Cooperative Equilibria in Iterated Social Dilemmas

Valerio Capraro[1], Matteo Venanzi[2], Maria Polukarov[2], and Nicholas R. Jennings[2]

[1] Mathematics, University of Southampton, United Kingdom
V.Capraro@soton.ac.uk
[2] Electronics and Computer Science, University of Southampton, United Kingdom
{mv1g10,mp3,nrj}@ecs.soton.ac.uk

Abstract. The implausibility of the extreme rationality assumptions of Nash equilibrium has been attested by numerous experimental studies with human players. In particular, the fundamental social dilemmas such as the Traveler's dilemma, the Prisoner's dilemma, and the Public Goods game demonstrate high rates of deviation from the unique Nash equilibrium, dependent on the game parameters or the environment in which the game is played. These results inspired several attempts to develop suitable solution concepts to more accurately explain human behaviour. In this line, the recently proposed notion of cooperative equilibrium, [5], [6], based on the idea that players have a natural attitude to cooperation, has shown promising results for single-shot games. In this paper, we extend this approach to iterated settings. Specifically, we define the Iterated Cooperative Equilibrium (ICE) and show it makes statistically precise predictions of population average behaviour in the aforementioned domains. Importantly, the definition of ICE does not involve any free parameters, and so it is fully predictive.

1 Introduction

The standard assumption of economic models that players in strategic situations act perfectly rationally has been constantly rejected by numerous experiments over the years. These experiments, typically conducted on the fundamental social dilemmas such as the Prisoner's dilemma, the Traveler's dilemma, and the Public Goods game, have shown that cooperation between players (associated with the deviation from the unique, but inefficient, Nash equilibrium) is frequent, and appears to depend on both the game parameters and the environment in which the game is played. In particular, it has been observed that the rate of cooperation in the Traveler's dilemma depends on the bonus/penalty value, whenever the game is single-shot or iterated [7], [12]; the rate of cooperation in the Prisoner's dilemma depends on the payoff parameters or the way the players are matched to play together [11], [32]; and the rate of cooperation in the Public Goods game depends on the marginal return or on the frequency of interaction between free-riders and cooperators [13], [14] [17].

Considerable research efforts have been made in attempt to explain deviations from Nash equilibria. Some methods developed to this end are based on the idea that humans have bounded rationality and/or can make mistakes in computations[3] [4], [9], [20], [25]; others explain cooperation in terms of evolution [1], [3], [10], [19], [21],

[3] See [31] for a recent parallelism among these approaches.

[22], [23], [29]. Finally, much of work has been directed towards defining profoundly different solution concepts [24], [26], especially in the recent algorithmic game theory and artificial intelligence communities [2], [8], [15], [16], [18], [27], [30]. This interest is particularly motivated by the emerging applications of human-agent collectives, where artificial agents interact with humans. To build such systems effectively, it is highly important to understand and find accurate methods to predict human behaviour.

To this end, a new solution concept, termed *cooperative equilibrium*, has been recently proposed for one-shot games [5], [6]. This approach is inspired by the aforementioned experimental findings, which suggest that players are conditionally cooperative—that is, the same player may act more or less cooperatively in the same game scenario, depending on the actual payoffs. In other words, humans have an attitude to cooperation by nature: they do not act a priori as single players, but rather forecast how the game would have been played if they formed coalitions and then select actions according to their best forecast. It turns out, that direct implementation of this idea can predict human behaviour with impressively high precision, as demonstrated in [5], [6] on the aforementioned social dilemmas.

In this paper, we further explore this direction and extend the cooperatve equilibrium approach to iterated settings. Specifically, we define the Iterated Cooperative Equilibrium (ICE), that combines this concept with some ideas developed in [7] for iterated games. Importantly, in contrast to other methods, ICE does not use any free parameters, and thus is fully predictive.

We then evaluate our method on the iterated Traveler's dilemma, the Prisoner's dilemma, and the Public Goods game. To this end, we make use of the experimental data provided in [7], [32] and [14] for these three domains, respectively.[4] Our results confirm that the ICE makes accurate predictions of population average behaviour in social dilemmas. In particular, it clearly outperforms the Logit Learning Model (LLM) developed in [7] for the Traveler's dilemma.

The paper unfolds as follows. In Section 2 we define the social dilemmas in consideration. In Section 3 we formalise our approach. We then apply it to the iterative Traveler's dilemma in Section 4, to the Prisoner's dilemma in Section 5, and to the Public Goods game in Section 6. Section 7 concludes with directions for future work.

2 Preliminaries

We start with the definitions of the social dilemmas in consideration of this paper.

Prisoner's dilemma. Two players can choose to either cooperate (C) or defect (D). If both players cooperate, each receives the monetary reward, R, for cooperating. If one player defects and the other cooperates, then the defector receives the temptation payoff, T, while the other receives the sucker payoff, S. If both players defect, they both receive the punishment payoff, P. Payoffs are subjected to the condition $T > R > P > S$.

Traveler's dilemma. Two travelers need to claim for a reimbursement between L and H monetary units for their (identical) luggage that has been lost by the same

[4] These were the only sources we could find that reported sufficient data for our purposes.

air company. To avoid high claims, the air company employs the following rule: the traveler who makes a lower claim, say m, gets a reimbursement of $m + b$ monetary units, and the other one gets a reimbursement of $m - b$ monetary units, for a fixed value of bonus/penalty, b. If both players claim the same amount, m, then they both get reimbursed by m monetary units.

Public Goods game. n players receive an initial endowment of $y > 0$ monetary units each and simultaneously choose an amount $0 \leq x_i \leq y$ to contribute to a public pool. The total amount in the pot is multiplied by α_0 and then divided equally among all group members. Thus, player i's utility is $u_i(x_1, \ldots, x_n) = y - x_i + \alpha(x_1 + \ldots + x_n)$, where $\alpha = \frac{\alpha_0}{n}$. The number α is termed the *constant marginal return* and assumed to belong to the interval $\left(\frac{1}{n}, 1\right)$.

3 Iterated Cooperative Equilibrium

We now introduce the concept of iterated cooperative equilibrium for the aforementioned social dilemmas.

Let $\mathcal{G} = (N, (S_i, u_i)_{i \in N})$ be a normal-form game with a set N of n players, and for all $i \in N$, a finite set of strategies S_i and a monetary payoff function $u_i : S \to \mathbb{R}$, where $S = \times_{j \in N} S_j$. As usual, we use $-i$ to denote the set $N \setminus \{i\}$ of all players but i. We denote by $\Delta(X)$ the set of probability distributions on a finite set X. Thus, $\Delta(S_i)$ defines the set of mixed strategies for player $i \in N$, and his expected payoff from a mixed strategy profile σ is given by $u_i(\sigma) = \sum_{s \in S} u_i(s)\sigma_1(s_1) \cdot \ldots \cdot \sigma_n(s_n)$.

The idea behind our approach is as follows. Suppose each agent i simply considers two possible scenarios: the fully selfish play p_s, where players take individual actions pursuing their private interests, and the fully cooperative play p_c, where players are assumed to pursue the collective interest. With each scenario p we associate a value $v_i(p)$, defined as an average $v_i(p) = e_i(p)\tau_i(p) + e_i(\overline{p})\tau_i(\overline{p})$, where, roughly speaking,

- $\tau_i(p)$ is the probability that all players follow scenario p, and $\tau_i(\overline{p}) = 1 - \tau_i(p)$ is the probability that (at least one of) the players $-i$ will deviate from p for the sake of their individual interests, knowing that player i follows scenario p. In particular, this implies that $\tau_i(\overline{p_s}) = 0$, since a Nash equilibrium cannot be improved by unilateral deviations;
- $e_i(p)$ is the payoff of i when scenario p is realised, and $e_i(\overline{p})$ is the infimum of gains player i achieves when other players deviate from p.

Then, the values $v_i(p)$ determine each player i's strategy as follows. Let $p_i^* \in \{p_s, p_c\}$ be the scenario that maximises the function v_i, and define the induced game $\mathcal{G}(p_i^*)$ to be the restriction of $\mathcal{G}$ where the set of allowed mixed strategy profiles is given by $\{\sigma | u_i(\sigma) \geq v_i(p_i^*)\}$. Since this set is convex and compact, the induced game has Nash equilibria. The cooperative equilibrium is then given by a combination of strategies where each player i plays according to a Nash equilibrium of his induced game.

Formalising this idea is not completely trivial: while the payoffs e_i seem straightforward to define, the probabilities τ_i are much more delicate, since the event "players $-i$ deviate from scenario p_c" is not measurable in any universal sense. In iterated settings, we can approach this problem applying a sort of fictitious play. Specifically, we start with initial values $\tau_i(p_c) = \tau_i(\overline{p_c}) = \frac{1}{2}$, and then at each step we update these probabilities using observations made in previous rounds. To this end, we use the standard method for probabilistic modelling of binary random events based on the beta family of probability density functions [28]. If in the first round player i has observed cooperation, then $\tau_i(p_c)$ grows from $\frac{1}{2}$ to $\frac{2}{3}$, otherwise it drops from $\frac{1}{2}$ to $\frac{1}{3}$ and so forth: that is, if k is the number of cooperative plays observed in periods from 1 to $t-1$, then $\tau_i^{(t)}(p_c)$ is updated to $\frac{k+1}{t+1}$. We now define this procedure in detail.

Let $\mathcal{G} \in \{\text{Prisoner's dilemma, Traveler's dilemma, Public Goods}\}$. Then, $\mathcal{G}$ has a unique Nash equilibrium, $NE(\mathcal{G})$. Moreover, there is also a unique Pareto optimal strategy profile, $OPT(\mathcal{G})$.

For each period $t \geq 1$, we set $v_i^{(t)}(p_s) = u_i\left(NE(\mathcal{G})\right)$ and $e_i^{(t)}(p_c) = u_i\left(OPT(\mathcal{G})\right)$. For other parameters, we consider the first and the later rounds separately.

Period 1. We define:

- $e_i^{(1)}(\overline{p_c}) = \inf\{u_i(\sigma)|\sigma_i = OPT(\mathcal{G})_i; \forall j \neq i, u_j(\sigma_j, OPT(\mathcal{G})_{-j}) \geq u_j(OPT(\mathcal{G}))\}$ is the infimum payoff that player i obtains when he plays according to the Pareto optimum, while other players deviate from this profile if the corresponding *unilateral* deviation weakly improves the payoff to each deviator;
- $\tau_i^{(1)}(p_c) = \tau_i^{(1)}(\overline{p_c}) = \frac{1}{2}$;
- $v_i^{(1)}(p_c) = \tau_i^{(1)}(p_c)e_i^{(1)}(p_c) + \tau_i^{(1)}(\overline{p_c})e_i^{(1)}(\overline{p_c})$;
- $v_i^{(1)} = \max\{v_i^{(1)}(p_s), v_i^{(1)}(p_c)\}$;
- Ind$(\mathcal{G}, i, 1)$ is the restriction of game $\mathcal{G}$ where the set of allowed mixed strategy profiles is limited to $\left\{\sigma | u_j(\sigma) \geq v_i^{(1)}, \forall j\right\}$.

Period t. We update payoffs e_i and probabilities τ_i as follows.

- Let σ_{-i} be the *average* of strategies played by players $-i$ in periods from 1 to $t-1$. Then, $e_i^{(t)}(\overline{p_c}) = u_i(OPT(\mathcal{G})_i, \sigma_{-i})$;
- Let $\sigma_{-i}^{(s)}$ be the strategy played by players $-i$ in period $s < t$. We say that $\sigma_{-i}^{(s)}$ is a *cooperation* if there is a strategy $\sigma_i \neq (NE(\mathcal{G}))_i$ such that $(\sigma_i, \sigma_{-i}^{(s)})$ is allowed in Ind$(\mathcal{G}, i, s)$. Let k be the number of cooperations in periods from 1 to $t-1$. Then,
$$\tau_i^{(t)}(p_c) = \frac{k+1}{t+1};$$
- $\tau_i^{(1)}(\overline{p_c}) = 1 - \tau_i^{(t)}(p_c)$;
- $v_i^{(t)}(p_c)$, $v_i^{(t)}$ and Ind$(\mathcal{G}, i, t)$ are determined analogously to Period 1.

Given this, we can now make the following definition.

Definition 1. The *iterated cooperative equilibrium* (ICE) of game $\mathcal{G}$ in period t is a strategy profile σ where strategy σ_i for each player $i \in N$ corresponds to the strategy he plays in the Nash equilibrium of the induced game Ind$(\mathcal{G}, i, t)$.

4 Traveler's Dilemma

In this section, we demonstrate the predictive power of cooperative equilibrium on the iterated Traveler's dilemma. We make use of the experimental data provided by Capra-Goeree-Gomez-Holt in [7] for the setting with $L = 80$ and $H = 200$, and compare ICE predictions with the *logit learning model* (LLM), proposed in [7] to explain these data.

There are two main differences between the LLM and ICE we would like to stress:

- First, as have been previously mentioned, ICE does not use any free parameter, while the LLM involves two free parameters, a *learning* parameter and a *error* parameter. In other words, ICE is a *predictive* model, and the LMM is *descriptive*.
- Second, the models are different conceptually. ICE applies the idea that people have an *attitude to cooperation*: they do not act a priori as single players, but rather forecast how the game would be played if they formed coalitions, and then play according to their best forecast. In contrast, the LLM assumes *selfish*, *individual* decisions, and explains deviations from Nash equilibrium in terms of mistakes.

We now proceed to compare between the ICE and the LLM predictions, based on the experimental data collected in [7]. In this experiment, groups of 9, 10 and 12 subjects played a 10 rounds Traveler's dilemma with low ($b \in \{5, 10\}$), intermediate ($b \in \{20, 25\}$) or high ($b \in \{50, 80\}$) bonus/penalty values. After each round, the subjects' claims were casually matched to determine their payoffs. In this paper, we exclude the case with $b = 10$ since it involved an odd number of participants (9 players), and so at each turn one player remained unmatched and his payoff was not determined; we therefore cannot compute the ICE in this case. Following [7], the LLM predictions are calculated using the values $\rho = 0.75$ and $\mu = 10.9$ for the learning/error parameters.

Recall that the Traveler's dilemma has a unique Nash equilibrium where each player chooses the minimal claim of $L = 80$, whichever is the value of bonus/penalty, b. The results in [7] show that in practice the players' behaviour is not independent of the value of b. Indeed, when the bonus/penalty value is low, the players tend to make very high claims, especially in the last rounds; this to some extent is supported by the logit learning model proposed in [7]. However, as can be seen from Table 1 and Figure 1, for $b = 5$ the ICE predicted values fall much closer to the average observed claims than the LLM predictions.

Period	Average observed claim	ICE prediction	LLM prediction
1	180.08	195.00	167.75
2	180.00	182.06	175.09
3	185.30	185.77	179.53
4	191.34	188.15	181.88
5	194.98	190.03	183.81
6	196.62	191.35	185.14
7	196.86	192.70	186.32
8	196.68	193.34	186.82
9	195.48	194.05	187.02
10	194.34	194.03	186.80

Table 1: Observed and predicted claims in Traveler's dilemma with low bonus/penalty of $b = 5$.

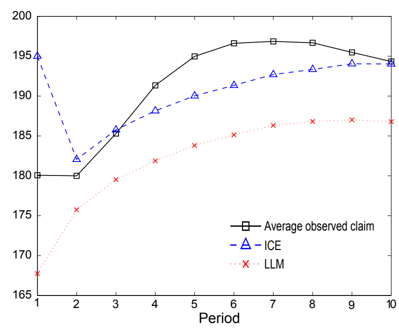


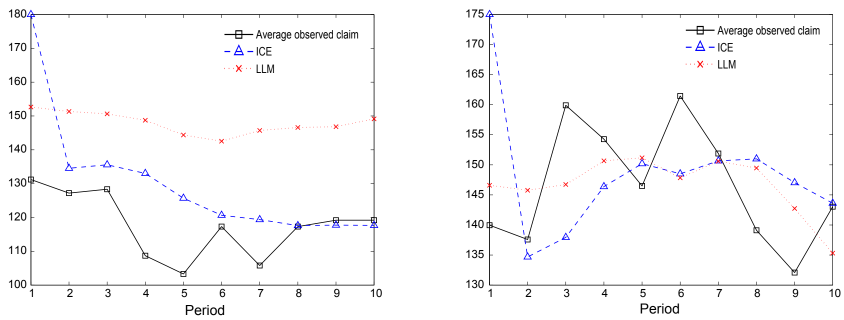
Fig. 1: ICE vs. LLM for $b = 5$. The solid line corresponds to the actual data. The ICE predictions are represented by the dashed line, and the LLM predicted values are depicted by the dotted line.

Table 2 and Figure 2 below present the data and predictions for the two cases with intermediate bonus/penalty values of $b = 20$ and $b = 25$. For $b = 20$, ICE again clearly

Period	b = 20			b = 25		
	Average observed claim	ICE prediction	LLM prediction	Average observed claim	ICE prediction	LLM prediction
1	131.20	180.00	152.64	139.96	175.00	146.60
2	127.20	134.53	151.32	137.59	134.68	145.77
3	128.35	135.57	150.63	159.90	137.94	146.73
4	108.70	133.02	148.74	154.27	146.38	150.66
5	103.30	125.69	144.38	146.49	150.19	151.17
6	117.30	120.66	142.55	161.44	148.51	147.84
7	105.80	119.37	145.71	151.88	150.65	150.60
8	117.30	117.60	146.60	139.12	150.99	149.47
9	119.20	117.73	146.82	132.09	147.04	142.74
10	119.20	117.66	149.14	143.04	143.62	135.32

Table 2: Observed and predicted claims in Traveler's dilemma with intermediate bonus/penalty.

outperforms the LLM, as shown in Figure 2a. For $b = 25$, the two models show similar performance: ICE is closer to the actual average claim in periods 2, 5, 6, 7 and 10, while the LLM performs better in periods 1, 3, 4, 8, and 9 (see Figure 2b). Note that the observed data in this case is very noisy, with no clear tendency towards higher or lower claims across the rounds of the experiment.



(a) $b = 20$ (b) $b = 25$

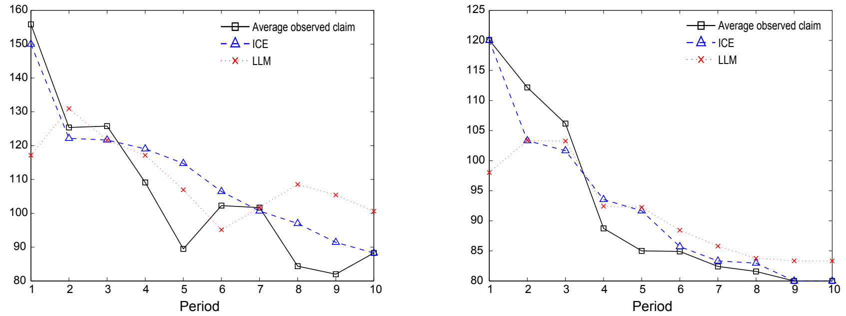
Fig. 2: ICE vs. LLM in Traveler's dilemma for $b \in \{20, 25\}$. Solid lines correspond to the actual data. The ICE predictions are represented by dashed lines, and the LLM predicted values are depicted by dotted lines.

As the bonus/penalty values get higher, the players reduce their claims, and actually converge to the Nash equilibrium solution in the last rounds of the experiment for high b (see Table 3 and Figure 3). While both ICE and the LLM capture this tendency, yet again, the ICE predictions appear to be closer to the experimental data.

In conclusion, the ICE model is much more accurate than the LLM in the prediction of population average behaviour in the Traveler's dilemma. Next we show that it can be successfully applied to other relevant social dilemmas, such as in fact the Prisoner's dilemma and the Public Goods game.

Period	$b = 50$			$b = 80$		
	Average observed claim	ICE prediction	LLM prediction	Average observed claim	ICE prediction	LLM prediction
1	155.86	150.00	117.17	120.07	120.00	98.04
2	125.37	122.15	130.95	112.18	103.33	103.38
3	125.77	121.66	121.63	106.16	101.66	103.26
4	109.13	119.06	117.15	88.75	93.55	92.43
5	89.47	114.75	106.95	85.00	91.66	92.23
6	102.26	106.46	95.13	84.91	85.71	88.44
7	101.68	100.67	101.74	82.41	83.33	85.79
8	84.38	96.99	108.54	81.58	82.96	83.77
9	82.00	91.43	105.42	80.00	80.00	83.35
10	88.27	88.27	100.63	80.00	80.00	83.34

Table 3: Observed and predicted claims in Traveler's dilemma with high bonus/penalty.



(a) $b = 50$ (b) $b = 80$

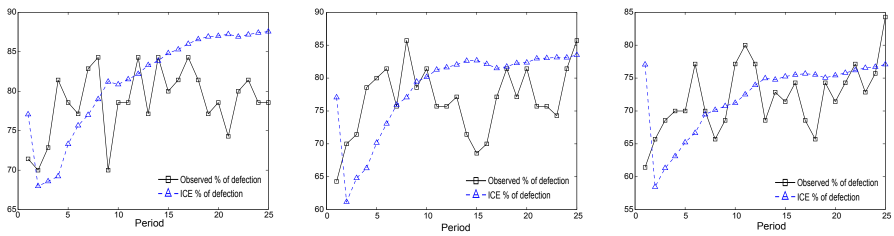
Fig. 3: ICE vs. LLM in Traveler's dilemma for $b \in \{50, 80\}$. Solid lines correspond to the actual data. The ICE predictions are represented by dashed lines, and the LLM predicted values are depicted by dotted lines.

5 Prisoner's Dilemma

In this section, we test our method on the iterated Prisoner's dilemma, using the experimental data provided by Yang-Yue-Yu in [32]. Although it is a dominant strategy for each player to defect, irrespective of payoffs or any other factors, human behaviours observed in experimental studies show considerable rates of cooperation, which appear to depend on game parameters or the environment in which it is played. The study in [32] is particularly focused on the way the players are matched to play together. This feature is crucial since different matching rules entail different histories for a player, and hence, different beliefs regarding his opponents' play. These, in turn, ultimately reflect on the player's strategic decisions. Therefore, it is of great importance to provide prediction methods that would achieve robust performance in different environments. As we show, ICE can successfully tackle this challenge.

The experiment involved 70 subjects that played a 25 rounds Prisoner's dilemma with payoffs parameters $T = 12$, $R = 8$, $D = 3$ and $S = 1$, under different matching schemes. Specifically, it included the following treatments: (i) the *random matching* (RM) where subjects were randomly paired in each period; (ii) the *one-period correlated matching* (OP) where subjects who have selected identical strategies in a given round are randomly paired with one another in the next period; and (iii) *weighted-history correlated matching* (WH) where, after every round, subjects are matched with a player who has been choosing *similar* strategies in the previous five periods. In more detail, the history is weighted using Fibonacci numbers as follows. Each subject starts with a sorting score $T(t) = 0$, for all $t \leq 1$. At each round t, his score is updated to $T(t) = 5a(t-1) + 3a(t-2) + 2a(t-3) + 1a(t-4) + 1a(t-5)$, where $a(s)$ is 0 if he plays defection in period s, and 1 otherwise. In each period, subjects are paired in the order of their current scores.

Table 4 and Figure 4 summarise the data collected in this experiment, along with the corresponding values of iterated cooperative equilibrium. As these results demonstrate, ICE accurately predicts the players' behaviour in Prisoner's dilemma, especially for cases with correlated matching (see Figures 4b and 4c). In the case where the players were matched randomly (Figure 4a), the ICE predictions in the last rounds of the experiment appear slightly more pessimistic than the actual data, which is implied by relatively high rates of defection observed in the intermediate rounds.



(a) RM treatment (b) OP treatment (c) WH treatment

Fig. 4: ICE in Prisoner's dilemma. Solid lines correspond to the actual data. The ICE predictions are represented by dashed lines.

6 Public Goods Game

In this section, we use ICE predictions to explain the experimental data on multi-round Public Goods game presented by Gunnthorsdottir-Houser-McCabe in [14].

The experiment consisted of three treatments with different constant marginal returns of $\alpha = 0.3$, $\alpha = 0.5$, and $\alpha = 0.75$. The first and the third treatment involved 36 subjects each, and 60 subjects participated in the second treatment. The subjects played a 10 rounds Public Goods game in groups of 4, to which they were randomly matched in each round.

The average observed contributions and the corresponding ICE predictions for the first treatment with $\alpha = 0.3$ are given in Table 5 and Figure 5.

Period	RM		OP		WH	
	Observed % of defection	ICE % of defection	Observed % of defection	ICE % of defection	Observed % of defection	ICE % of defection
1	71.43	77.07	64.29	77.06	61.42	77.06
2	70.00	67.97	70.00	61.18	65.71	58.46
3	72.86	68.59	71.43	64.80	68.57	61.31
4	81.43	69.22	78.57	66.29	70.00	63.11
5	78.57	73.31	80.00	70.16	70.00	65.22
6	77.14	75.66	81.43	73.07	77.14	66.65
7	82.86	77.00	75.71	75.94	70.00	69.47
8	84.29	78.99	85.71	77.06	65.71	70.15
9	70.00	81.21	78.57	79.46	68.57	70.74
10	78.57	80.87	81.43	80.17	77.14	71.22
11	78.57	81.51	75.71	81.28	80.00	72.51
12	84.29	82.18	75.71	81.62	77.14	73.93
13	77.14	83.31	77.14	82.04	68.57	74.97
14	84.29	83.85	71.43	82.62	72.85	74.76
15	80.00	84.82	68.57	82.68	71.42	75.24
16	81.43	85.29	70.00	82.13	74.28	75.50
17	84.29	86.00	77.14	81.49	68.57	75.67
18	81.43	86.58	81.43	81.74	65.71	75.52
19	77.14	86.92	77.14	82.28	74.28	75.05
20	78.57	87.00	81.43	82.36	71.42	75.41
21	74.29	87.22	75.71	82.98	74.28	75.79
22	80.00	86.91	75.71	83.04	77.14	76.18
23	81.43	87.18	74.29	83.14	72.85	76.55
24	78.57	87.41	81.43	83.12	75.71	76.70
25	78.57	87.56	85.71	83.52	84.28	77.08

Table 4: Observed and predicted behaviour in iterated Prisoner's dilemma.

Period	Average observed contribution	Standard deviation	ICE prediction
1	41.00	18.92	0.00
2	29.36	18.11	18.50
3	31.89	17.89	15.86
4	27.80	20.55	17.37
5	16.97	15.24	11.09
6	10.50	9.80	7.24
7	10.33	8.10	5.16
8	7.91	5.45	4.42
9	6.39	9.54	1.56
10	4.39	7.41	1.77

Table 5: Observed and predicted contributions in Public Goods with marginal return of $\alpha = 0.3$.

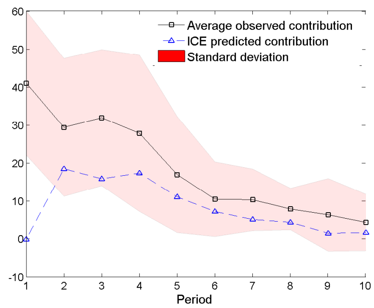
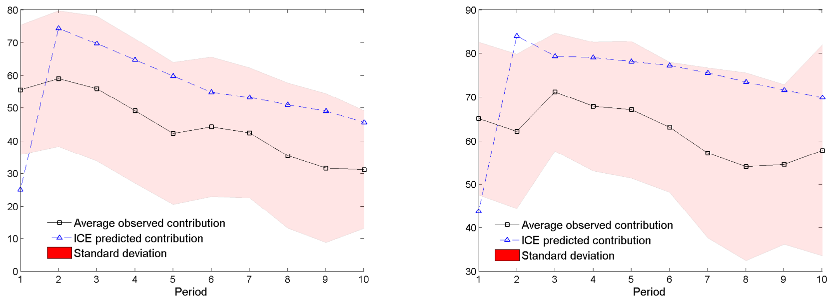


Fig. 5: ICE in Public Goods with $\alpha = 0.3$. The actual data are represented by the solid line. The shaded area shows the standard deviation. The dashed line corresponds to the ICE predictions.

The data are very heterogenous – note the high rates of standard deviation. This is reflected on the fact that the ICE's predictions in this setting seem less accurate than in previously considered domains. Notice, however, that in all game rounds (except of the very first one where the players beliefs are yet completely fictitious), the ICE values fall within the standard deviation interval and their error decreases as the number of periods increases. Similar performance is also showed in treatments with higher marginal returns, presented in Table 6 and Figure 6.

Period	$\alpha = 0.5$			$\alpha = 0.75$		
	Observed contribution	Standard deviation	ICE prediction	Observed contribution	Standard deviation	ICE prediction
1	55.48	19.77	25	65.00	17.47	43.75
2	58.88	20.69	74.31	62.08	17.67	84.01
3	55.83	22.09	69.57	71.11	13.53	79.29
4	49.03	22.06	64.62	67.78	14.76	78.97
5	42.16	21.67	59.67	67.02	15.64	78.07
6	44.16	21.29	54.75	63.02	14.88	77.18
7	42.33	19.84	53.10	57.16	19.47	75.48
8	35.38	22.17	50.86	54.02	21.46	73.40
9	31.60	22.72	48.94	54.52	18.26	71.56
10	31.10	17.93	45.53	57.78	24.12	69.81

Table 6: Observed and predicted contributions in Public Goods with constant marginal returns of $\alpha = 0.5, 0.75$.



(a) $\alpha = 0.5$ (b) $\alpha = 0.75$

Fig. 6: ICE in Public Goods with $\alpha = 0.5, 0.75$. The actual data are represented by solid lines. Shaded areas show the standard deviation. Dashed lines correspond to the ICE predictions.

7 Conclusions

In this paper, we introduced the Iterated Cooperative Equilibrium (ICE) which extends the approach of players' natural attitude to cooperation to games played in iterated fashion. In each round, the players forecast how the game would be played if they formed coalitions, and select their actions accordingly. The beliefs are initially defined through a sort of fictitious play, and then get updated at each step of the game, based on previous observations. We applied this concept to three fundamental social dilemmas: the Prisoner's dilemma, the Traveler's dilemma, and the Public Goods game. The novel and most important features of the ICE is that (1) it does not use any free parameters and so it is completely predictive; (2) it makes statistically precise predictions of population average behaviour in the aforementioned domains.

This work opens a number of research directions, from the extension of the ICE to include other relevant game models to theoretical questions concerning, for instance, convergence of the iterative procedure. Regarding the latter point, one can easily see that the ICE can converge only to one of Rabin's fairness equilibria [24]: in the Traveler's dilemma, ICE can converge either to $(200, 200)$ or to $(80, 80)$; in the Prisoner's dilemma, ICE can converge either to (C, C) or to (D, D). But can actual human behaviour converge to a different strategy? The intuition suggests that the answer to this question is negative and that, in general, human behaviour may not converge at all. Indeed, if a player in the Traveler's dilemma would know that his opponent plays an intermediate strategy, say $s = 175$, then he would either reduce his claim to achieve a larger gain (which would finally lead the players to the Nash equilibrium), or rather decide to increase it to show his opponent that they both can gain more. This way of reasoning generates an oscillation, that is perfectly coherent with and reflected by ICE.

Finally, it would also be interesting to try and combine ICE with evolutionary models, in order to tackle the "cold start" effect—i.e., inaccurate predictions in early iterations. Indeed, ICE typically starts showing high performance only after a few rounds of iteration, since players have to form statistically robust beliefs. Now, in [19] the authors use an evolutionary model to explain the experimental data in the first two periods of iterated Traveler's dilemma presented in [7]. So, it is plausible that a clever combination of ICE with evolutionary models can fit the experimental data even better.

References

1. Axelrod, R. The Evolution of Cooperation. Basic Books (1984)
2. Benisch, M., Davis, G.B., Sandholm, T. Algorithms for rationalizability and CURB sets. AAAI 2006, 598–604 (2006)
3. Boyd, R., Gintis, H., Bowles, S., Richerson, P.J. The evolution of altruistic punishment. Proc. Nat. Acad. Sci. USA 100, 3531–3535 (2003)
4. Camerer, C., Ho, T., Chong, J. A cognitive hierarchy model of games. Quaterly J. of Economics 119 (3), 861–898 (2004)
5. Capraro, V. A solution concept for games with altruism and cooperation. http://arxiv.org/pdf/1302.3988v2.pdf
6. Capraro, V. A Model of Human Cooperation in Social Dilemmas. Available at SSRN: http://ssrn.com/abstract=2246094

7. Capra, M., Goeree, J.K., Gomez, R., Holt, C.A. Anomalous Behavior in a Travelers Dilemma? American Economic Review 89 (3), 678–690 (1999)
8. Conitzer, V., Sandholm, T. A generalized strategy eliminability criterion and computational methods for applying it. AAAI 2005, 483488 (2005)
9. Costa-Gomes, M., Crawford, V., Broseta, B. Cognition and behavior in normal form games: An experimental study. Econometrica 69 (5), 1193–1235 (2001)
10. Fowler, J.H. Altruistic punishment and the origin of cooperation. Proc. Nat. Acad. Sci. USA 102, 7047–7049 (2005)
11. Fudenberg, D., Rand, D.G., Dreber, A. Slow to Anger and Fast to Forgive: Cooperation in an Uncertain World. American Economic Review 102, 720–749 (2012)
12. Goeree, J.K., Holt, C.A. Ten Little Treasures of Game Theory and Ten Intuitive Contradictions. American Economic Review 91, 1402–1422 (2001)
13. Goeree, J.K., Holt, C.A., Laury, S.K. Private Costs and Public Benefits: Unraveling the Effects of Altruism and Noisy Behavior. Journal of Public Economics, 83 (2), 255–276 (2002)
14. Gunnthorsdottir, A., Houser, D., McCabe, K. Disposition, History and Contributions in Public Goods Experiments. Journal of Behavior and Organization 62, 304–315 (2007)
15. Halpern, J.Y., Pass, R. Iterated Regret Minimization: a new solution concept. Games and Economic Behavior 74 (1), 184–207 (2012)
16. Hyafil, N., Boutilier, C. Regret minimizing equilibria and mechanisms for games with strict type uncertainty. UAI 2004, 268-277 (2004)
17. Isaac, M.R., Walker, J. Group size effects in public goods provision: The voluntary contribution mechanism. Quarterly Journal of Economics 103, 179–200 (1988)
18. Jamroga, W., Melissen, M. Doubtful Deviations and Farsighted Play. Progress in Artificial Intelligence, Lecture Notes in Computer Science, Volume 7026, 506–520 (2011)
19. Manapat, M.L., Rand, D.G., Pawlowitsch, C., Nowak, M.A. Stochastic evolutionary dynamics resolve the Travelers Dilemma. Journal of Theoretical Biology 303, 119–127 (2012)
20. McKelvey, R., Palfrey, T. Quantal response equilibria for normal form games. Games and Economic Behavior 10 (1), 6–38 (1995)
21. Nowak, M.A., Sigmund, K. Evolution of indirect reciprocity. Nature 437, 1291-1298 (2005)
22. Nowak, M.A. Five rules for the evolution of cooperation. Science 314, 1560–1563 (2006)
23. Panchanathan, K., Boyd, R. Indirect reciprocity can stabilize cooperation without the second-order free rider problem. Nature 433, 499–502 (2004)
24. Rabin, M. Incorporating Fairness into Game Theory and Economics. American Economic Review LXXXIII, 1281–1302 (1993)
25. Stahl, D., Wilson, P. Experimental evidence on players' models of other players. J. Economic Behavior and Organization 25 (3), 309–327 (1994)
26. Renou, L., Schlag, K.H. Minimax regret and strategic uncertainty, Journal of Economic Theory 145, 264–286 (2009)
27. Tennenholtz, M. Competitive safety analysis: Robust decision-making in multi-agent systems. JAIR 17, 363–378 (2002)
28. Teacy, WT Luke and Patel, Jigar and Jennings, Nicholas R and Luck, Michael. Travos: Trust and reputation in the context of inaccurate information sources. Autonomous Agents and Multi-Agent Systems 12, 183–198 (2006)
29. Trivers, R.L. The evolution of reciprocal altruism. Q. Rev. Biol. 46, 35–57 (1971)
30. Wellman, M., Greenwald, A., Stone, P. Autonomous Bidding Agents: Strategies and Lessons from the Trading Agent Competition. MIT Press (2007)
31. Wright, J.R., Leyton-Brown, K. Beyond Equilibrium: Predicting Human Behavior in Normal-Form Games. AAAI 2010, 901–907 (2010)
32. Yang, C.L., Yue, C.D.J., Yu, I.T. The rise of cooperation in correlated matching prisoners dilemma: An experiment. Experimental Economics 10, 3–20 (2007)