the Grim and Payback strategies.

Grim For the Grim strategy, the equilibrium path payoff is

$$V_C = \frac{1}{1-\delta}.$$

The only relevant IC constraint is in the cooperative state and Grim is an equilibrium if and only if the deviation payoff in the cooperative state is lower than the on-path payoff, which implies the critical discount factor:

$$1 + g \leq V_C \iff \delta \geq \frac{g}{1+g}.$$

Payback The Payback strategy has three states: (C, P, R) (Cooperation, Punishment and Restitution). The on-path payoffs in these three states are:

$$\begin{aligned} V_C &= \frac{1}{1-\delta}, \\ V_P &= 1 + g + \delta V_C, \\ V_R &= -l + \delta V_C. \end{aligned}$$

The IC constraint in the P state is always satisfied because the player is recommended the dominant stage-game action, and the continuation payoff does not depend on his action.

The IC constraint in the C state is

$$1 + g + \delta V_R \leq 1 + \delta V_C \iff \delta \geq \frac{g}{1+l}.$$

The IC constraint in the R state is

$$0 + \delta V_R \leq -l + \delta V_C \iff \delta \geq \frac{l}{1+l}.$$

Combining these two constraints, the threshold discount factor for Payback to be a subgame perfect equilibrium is:

$$\frac{\max\{l, g\}}{1+l}.$$

For part (i) note that since $g < 1+l$, the ratio is guaranteed to be less than 1. For part (ii), note that if $g = l$, the threshold discount factors for the two strategies coincide. However, for all $g \neq l$ the threshold discount factor for Payback is higher. $\square$

A.2 Imperfect Monitoring

Recall that in Section 3.3 we assume $g = l$. We claim that whenever $g = l$, an equivalent representation of the Prisoner's Dilemma is the one in table 12b, in which cooperation and defection take the "benefit/cost" form, in which to cooperate, players have to pay a cost c to give a benefit b to the other player, while defection gives 0 to each party.

Player 1 \ Player 2	C	D
C	$1, 1$	$-l, 1+g$
D	$1-g, -l$	$0, 0$

12a) Standardized PD.

Player 1 \ Player 2	C	D
C	$b-c, b-c$	$-c, b$
D	$b, -c$	$0, 0$

12b) Benefit-cost Representation.

Table 12 Equivalence between two representations of a PD.

Specifically, to move from the first representation to the second, substitute $g = l = \frac{c}{b-c}$ and multiply all the payoffs by $b - c$. We assume $b > c \geq 0$ so that the stage game in this new representation is still a PD.

Recall that we also assume that in every period, each intended action is independently changed with probability $E < \frac{1}{2}$. Denoting $G \equiv \frac{b}{c} = \frac{1+g}{g} > 1$, we can write the stage-game expected payoffs of player 1 as:

Player 1 \ Player 2	C	D
C	$c(G-1)(1-E)$	$c(E(1+G)-1)$
D	$c(G-E(1+G))$	$cE(G-1)$

Table 13 Expected Payoffs of Player 1 in the PD with Imperfect Monitoring.

These expected stage-game payoffs constitute a PD because $E \in (0, \frac{1}{2})$ and $G > 1$. Since c scales all expected payoffs equally, without loss of generality, we further normalize $c = 2$ (as in Fudenberg et al., 2012) to describe the stage game with just two parameters (G and E).

A.2.1 Proof of Proposition 2

Payback The expected payoffs in the three states of Payback are, respectively:

$$\begin{aligned} V_C &= 2(G-1)(1-E) + \delta(V_C(1-E)^2 + (1-E)EV_P + (1-E)EV_R + E^2V_C), \\ V_P &= 2G - 2E(1+G) + \delta(V_C(1-E)^2 + (1-E)EV_C + (1-E)EV_P + E^2V_P), \\ V_R &= 2(E(1+G)-1) + \delta(V_C(1-E)^2 + (1-E)EV_C + (1-E)EV_R + E^2V_R). \end{aligned}$$

These expressions can be simplified to:

$$\begin{aligned} V_C &= 2\frac{(G-1)(1-E)}{(1-\delta)(1+E\delta(1-2E))} \\ V_P &= \frac{2G(1-E) - 2E + \delta(1-E)V_C}{1-\delta E} \\ V_R &= \frac{2(E(1+G)-1) + \delta(1-E)V_C}{1-\delta E} \end{aligned}$$

To check whether Payback is a (Perfect Public) equilibrium, we need to check the IC constraints in the three states:

Step 1: IC constraints in State C.

A defection in State C gives payoff

$$V_{DC} = 2G - 2E(1+G) + \delta(V_R(1-E)^2 + (1-E)EV_C + (1-E)EV_C + E^2V_P)$$

For this not to be profitable, we must have

$$V_{DC} \leq V_C$$

The threshold discount factor for state C solves:

$$\begin{aligned} & 2G - 2E(1+G) + \delta(V_R(1-E)^2 + (1-E)EV_C + (1-E)EV_C + E^2V_P) \\ = \; & 2(G-1)(1-E) + \delta(V_C(1-E)^2 + (1-E)EV_P + (1-E)EV_R + E^2V_C). \end{aligned}$$

The solution is:

$$\delta_C = \frac{2}{V_C - V_R + E(V_P + V_R - 2V_C)} \tag{1}$$

Plugging in the formulas for the payoffs and solving for the threshold δ we get that it is:

$$\delta_C = \frac{-G(1-2E)(1-E) - E + \sqrt{(G(1-2E)(1-E)+E)^2 + 4E(1-2E)(1-E)}}{2E(1-2E)(1-E)}$$

Step 2: IC constraints in State P.

The IC constraint is always satisfied in state P. Note that the transitions from that state are independent of player 1 actions. Hence the continuation payoff of that player does not depend on his action and hence the best response is just myopic best response which is the recommended action in that state.

Step 3: IC constraints in State R.

In state R, the recommended action is C. A deviation to D yields expected payoff:

$$V_{DR} = 2E(G-1) + \delta(V_R(1-E) + EV_C).$$

For the deviation not to be profitable, we must have

$$V_{DR} \leq V_R = 2(E(1+G) - 1) + \delta(V_C(1-E) + EV_R).$$

The threshold discount factor in that state is:

$$\delta_R = \frac{2}{V_C - V_R}.$$

Recall from (1) that the threshold discount factor in state C satisfies

$$\delta_C = \frac{2}{V_C - V_R + E(V_P + V_R - 2V_C)}.$$

Since

$$V_P + V_R - 2V_C = -2\frac{(G-1)(1-2E)}{1+\delta E - 2E^2\delta} < 0,$$

we have that $\delta_C \geq \delta_R$ for all parameters. Hence Payback is an equilibrium if and only if

$$\delta \geq \delta_{PPE,Payback} \equiv \frac{-G(1-2E)(1-E) - E + \sqrt{(G(1-2E)(1-E)+E)^2 + 4E(1-2E)(1-E)}}{2E(1-2E)(1-E)}.$$

Grim. The Grim Strategy has two states: (C, P) (Cooperate, Punishment). The expected payoffs in the two states are, respectively:

$$\begin{aligned} V_C &= 2(G-1)(1-E) + \delta(V_C(1-E)^2 + (1-(1-E)^2)V_P), \\ V_P &= \frac{2E(G-1)}{1-\delta}. \end{aligned}$$

Solving for V_C we get:

$$V_C = 2(G-1)\frac{(1-\delta)(1-E) + E^2\delta(2-E)}{(1-\delta)(1-\delta+E\delta(2-E))}.$$

The IC constraint in the Punishment state is always satisfied since the players play infinite repetition of the static Nash equilibrium. So to check under what conditions Grim is a PPE, we need to only check the IC constraint in the Cooperate state, which is

$$2G - 2E(1+G) + \delta((1-E)EV_C + (1-(1-E)E)V_P) \leq V_C.$$

It holds when:

$$\delta \leq \delta_{PPE,Grim} \equiv \frac{1}{(1-E)(G(1-2E)+E)}.$$

Threshold $\delta_{PPE,Grim}$ is less than 1 if and only if $G = \frac{1+g}{g} > \frac{E}{(1-2E)(1-E)}$, as claimed in part (i) of Proposition 2 (we show next that the threshold for Payback is lower, hence if δ is high enough for Grim to be a PPE, it is also high enough for Payback to be PPE).

Comparing the threshold factors for Grim and Payback Define $X \equiv (G(1-2E)(1-E)+E)$ and $A \equiv E(1-2E)(1-E)$. In this notation, we can write:

$$\begin{aligned} \delta_{PPE,Payback} &\equiv \frac{-X+\sqrt{X^2+4A}}{2A}, \\ \delta_{PPE,Grim} &= \frac{1}{X-E^2}. \end{aligned}$$

Let Z be the ratio of these thresholds:

$$Z \equiv \frac{\delta_{PPE,PB}}{\delta_{PPE,Grim}} = \frac{(X-E^2)(-X+\sqrt{X^2+4A})}{2A}.$$

First, note that the limit of Z as $G \to \infty$ is 1:

$$\lim_{X \to +\infty} \frac{(X - E^2)(-X + \sqrt{X^2 + 4A})}{2A} = 1.$$

Using this limit, we show that $Z < 1$ for all parameters by showing that Z is increasing in G.

The derivative of the numerator of Z with respect to X is:

$$\begin{aligned} & \frac{\partial}{\partial X}((X - E^2)(-X + \sqrt{X^2 + 4A})) \\ = \quad & -\frac{(2X - E^2)\sqrt{X^2 + 4A} - 2X^2 - 4A + E^2 X}{\sqrt{X^2 + 4A}}. \end{aligned}$$

Note that this expression is zero at $A = 0$, and positive for all $A > 0$. Therefore, for all parameters $\frac{\partial Z}{\partial X} > 0$, which implies that Z is strictly increasing in G. Hence, for all parameters, $Z < 1$. This establishes the rest of the claims in Proposition 2. □

A.2.2 Proof of Proposition 3

Payback vs. Grim The expected PPE payoff for Payback is:

$$V_{Payback} = 2\frac{(G - 1)(1 - E)}{(1 - \delta)(1 + E\delta(1 - 2E))}.$$

while for Grim:

$$V_{Grim} = 2(G - 1)\frac{(1 - \delta)(1 - E) + E^2\delta(2 - E)}{(1 - \delta)(1 - \delta + E\delta(2 - E))}.$$

Therefore, to show $V_{Payback} > V_{Grim}$ we need to show that:

$$\frac{1 - E}{1 + E\delta(1 - 2E)} > \frac{(1 - \delta)(1 - E) + E^2\delta(2 - E)}{(1 - \delta + E\delta(2 - E))}.$$

An algebraic manipulation shows that this condition holds if and only if:

$$\delta(1 - E - 2E^2 + E^3) + 1 > 0.$$

For all $E < \frac{1}{2}$, this expression is positive, and that establishes the comparison between the equilibrium payoffs for Payback and Grim.

Payback vs. TFT If players follow TFT (starting in the Cooperative state), the expected payoff is:

$$V_{TFT} = 2(G-1)\frac{1-E+\delta(2E-1)}{(1-\delta)(1+\delta(2E-1))}.$$

Using our computation for $V_{Payback}$, we need to show that:

$$\frac{1-E}{(1+E\delta(1-2E))} > \frac{1-E+\delta(2E-1)}{1+\delta(2E-1)}.$$

After some algebra, this is equivalent to

$$\delta > -\frac{E^2}{E(1-2E)},$$

which is always satisfied since $\delta > 0$ and the r.h.s. is negative. $\square$

A.2.3 Proof of Proposition 4

We start with calculations for the size of the basin of attraction (BAD) for Payback, Grim and TFT when compared to AllD.

BAD for Payback. When I play Payback and meet another player that plays Payback, my expected payoff is

$$V_{Payback,Payback} = 2\frac{(G-1)(1-E)}{(1-\delta)(1+E\delta(1-2E))}.$$

When I meet a player that plays AllD, my expected payoff is a solution to the system that describes expected payoffs in the three states of Payback:

$$\begin{aligned}
U_C &= 2(E(1+G)-1)+\delta((1-E)^2U_P+2E(1-E)U_C+E^2U_R),\\
U_P &= 2E(G-1)+\delta((1-E)U_P+EU_C),\\
U_R &= 2(E(1+G)-1)+\delta((1-E)U_C+EU_R).
\end{aligned}$$

The solution for the expected payoff at the beginning of the game is:

$$V_{Payback,ALLD} = 2\frac{\delta(1-E)^2(2\delta E^2-\delta E+1)-2\delta E^2G(1-E)+E(1+G)-1}{(1-\delta)(1-2\delta E(1-E))}.$$

If instead I play AllD and I meet a player that plays AllD, I get a payoff:

$$V_{ALLD,ALLD} = \frac{2E(G-1)}{1-\delta}.$$

Finally, if I play AllD and I meet a player that plays Payback, I get a payoff that is a solution to the following system of equations (where subscripts now represent the state of Payback of the other player).

$$\begin{aligned}
U_C &= 2G - 2E(1+G) + \delta((1-E)^2 U_P + 2E(1-E)U_C + E^2 U_R), \\
U_P &= 2E(G-1) + \delta((1-E)U_P + EU_C), \\
U_R &= 2G - 2E(1+G) + \delta((1-E)U_C + EU_R).
\end{aligned}$$

Solving this system yields expected payoff at the beginning of the game:

$$V_{ALLD,Payback} = 2\frac{G(1-\delta) - (1+G)E + EG(1-2E)(1-E)^2\delta^2 - E\delta(-2G + EG - 2E + 2E^2)}{(1-\delta)(1-2\delta E(1-E))}.$$

Let α be the fraction of other players using the Payback strategy. Player 1 is indifferent between AllD and Payback if

$$\alpha V_{Payback,Payback} + (1-\alpha)V_{Payback,ALLD} = \alpha V_{ALLD,Payback} + (1-\alpha)V_{ALLD,ALLD}.$$

Solving for α we get:

$$\alpha^*_{Payback} = \frac{(1-\delta E(1-E))(1+\delta E(1-2E))(1-\delta(1-E))}{(1-2E)(1-E)(1-\delta^2 E^2(1-E))(G-1)\delta}.$$

BAD for Grim. When I play Grim and meet another player that plays Grim, my expected payoff is

$$V_{Grim,Grim} = 2(G-1)\frac{(1-\delta)(1-E) + E^2\delta(2-E)}{(1-\delta)(1-\delta+E\delta(2-E))}.$$

When I play Grim and meet a player that plays AllD, my expected payoff is a solution to the system (where subscripts represent the two states of my Grim strategy):

$$\begin{aligned}
U_C &= 2(E(1+G)-1) + \delta((1-E)EU_C + (1-(1-E)E)U_P) \\
U_P &= \frac{2E(G-1)}{1-\delta},
\end{aligned}$$

The resulting expected payoff is:

$$V_{Grim,ALLD} = 2\frac{\delta E^2(G-1)(-1+E) + (G+1-2\delta)E - 1 + \delta}{(1-\delta)(1-\delta E+\delta E^2).}$$

Finally, if I play AllD and I meet a player that plays Grim, I get a payoff that is a solution to the following system of equations (where subscripts represent now the states of the other player):

$$\begin{aligned} U_C &= 2G - 2E(1+G) + \delta((1-E)EU_C + (1-(1-E)E)U_P), \\ U_P &= \frac{2E(G-1)}{1-\delta}. \end{aligned}$$

The resulting expected payoff is:

$$V_{ALLD,Grim} = 2\frac{G(1-\delta) - \delta E^2(G-1)(1-E) - (1+G(1-2\delta))E}{(1-\delta)(-\delta E + \delta E^2 + 1)}.$$

Let α be the fraction of other players using the Grim strategy. Player 1 is indifferent between AllD and Grim if

$$\alpha V_{Grim,Grim} + (1-\alpha)V_{Grim,ALLD} = \alpha V_{ALLD,Grim} + (1-\alpha)V_{ALLD,ALLD}.$$

Solving for α we get:

$$\alpha^*_{Grim} = \frac{1-\delta+E\delta(2-E)}{\delta(1-2E)(1-E)(G-1)}.$$

BAD for TFT. When two players that play TFT meet, each has expected payoff of:

$$V_{TFT,TFT} = 2(G-1)\frac{1-E+\delta(2E-1)}{(1-\delta)(1+\delta(2E-1))}$$

When I play TFT and meet a player that plays AllD, my expected payoff is:

$$V_{TFT,ALLD} = \frac{2(E(G+1)-1)+\delta(1-E)(1-2E)}{1-\delta}$$

If I play AllD and meet a player that plays TFT, I get expected payoff :

$$V_{ALLD,TFT} = \frac{2(G(1-E)(1+\delta(-1+2E))-E)}{1-\delta}$$

Let α be the fraction of other players using the TFT strategy. Player 1 is indifferent between AllD and TFT if

$$\alpha V_{TFT,TFT} + (1-\alpha)V_{TFT,ALLD} = \alpha V_{ALLD,TFT} + (1-\alpha)V_{ALLD,ALLD}.$$

Solving for α we get:

$$\alpha^*_{TFT} = \frac{1-\delta+2\delta E}{(G-1)\delta(1-2E)}$$

Comparison: Grim vs. Payback Using our formulas for α^*_{Grim} and $\alpha^*_{Payback}$ we get that the difference is:

$$\begin{aligned} & \alpha^*_{Grim} - \alpha^*_{Payback} \\ = \quad & E\frac{E^4\delta^2 + 2\delta(1-\delta)E^3 + 1 - \delta E(1+E(1-\delta))}{(1-2E)(1-E)(1-\delta^2E^2+\delta^2E^3)(G-1)} \end{aligned}$$

The denominator is positive, and the numerator is positive because

$$1 - \delta E(1+E(1-\delta)) \geq 1 - E > 0.$$

Comparison: TFT vs. Payback We want to show that $\alpha^*_{Payback} < \alpha^*_{TFT}$, which is:

$$\frac{(1-\delta E(1-E))(1+\delta E(1-2E))(1-\delta(1-E))}{(1-E)(1-\delta^2E^2(1-E))} < 1 - \delta(1-2E)$$

which can be rewritten as:

$$\frac{(1-\delta E(1-E))}{(1-\delta^2E^2(1-E))} < \frac{(1-\delta(1-2E))(1-E)}{(1+\delta E(1-2E))(1-\delta(1-E))}.$$

Since the l.h.s. is lower than 1, it is enough to show that the r.h.s. is greater than 1, which is equivalent to:

$$1 - \delta(1-2E) - E + E\delta(1-2E) > 1 + \delta E(1-2E) - \delta(1-E)(1+\delta E(1-2E)).$$

Collecting terms, this is equivalent to:

$$1 < 1 + (1-E)E\delta,$$

which is true for any $(E,\delta) \in (0, \frac{1}{2}) \times (0,1)$. $\square$

References

Abreu, D. (1988). On the theory of infinitely repeated games with discounting. *Econometrica*, *56*(2), 383–396.

Aoyagi, M., Bhaskar, V., & Fréchette, G. R. (2019). The impact of monitoring in infinitely repeated games: Perfect, public, and private. *American Economic Journal: Microeconomics*, *11*(1), 1–43.

Aoyagi, M., Fréchette, G. R., & Yuksel, S. (2024). *Beliefs in Repeated Games* (Working Paper).

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

Athey, S., & Bagwell, K. (2001). Optimal collusion with private information. *RAND Journal of Economics*, 428–465.

Bigoni, M., Camera, G., & Casari, M. (2013). Strategies of cooperation and punishment among students and clerical workers. *Journal of Economic Behavior & Organization*, *94*, 172–182.

Blonski, M., Ockenfels, P., & Spagnolo, G. (2011). Equilibrium selection in the repeated prisoner's dilemma: Axiomatic approach and experimental evidence. *American Economic Journal: Microeconomics*, *3*(3), 164–92.

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

Boczo, M., Vespa, E., Weidman, T., & Wilson, A. J. (2023). Testing models of strategic uncertainty: Equilibrium selection in repeated games. *Working Paper*.

Boerlijst, M. C., Nowak, M. A., & Sigmund, K. (1997). The logic of contrition. *Journal of Theoretical Biology*, *185*(3), 281–293.

Boyd, R. (1989). Mistakes allow evolutionary stability in the repeated prisoner's dilemma game. *Journal of Theoretical Biology*, *136*(1), 47–56.

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

Calford, E., & Oprea, R. (2017). Continuity, Inertia, and Strategic Uncertainty: A Test of the Theory of Continuous Time Games. *Econometrica*, *85*(3), 915–935.

Charness, G., & Rabin, M. (2002). Understanding social preferences with simple tests. *The Quarterly Journal of Economics*, *117*(3), 817–869.

Dal Bó, P., & Fréchette, G. R. (2011). The evolution of cooperation in infinitely repeated games: Experimental evidence. *American Economic Review*, *101*(1), 411–29.

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

Dal Bó, P., & Fréchette, G. R. (2019). Strategy choice in the infinitely repeated prisoner's dilemma. *American Economic Review*, *109*(11), 3929–52.

Dreber, A., Fudenberg, D., & Rand, D. G. (2014). Who cooperates in repeated games: The role of altruism, inequity aversion, and demographics. *Journal of Economic Behavior & Organization*, *98*, 41–55.

Dvorak, F., & Fehrler, S. (2023). Negotiating cooperation under uncertainty: Communication in noisy, indefinitely repeated interactions. *American Economic Journal: Microeconomics, forthcoming.*

Embrey, M., Frechette, G. R., & Yuksel, S. (2017). Cooperation in the finitely repeated prisoner's dilemma. *The Quarterly Journal of Economics*, *133*(1).

Embrey, M., Mengel, F., & Peeters, R. (2016). Eliciting Strategies in Indefinitely Repeated Games of Strategic Substitutes and Complements. *SSRN Scholarly Paper: ID 2896321.*

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, *114*(3), 817–868.

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

Fudenberg, D., Rand, D. G., & Dreber, A. (2012). Slow to anger and fast to forgive: Cooperation in an uncertain world. *American Economic Review*, *102*(2), 720–49.

Ghidoni, R., & Suetens, S. (2022). The Effect of Sequentiality on Cooperation in Repeated Games. *American Economic Journal: Microeconomics*, *14*(4), 58–77.

Graser, C., & van Veelen, M. (2024). Repeated prisoner's dilemmas with errors: How much subgame-perfection, how much forgiveness, and how much cooperation? *Tinbergen Institute Discussion Paper: TI 2024-022/I.*

Green, E. J., & Porter, R. H. (1984). Noncooperative collusion under imperfect price information. *Econometrica*, 87–100.

Harrington, J. E. (2006). How do cartels operate? *Foundations and trends in microeconomics* (pp. 1–105). Now Publishers Inc.

Harrington, J. E., & Skrzypacz, A. (2007). Collusion under monitoring of sales. *The RAND Journal of Economics*, *38*(2), 314–331.

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

Levin, J. (2003). Relational incentive contracts. *American Economic Review*, *93*(3), 835–857.

Martinez-Martinez, I., & Normann, H.-T. (2024). Cooperation in Multiplayer Dilemmas. *Working Paper.*

Normann, H.-T., & Sternberg, M. (2023). Human-algorithm interaction: Algorithmic pricing in hybrid laboratory markets. *European Economic Review*, *152*, 104347.

Oprea, R. (2020). What makes a rule complex? *American Economic Review*, *110*(12), 3913–51.

Proto, E., Rustichini, A., & Sofianos, A. (2021). Intelligence, Errors, and Cooperation in Repeated Interactions. *The Review of Economic Studies*.

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

Romero, J., & Rosokha, Y. (2019). The Evolution of Cooperation: The Role of Costly Strategy Adjustments. *American Economic Journal: Microeconomics*, *11*(1), 299–328.

Romero, J., & Rosokha, Y. (2023). Mixed Strategies in the Indefinitely Repeated Prisoner's Dilemma. *Econometrica*, *91*(6), 2295–2331.

Salant, Y. (2011). Procedural analysis of choice rules with applications to bounded rationality. *American Economic Review*, *101*(2), 724–48.

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

Sugden, R. (1986). *The economics of rights, cooperation and welfare*. Springer.

Van Damme, E. (1989). Renegotiation-proof equilibria in repeated prisoners' dilemma. *Journal of Economic theory*, *47*(1), 206–217.

Wu, J., & Axelrod, R. (1995). How to cope with noise in the iterated prisoner's dilemma. *Journal of Conflict Resolution*, *39*(1), 183–189.