ACTION-ASSORTATIVE MATCHING AND COOPERATION IN PRISONER'S DILEMMA*

Mürüvvet Büyükboyacı† Mehmet Y. Gürdal‡

March 28, 2021

Abstract

We experimentally study the effect of action-assortative matching and payoffs from mutual cooperation in the Prisoner's Dilemma game. Under action-assortative matching, cooperators are more likely to be matched with cooperators whereas defectors are more likely to be matched with defectors. Subjects first played a baseline Prisoners Dilemma game and then a modified game with three possible levels of assortativity, termed as low, medium and high. We have three main behavioral observations from our experiments: 1) We observe both cooperators and defectors under all three levels of assortativity. 2) Cooperation rates increase with the payoff from mutual cooperation. 3) Cooperation rates increase with the level of assortativity. Furthermore, although we observe a decay in cooperation over time in the baseline and the low assortative treatments, it is possible to sustain cooperation in the medium and the high assortative treatments. Estimations based on the reinforcement learning model suggests differences across treatments in terms of the initial attractions and the effect of recent payoffs. Finally, we demonstrate that our behavioral findings could be explained with a model based on Quantal Response Equilibrium or a model with subjective interpretation of assortativity.

Keywords: prisoner's dilemma, cooperation, assortative matching.
JEL Classification Numbers: C70, C72, C73, C91, D02.

*Büyükboyacı thanks GAP-403-2018-2806, Gürdal thanks Boğaziçi University (BAP 13740) and Austrian Academy of Sciences (JESH 2020). We thank Nick Feltovich, Ben Greiner, Gergely Hajdu, Doruk Iris, Melis Kartal, Rupert Sausgruber and participants at two seminars for helpful comments and suggestions.

†Department of Economics, METU; e-mail: muruvvet@metu.edu.tr
‡Department of Economics, Boğaziçi University; e-mail: mehmet.gurdal@boun.edu.tr

1 Introduction

Many interactions in the society (from personal relationships to workplace collaborations, from environmental conservation to political participation, international relations, and price competition in markets etc.) involve a conflict between society's aggregate welfare and the welfare of separate individuals. These situations are often modelled as social dilemma games where everybody benefits from collective cooperation, but each individual has an incentive to deviate. An extensive set of theoretical and experimental studies provide us remedies such as sanctioning and repeated interaction to facilitate cooperation in these environments. A relatively less studied alternative remedy is non-random encounters among individuals, also termed as assortative matching (Bergstrom, 2003; Bergstrom, 2013; Bilancini et al., 2018; Yang and Yue, 2019). Under assortative matching, individuals with similar traits across some dimensions are more likely to interact with each other compared to random matching.

In this paper, we employ an assortative matching procedure where choosing an action increases the likelihood of a player for being matched with another player choosing the same action.[1] Our objective is to study experimentally the joint role of assortative matching and cooperation payoffs(c) on the cooperation rates in Prisoner Dilemma (PD hereafter) games depicted in Figure 1. In PD games, the strategy Defect (D) strictly dominates the strategy Cooperate (C) for both players, that is, $b > c$ and $d > a$. However, the Nash equilibrium is not efficient, since $c > d$. Consequently, under random matching, the expected behavior is universal defection. In our experiments, subjects play the standard PD game (baseline treatments) as well as a modified PD game (assortative treatment) where players are matched based on their choices of action. This process is termed as action-assortativity as in Nax et al. (2014), which is based on the theoretical framework provided in Bergstrom (2003). In our experiments, players first choose their action and then assigned either to an *assortative pool*, involving only members from the same type, or to a *common pool*, involving members from all types. Next, players in the same pool are matched in pairs and the payoffs are realized.[2] Thus, subjects who choose C (D) are more likely to be matched with subjects who also choose C (D).

Figure 1: Prisoner's Dilemma

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	(c, c)	(a, b)
	Defect	(b, a)	(d, d)

When the probability of being assigned to the assortative and common pools are F and $1 - F$, respectively, C strictly dominates D when F, c and the overall fraction of C choices are sufficiently high.

[1] Similar matching mechanisms were also used Gunnthorsdottir et al. (2010) and Nax et al. (2014). In either of the games players first commit to a contribution level in a public good game and are then matched based on their decisions.

[2] If they are in the assortative pool, only possible pairs are (C, C) or (D, D), if they are in the common pool any pairing is possible.

Motivated by this, we employ a 3x3 design: In one dimension we changed the level of assortativity, F, and in the other dimension we changed the payoff from mutual cooperation, c. Treatment variables are explained in more detail in the experimental design part.[3] We refer to the values of F as low, medium or high assortativity and the values of c as low, medium or high cooperation payoffs. In the high and the medium assortativity treatments, for all cooperation payoffs (low, medium, high), the expected payoff from playing C is higher than the expected payoff from playing D, whereas in the low assortativity and the baseline treatments, the reverse is true.

We have three main findings from our experiments. First, although we expect full cooperation (defection) under high and medium (low assortativity and baseline) assortativity, we observe a substantial fraction of defectors (cooperators). Second, cooperation rates weakly increase with the payoff from mutual cooperation, c. This is reminiscent of the result reported in Charness et al. (2016) where they compare the effect of cooperation payoffs on the cooperation rates in one-shot PD games. Third, cooperation rate is significantly and positively affected by the level assortativity, F. We show that our behavioral observations can be organized through two different models: Quantal Response Equilibrium (McKelvey and Palfrey, 1995, 1998) and an assortative matching model in which individuals interpret assortativity subjectively, similar to the magical thinking model by Daley & Sadowski (2017).

When we look at the dynammic cooperation behavior, we observe a decay in cooperation rates over time in the baseline and low assortativity treatment, similar to earlier findings from related social dilemma games (Ledyard, 1995). However, no such decay is observed in medium and high assortativity treatments. Differences in cooperation rates over time across treatments suggests that assortativity and cooperation payoffs may lead to differences in subjects' learning process. To understand this behavior better, we test this hypothesis by estimating the reinforcement learning model (Roth and Erev, 1998). The results reveal that in the low assortative treatments, subjects start game (initial attraction) by choosing defect and since all past payoffs are remembered equally (i.e the recency parameter ϕ is close to 1) they do not choose cooperate in the later periods. However, in the medium and high assortative treatments, subjects start the game by choosing cooperate and since ϕ value is lower in these treatments, relatively recent payoffs become more important, and hence it is possible to sustain cooperation.

The rest of the paper is organized as follows. In Section 2, we overview the results from two related strands of literature. In Section 3, we lay out the theoretical implications of action-assortative matching in PD games. In Section 4, we describe our experimental design. In Section 5 we present our results, in Section 6 we describe two alternative theoretical models that can explain our data and we conclude in Section 7.

2 Related Literature

Our study is related to two strands from previous literature: The first strand focuses on the effect of the relative values of payoff parameters in PD games on the cooperation rates. The second strand involves

[3] In all sessions we used fixed values for a, b, and d which were $5, 28$, and 9 respectively. c and F are changed across sessions: $16, 20$, and 24 are used for c and $9/11, 2/3$ and $3/17$ are used for F.

theoretical and experimental studies of assortative matching.

The studies in the first strand generally define an index based on payoffs in the game and show how people's cooperation rates change depending on this index. Rapoport (1967) defines an Index of Cooperation, equal to $(c-d)/(b-a)$ in our representation, as gains from cooperation normalized by the difference between highest and lowest payoffs in the game. He finds that cooperation in one-shot PD games increases with this index. Mengel (2018) defines *Risk* as the loss due to unilaterally cooperating against a defector, *Efficiency* as how much can be gained by mutual cooperation as opposed to mutual defection, and *Temptation* as the gain due to unilaterally defecting against a cooperator. In our representation, Risk equals to $(d-a)/d$, Efficiency equals to $(c-d)/c$ and Temptation equals to $(b-c)/b$. For the case of one-shot PD games, she finds a significant effect of Risk and for the case of repeated games, she finds a significant effect of Temptation on cooperation rates. Schmidt et al. (2001) compare the impact of payoff parameters in six different experimental games using two 'stranger' designs with exogenous and endogenous matching (where players can choose their match based on the history of play) and one 'partner' setting with endogenous matching. They focus on three indices, called 'greed', 'fear' and 'cooperator's gain' similar to temptation, risk and efficiency indices, respectively in Mengel (2018). Nevertheless, the variation in their indices is low. They find that cooperation rates correlate with all three indices, but do not find systematic differences across their different designs. Capraro et al. (2014) ask participants in their experiment to choose how much of an endowment to spend on helping the other player, with every c units spent resulting in the other person gaining b units, instead of asking participants to choose between C or D as typically done for PD games. They also vary the b/c ratio, and ask how the distribution of cooperation levels changes as a result. They find that the 'benefit to cost ratio' increases cooperation rates. There are other studies using public good or trust games with similar notions. Dawes and Thaler (1988) try to disentangle 'greed' and 'fear' in 7-player public good games. Snijders and Keren (1999) study the importance of risk and temptation by varying payoff parameters in (one-shot) trust games. They find that potential losses for a trustor are important in determining behaviour, which seems to indicate that risk might play an important role in this game. In this strand of the literature, the closest study to ours belongs to Charness et al. (2016). In their study, with the values in Figure 1, they keep parameters a, b, and d fixed and change c across treatments. They define payoffs from joint cooperation as social surplus, and find that an increase in social surplus yields higher cooperation rates in a one-shot PD game. We replicate Charness et al. (2016)'s study for different c values and whether cooperation due to higher social surplus can be sustainable over time, in one dimension of our treatments.

The second strand involves theoretical and experimental studies of assortative matching. Bergstrom (2003) provides a theoretical outline of assortative matching and discusses its implications for cooperation and evolutionary dynamics under these type of encounters. He also defines *index of assortativitiy*, which is one of our treatment variables in the current paper. With the same idea, Bilancini et al. (2018) theoretically analyzed the role of assortative matching on cooperation rates with two different cultural groups in the society. They find that when cultural intolerance is sufficiently strong, homophily emerges: if all agents from one cultural group cooperate, while all agents from the other cultural group defect, then interactions among agents within the same cultural group are more frequent. In our study, we have three different assortativity levels, i.e., the likelihood that cooperators (defectors) are matched with cooperators (defectors) changed from treatment to treatment. These treatments can be thought as cultural groups with different cooperative incentives. Although we did not make different groups with different assortativity levels interact among each other, we observe different cooperation rates within the same group depending on the assortativity level.[4] Nax & Rigos (2016) focus on several social dilemma games and allow assortativity to evolve endogeneously (i.e., increase or decrease) through voting among players. They demonstrate the stability of two possible outcomes: full assortativity or no assortativity. All studies mentioned up to now were of theoretical nature and to our knowledge, there are few experimental studies related to assortative matching. Yang et al. (2007) is a relatively early experimental study of assortative matching in PD games. Across different treatments, they change the length of previous rounds of play used to match players (one round vs. five rounds). They find a slight positive effect on cooperation for the one-round case and a more pronounced effect for the case of five rounds. In a more recent and related study, Yang & Yue (2019) introduce prosocial dummy players in the assortative matching treatments (again assortativity is achieved based on history of play in the last 5 rounds) and document a positive effect of these on cooperation rates. The current paper differs from these experimental studies in the sense that we achieve assortativity not using the history of play but through an algorithm based on action choices and a predetermined index of assortativity which will be described in more detail in Section 4.

3 Theoretical Implications of Action-Assortative Matching

In this section we describe the two-pool assortative matching process (Cavalli & Sforza, 1981; Bergstrom, 2003) and discuss its payoff implications for the Prisoner's Dilemma games. Take a population involving two types of players. A two pool assortative matching process first assigns members from each type either to an *assortative pool*, involving only members from the same type, or to a *common pool*, involving members from both types. Next, players are matched in pairs with some other player from the same pool that they are assigned to. When the probabilities of being assigned to the assortative pool and the common pool are constant and equal to F and $1 - F$, respectively, we obtain a two-pool assortative matching process with uniform assortativity F. The two-pool assortative matching process generates a constant *index of assortativity*, which is defined as the difference between the probability that a cooperator meets a cooperator, $P(C|C)$, minus the probability that a defector meets a cooperator, $P(C|D)$.[5] That is, index of assortativity $\alpha = P(C|C) - P(C|D)$. Given that the fraction of cooperators in the population is p, we have $P(C|C) = F + (1 - F)p$ and $P(C|D) = (1 - F)p$, and $\alpha = F$. Therefore in the process described above, F is also equal to the index of assortativity, α. In addition, the expected payoffs from cooperation and defection are as follows:

[4]The scholars studying evolutionary dynamics also focused on assortative matching (Hamilton and Taborsky, 2005a, Hamilton and Taborsky, 2005b; Bergstrom, 2003).

[5]Due to symmetry, the same index can also be defined as the probability that a defector meets a defector, $P(D|D)$, minus the probability that a cooperator meets a defector, $P(D|C)$.

$$E(C) = P(C|C)c + P(D|C)a \quad (1)$$
$$= (F + (1 - F)p)c + (1 - F)(1 - p)a \quad (2)$$
$$= (c - a)(F + p - Fp) + a \quad (3)$$

and

$$E(D) = P(D|D)d + P(C|D)b \quad (4)$$
$$= (F + (1 - F)(1 - p))d + (1 - F)pb \quad (5)$$
$$= (b - d)(p - Fp) + d \quad (6)$$

Define $\delta(p) = E(C) - E(D)$. Classical game theory predicts that, as long as $\delta(p) < 0$, rational agents should choose D. Nevertheless, we do not observe this strict behavior in many PD games: the expected payoff from defection is higher than the expected payoff from cooperation in standard PD games, but many subjects in economics experiments start by choosing cooperation, C and as they gain experience in the game individuals switch to defect, D (Ledyard, 1995; Andreoni and Miller, 1993; Bohnet and Kuebler, 2005; Dal Bo, 2005; Grimm and Mengel, 2009). Further, Charness et al. (2016) showed that cooperation rates in a one-shot PD games increase with c, even though it is a dominant strategy to choose D for all values of $c < b$. One possible interpretation of their results can be that as the expected payoff difference between defection and cooperation decreases[6], salience of choosing cooperation could be enhanced. This is the main idea behind models such as Quantal Response Equilibrium models as well: Players' error rates drop by the expected payoff difference among their strategies, i.e., the higher incentive a person has to choose the right strategy, the more likely she will choose that strategy. Consider our experiment, since $d\delta(p)/dc = F + p(1 - F) > 0$, i.e., the difference between the expected payoffs from cooperation and defection increases as c increases, players would be more likely to cooperate with higher c. A similar explanation is also valid for F: $d\delta(p)/dF = (c-a)(1-p)+(b-d)p > 0$, since $(c-a) > 0$ and $(b-d) > 0$, meaning the difference between the expected payoffs from cooperation and defection increases with the index of assortativity, F. In order to check whether the cooperation rates are affected from the extent of the difference in expected payoffs, we modified the values of both F and c in our treatments.

We close this section by laying out the implications of the matching procedure described above for the population frequencies of particular pairs of players. Let $\pi_{s_i s_j}$ denote the probability that a randomly chosen pair has one player choosing s_i and another player choosing s_j. The probability that a randomly chosen player is a cooperator ($s_i = C$) who is matched to defector ($s_j = D$), is given by $p(1 - F)(1 - p)$. Similarly, the probability that a randomly chosen player is a defector ($s_i = D$) who is matched to defector ($s_j = D$), is given by $(1 - p)(1 - F)p$. This and similar calculations yield:

[6]While having higher expected payoff for defection than cooperation.

$$\pi_{CD} = 2p(1-p)(1-F) \tag{7}$$

$$\pi_{CC} = p(F + (1-F)p) \tag{8}$$

$$\pi_{DD} = (1-p)(F + (1-F)(1-p)) \tag{9}$$

Since $d\pi_{CD}/dF = -2p(1-p) < 0$, the total proportion of assortative matchings (the sum of π_{CC} and π_{DD}) will increase with the index of assortativity, F.

4 Experimental Design and Hypotheses

Each session started with a baseline treatment and this was followed by an assortative matching treatment, both of which lasted for 15 periods. In each period of the baseline treatment subjects were matched in pairs, and played the PD game given in Figure 1. A stranger matching protocol was used in all treatments. In the assortative matching treatments, subjects first chose their strategies, and then with probability F, they were matched with someone from an assortative pool, which included only subjects choosing the same strategy as themselves, and with probability $1-F$, they were matched with someone from the common pool which included both types of subjects. Subjects' payoffs from the experiment is determined according to the sum of their payoffs from one randomly picked period among the first 15 periods (baseline), one randomly picked period among the second 15 periods (assortative treatment), and the participation fee.

In our experiments, we changed c (the payoff from mutual cooperation) and F (the probability of being matched from the assortative pool) across treatments. We used 3 different values for these parameters, resulting in a 3x3 design. Other parameters of the game were the same for each session and were set as follows: $a = 5$, $b = 28$, $d = 9$. At a given session c was fixed and could be one among $\{16, 20, 24\}$ and was the same both for the baseline and the assortative matching treatments. F on the other hand was also fixed for a given session and chosen among $\{9/11, 2/3, 3/17\}$. We refer to these values as *high assortativity*, *medium assortativity* and *low assortativity*.

The realization of these probabilities were implemented as follows: Subjects first chose a strategy, and then those who chose cooperate were assigned a random score between $[x, 100]$, and those who chose defect were assigned a score between $[0, 100-x]$. Afterwards, subjects were ranked according to these scores, with ties broken arbitrarily. Next, subjects ranking 1st and 2nd were matched to each other, those ranking 3rd and 4th were matched to each other, and so on. When $x \in (0, 50)$, this implies that $F = x/(100-x)$ and $1-F = (100-2x)/(100-x)$. For each session, x value was fixed and could be either 45, 40, or 15, generating the F values $9/11$, $2/3$, or $3/17$, respectively. We use $G(c, F)$ to denote a PD game implemented with the cooperation payoff c and the uniform assortativity parameter F. In this notation, $F = 0$ for baseline treatments. In Table 1, we present the expected payoffs from cooperation and defection for risk neutral players among all treatments for different possible levels of the population frequency of cooperators, p. In particular, we calculated the expected payoffs when $p = 0.1$, $p = 0.5$, or $p = 0.9$. Using these numbers, the probability of particular game outcomes (a cooperator being matched with a defector, etc.), hence the expected payoffs can be calculated. We chose assortativity levels according

Table 1: Treatments and Expected Payoff Calculations for Different Cooperator Rates

Baseline	Assortative	Sessions	Subjects	p	$E(C)$	$E(D)$	$E(C)-E(D)$
$G(16,0)$	$G(16,3/17)$	2	32	0.1	7.85	10.56	-2.72
				0.5	11.47	16.82	-5.35
				0.9	15.09	23.08	-7.99
$G(20,0)$	$G(20,3/17)$	2	32	0.1	8.88	10.56	-1.68
				0.5	13.82	16.82	-3.00
				0.9	18.76	23.08	-4.32
$G(24,0)$	$G(24,3/17)$	2	32	0.1	9.92	10.56	-0.65
				0.5	16.18	16.82	-0.65
				0.9	22.44	23.08	-0.65
$G(16,0)$	$G(16,2/3)$	2	26	0.1	12.70	9.63	3.07
				0.5	14.17	12.17	2.00
				0.9	15.63	14.70	0.93
$G(20,0)$	$G(20,2/3)$	2	28	0.1	15.50	9.63	5.87
				0.5	17.50	12.17	5.33
				0.9	19.50	14.70	4.80
$G(24,0)$	$G(24,2/3)$	2	24	0.1	18.30	9.63	8.67
				0.5	20.83	12.17	8.67
				0.9	23.37	14.70	8.67
$G(16,0)$	$G(16,9/11)$	2	26	0.1	14.20	9.35	4.85
				0.5	15.00	10.73	4.27
				0.9	15.80	12.11	3.69
$G(20,0)$	$G(20,9/11)$	2	28	0.1	17.55	9.35	8.20
				0.5	18.64	10.73	7.91
				0.9	19.73	12.11	7.62
$G(24,0)$	$G(24,9/11)$	2	24	0.1	20.89	9.35	11.55
				0.5	22.27	10.73	11.55
				0.9	23.65	12.11	11.55

Notes: The table summarizes the PD games used in the baseline and the assortative treatments across different sessions, the number of subjects for these sessions, and the differences in expected payoffs from cooperation and defection for risk-neutral subjects under different levels of aggregate cooperation frequencies.

to the sign of the difference between the expected payoffs from cooperation and defection: in the high and the medium assortativity treatments, the expected payoff of cooperation is higher than the expected payoff from defection for all levels of c. In the low assortativity treatment, the expected payoff from defection is higher than the expected payoff from cooperation for levels of c. The levels of the payoff from joint cooperation are chosen similar to Charness et al. (2016), i.e., as in their case we have *high*, *medium*, and *low* cooperation payoffs. In both treatments, we are also interested in understanding whether cooperation rates change with the expected payoff difference between cooperation and defection.[7]

All experimental sessions were conducted using z-Tree (Fischbacher, 2007). The experiments were conducted in Boğaziçi University (in 2019, July) and Middle East Technical University (in 2020, February)[8]. We had 252 subjects in 18 sessions and each session had between 12 to 18 participants. Subjects read the instructions through their screens on their own. The experiment was conducted in Turkish and an English translation of the instructions in the low cooperation payoff-low assortativity treatment can be seen in the Appendix B. Sessions lasted for about 25 minutes and average earnings were 35.9 TL, which includes a 10 TL show-up fee.[9]

5 Results

We organize the results as follows. First, in section 5.1 we report our aggregate results. In particular, we report differences in frequency of cooperative plays and successful cooperation across treatments and how cooperation rates changes over time. In Section 5.2, we focus on the individual determinants of cooperation in the baseline and the assortative treatments using regression analysis. In Section 5.3, we analyze whether cooperation rates change differently over time across treatments due to differences in learning across treatments by using the reinforcement learning model.

5.1 Aggregate Results

Based on Bergstrom (2003)'s model, all individuals are expected to cooperate in the high and medium assortative treatments, whereas they are expected to defect in the baseline and low assortative treatments. In Table 2, we summarize the frequency of cooperation during different periods of our treatments. As can be seen from the table, cooperation rates are significantly greater than zero in the baseline (p-value < 0.01 for $c = 16$, $c = 20$ and $c = 24$) and the low assortativity treatments (p-value$= 0.04$ for $c = 16$, p-value$= 0.03$ for $c = 20$, p-value$= 0.05$ for $c = 24$) when cooperation frequencies are averaged over all periods. Likewise, cooperation frequencies are significantly smaller than 100% in the medium assortativity treatment for all cooperation payoff levels (p-value$= 0.02$ for $c = 16$, p-value< 0.01 for $c = 20$, p-value$= 0.08$ for $c = 24$) when cooperation rates are averaged over all periods. Cooperation rates are significantly smaller than 100% in the high assortativity treatment for all cooperation payoff levels except $c = 24$ (p-value$= 0.10$ for $c = 16$, p-value$= 0.05$ for $c = 20$, p-value$= 0.12$ for $c = 24$).[10] Further, we also check whether the frequency of subjects who always cooperate or the frequency of those who always defect change across treatments. For the baseline treatment, the frequencies of subjects who always defect were 32%, 35%, 22.5% for cooperation payoffs 16, 20, and 24, respectively. This points out that a considerable fraction of subjects always defect. In the low assortative treatment although the percentage of subjects

[7]This is why we have both high and medium assortativity levels and different cooperation payoff levels as treatment variables.

[8]The education in the university was face to face around the time.

[9]Hourly minimum wage at the time of the experiment was 12.26 TL.

[10]For significance test, one-sample t-test is used.

who always defect fell down, it is still substantially higher than 0, i.e., the corresponding frequencies were 12.5%, 16%, 12.5% for cooperation payoffs 16,

Table 2: Cooperation Frequencies

			Assortative		
Period	c	Baseline	Low	Medium	High
1	16	0.31	0.44	0.38	0.69
	20	0.47	0.38	0.50	0.61
	24	0.50	0.41	0.63	0.75
15	16	0.07	0.19	0.31	0.54
	20	0.06	0.13	0.54	0.57
	24	0.13	0.31	0.67	0.88
All	16	0.14	0.27	0.45	0.62
	20	0.16	0.18	0.55	0.64
	24	0.26	0.35	0.73	0.84

Notes: The table reports the cooperation frequencies in different treatments for periods 1, 15 and for all periods combined.

20, and 24, respectively. In neither case, we do not observe full defection as predicted. When we look at the percentage of full cooperators in the medium and the high assortative treatments we observe much lower percentages of full cooperators compared to full defectors in the baseline and the low assortativity treatments. In particular, the percentages of full cooperators in the medium and high assortative treatments were 4%, 4%, 0% for cooperation payoffs 16, 20, and 24, respectively. Hence, unlike the standard theoretical prediction, we observe neither full defection nor full cooperation in any of the treatments.

Observation 1: *Contrary to standard theoretical prediction, we do not observe full defection in the baseline and the low assortativity treatments or full cooperation in the medium and the high assortativity treatments.*

In Charness et al. (2016), it is shown that as the payoffs from mutual cooperation, c, increases relative to other payoff parameters, subjects choose to cooperate more often. To check whether this is the case in our setting, i.e., whether cooperation rates increase with the expected payoff difference due to increase in cooperation payoffs, c, we report cooperation rates across different cooperation payoffs. For the first period of the baseline treatment (the first column of Table 2) as cooperation payoffs change, we find that cooperation rates are significantly different when the cooperation payoffs 16 and 20 are compared ($p = 0.04$, $Z = -2.096$) and when cooperation payoffs 16 and 24 are compared ($p = 0.01$, $Z = -2.479$).[11] However-

[11]In the first period of the game (since the baseline treatments were run before the assortative treatments, we can do this only for the baseline treatments), one can take the independent observations as the individual, instead of the session since subjects

ever, this difference is insignificant when cooperation payoffs 20 and 24 are compared ($p = 0.66$ and $Z = -0.44$).[12] This finding is slightly different than the effect reported in Charness et al. (2016).

Observation 2: *Considering the first period of the baseline treatment, cooperation rate increases as the cooperation payoff increases from* 16 *to* 20 *or from* 16 *to* 24*, however, this effect is mitigated when the cooperation payoff increases from* 20 *to* 24*.*

According to the payoff-monotonicity assumption made by Weibull (1995) and Bergstrom (2003), the growth rate of the proportion of cooperators is positive (negative) if the expected payoff of cooperators is higher (lower) than the expected payoff of the defectors. According to our parameters, in our experiment we have $E(C) < E(D)$ in the baseline and the low assortative treatments, and $E(C) > E(D)$ in the medium and the high assortative treatments. So, we expect cooperation rates to decline throughout time during the baseline and the low assortative treatments, and to increase during the medium and the high assortative treatments. The behavior in baseline treatments exhibit a classical pattern, with substantial positive cooperation during the initial periods and a declining trend afterwards (see Figure 2 for related figures of average time trends). This trend drives the cooperation frequency around 14% in baseline as of 15th Period. In Table 2, we observe a clear drop in cooperation rates from Period 1 to Period 15 for the baseline and the low assortativity treatments across all cooperation payoff values, whereas cooperation rates either do not change significantly or increase from Period 1 to Period 15 during the medium and the high assortativity treatments. Figure 2 graphs the cooperation rates over 15 periods and reinforces the results of Table 2. For the baseline and the low assortativity treatments, we observe what is commonly observed in Prisoner's Dilemma games or other similar games representing social dilemma situations: a decay of cooperation rates over time. Figure 2 shows differences in cooperation rates depending on treatment which suggests a potential difference in learning paths across treatments. We analyze subjects' learning behavior in detail in Section 5.3 by using a reinforcement learning model.

Observation 3: *When the expected payoff from cooperation is smaller than that of defection, i.e., in the baseline and low assortative treatment, we observe a clear decline in the cooperation rates throughout time. When the expected payoff of cooperation is larger than the expected payoff of defection, i.e., in the medium and the high assortative treatments, cooperation rates stay stable by time and not increase as predicted in Weibull (1995) and Bergstrom (2003).*

The cooperation rates for the assortative treatments were previously given in Table 2, and a natural question here is how these cooperation frequencies translate into actual coordination among players which occurs when both players cooperate. We summarize the corresponding percentages of player pairs across different cooperation payoffs and assortativity levels in Table 3. High cooperation payoff and a high index

interact with each other just once yet. We check this behavior by using individual data from all periods in the baseline and assortative treatments, in the next section.

[12]We also report the effect of cooperation payoffs, c on cooperation rates by using individual data from all periods through regressions in the next period.

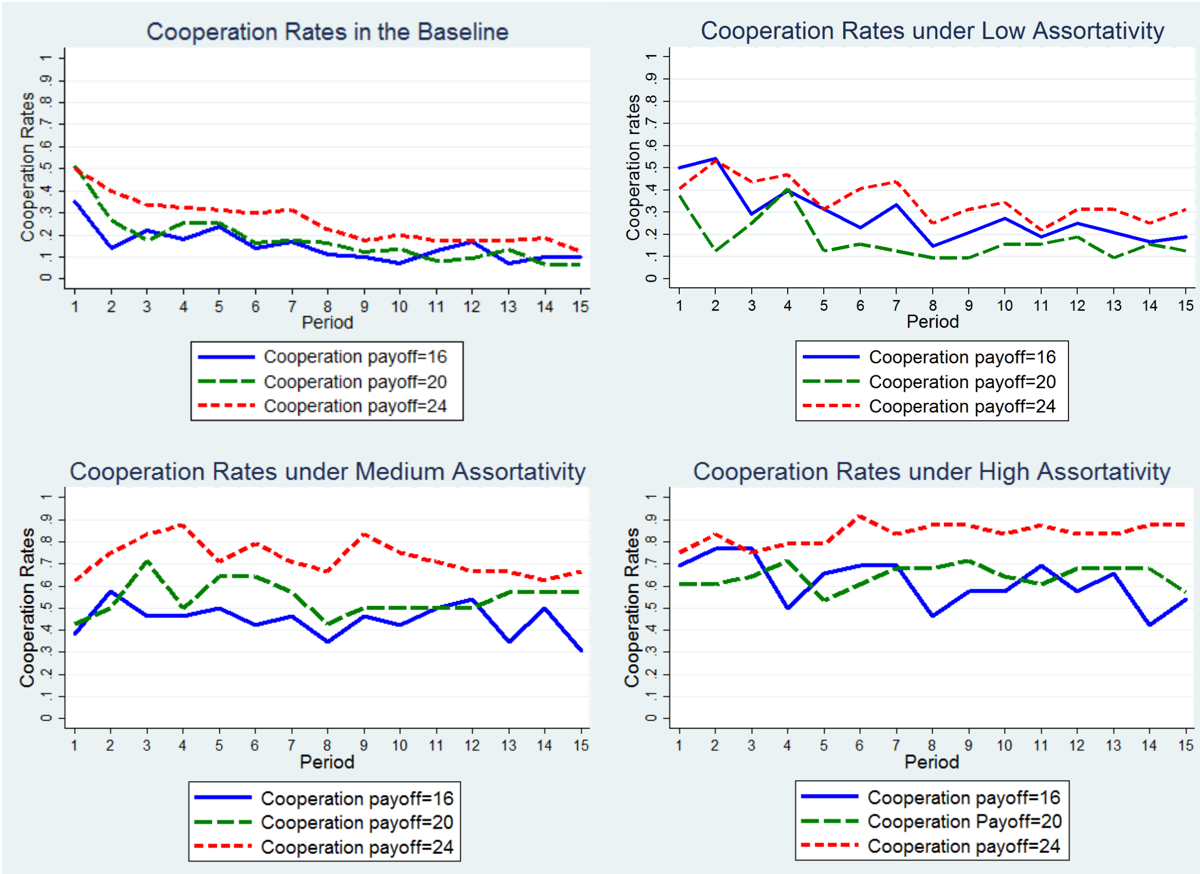

Figure 2: Cooperation Rates Across Time

of assortativity can drive the frequency of jointly cooperating pairs (π_{CC}) as close as to 80%, whereas this value could be around 10% when index of assortativity and cooperation payoffs are both low). In line with the theoretical implications of action assortative matching presented in section 3, the proportion of subjects in the nonassortative pool, (π_{CD}), declines as the index of assortativity, F, increases. That also means the proportion of subjects in the assortative pool, (π_{CC} or π_{DD}) increases.

Observation 4: *As the index of assortativity, F, increases from low to medium or to high, the proportion of subjects in the assortative pool increases, as predicted.*

5.2 Individual Results

We continue with regressions on the individual determinants of cooperation in baseline games. Our set of independent variables involve cooperation payoff (c), Period, experience during the previous period and other controls[13]. The results are summarized in Table 4. In all models, we find that $c = 24$ induces a higher

[13]The control variables we include are: indicator for male, age in years, number of siblings, number of economics courses taken, number of mathematics courses taken, Likert response to a general risk question ("How willing are you to take risks in general?"), and mean Likert response to a set of questions on norms of civic cooperation (Knack and Keefer, 1997), and number

Table 3: Percentage of Matching Pairs

			Assortative		
c	Outcome	Baseline	Low	Medium	High
16	π_{CC}	1.7	10.8	33.8	55.4
	π_{CD}	24.9	32.5	21.6	12.8
	π_{DD}	73.3	56.7	44.6	31.8
20	π_{CC}	3.6	4.2	41.9	57.1
	π_{CD}	25.3	26.6	26.2	14.3
	π_{DD}	71.1	69.2	31.9	28.6
24	π_{CC}	8.7	16.2	63.3	79.4
	π_{CD}	35	38.4	18.4	8.4
	π_{DD}	56.3	45.4	18.3	12.2

degree of cooperation compared to $c = 16$, but we fail to find a significant difference between $c = 20$ and $c = 16$. The pattern observed in Charness et al. (2016), is partially replicated here, with the game having the largest cooperation payoff generating substantially higher cooperation rate compared to the game with the smallest cooperation payoff. However, one should note that in the regressions we include all data (i.e. data coming from 15 periods) while Charness et al. (2016) use data coming from one shot Prisoner's dilemma. When we look at the cooperation rates in the first periods, although there is a significant difference between $c = 16$ and $c = 20$ and between $c = 16$ and $c = 24$, there is no significant difference between $c = 20$ and $c = 24$. Period variable has a significant and negative coefficient, indicating that experience leads to lower levels of cooperation during the baseline treatment. Models involving subject specific experience from the previous period as independent variables (Models 3, 4) indicate the following effects: Those who played cooperate in previous period are more likely to cooperate in the current period, whereas this effect is mitigated for those who cooperated but observed the opponent defected.[14] Having defected in the previous period and earning the temptation payoff (b) has no significant effect on the cooperation in current period. Among our control variables, we observe that males cooperate at a lower frequency, and the number of economics courses taken also has a negative effect on cooperation rates, as well.

We move on to regressions on the determinants of cooperation in assortative treatments. Our set of independent variables involve those we employed to analyze cooperation in the baseline treatment, and the controls for index of assortativity. The results are summarized in Table 5. In line with the results we found for the baseline treatment, we observe a positive and significant coefficient for cooperation payoff $c = 24$, but not for $c = 20$, indicating no significant difference with the latter and the case of $c = 16$ when controlling for index of assortativity. On the other hand, coefficients for both medium assortativity

of friends in the session). We include these variables as controls, rather than out of interest in their effects; hence we leave them out of our tables to save space.

[14] The difference between Model 3 and Model 4 is that we include control variables to the regression in Model 4.

Table 4: Cooperation in Baseline Treatments

	(1)	(2)	(3)	(4)
$c = 20$	0.0208	0.0520	-0.00137	0.0243
	(0.030)	(0.035)	(0.017)	(0.023)
$c = 24$	0.113***	0.143***	0.0617***	0.0869***
	(0.033)	(0.037)	(0.020)	(0.024)
Period	-0.0173***	-0.0173***	-0.00787***	-0.00818***
	(0.001)	(0.001)	(0.001)	(0.001)
Played C in $t-1$			0.289***	0.265***
			(0.022)	(0.025)
Played C and earned a in $t-1$			-0.115***	-0.111***
			(0.026)	(0.026)
Played D and earned b in $t-1$			0.0287	0.0319*
			(0.020)	(0.019)
N	3780	3780	3528	3528

Marginal effects from probit regressions reported. Dependent variable: 1=cooperation, 0:defection.
Standard errors are clustered at session level and reported in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

and high assortativity dummies turn out to be positive and significant. Further the coefficient of high assortativity dummy is higher than the coefficient of medium assortativity dummy, i.e., supporting that cooperation rates are affected not only from the sign of expected payoff difference between cooperation and defection but also how high this difference is.

We include subject's experience in previous period to Model 3 and Model 4 . According to these models, we find subject's experience in previous period also matters [15]: Those who cooperated in the previous Period are more likely to cooperate in the current period, and this effect is mitigated for subjects who cooperated but matched with a defecting opponent (thus earning the sucker payoff a). On the other hand, subjects who defected in the previous period and matched with a cooperator (thus earning the temptation payoff b), are less likely to cooperate at a given period. Note that we were unable to document such an effect for the baseline treatment. Unlike baseline treatments, we also don't find a gender effect, or an effect of the number economics classes taken. While the Period variable has a significant and negative coefficient, this effect solely comes from the low assortativity treatments. Running these regressions separately (results available upon request) for low, medium and high assortativity treatments, we find that the coefficient for Period is only significant for the case of the low assortativity treatment but not for the medium or the high assortativity treatments.

Observation 5: *As the index of assortativity F increases, cooperation rates increase.*

[15] The difference between Model 3 and Model 4 is that we include control variables to the regression in Model 4.

Observation 6: *When individual data from all periods are considered, cooperation rates increase significantly as cooperation payoff increases from* 16 *to* 24*, but the difference in cooperation rates is not significant when cooperation payoffs* 16 *to* 20 *are compared. This observations holds both for the baseline and the assortative treatments.*

Table 5: Cooperation in Assortative Treatments

	(1)	(2)	(3)	(4)
$c = 20$	0.00110	0.0293	0.00392	0.0227
	(0.042)	(0.041)	(0.024)	(0.023)
$c = 24$	0.182***	0.202***	0.111***	0.127***
	(0.046)	(0.046)	(0.028)	(0.028)
Medium Assortativity	0.288***	0.226***	0.150***	0.117***
	(0.032)	(0.036)	(0.021)	(0.025)
High Assortativity	0.406***	0.365***	0.196***	0.177***
	(0.037)	(0.040)	(0.029)	(0.034)
Period	-0.00749***	-0.00758***	-0.00534***	-0.00557***
	(0.002)	(0.002)	(0.001)	(0.002)
Played C in $t-1$			0.366***	0.342***
			(0.023)	(0.023)
Played C and earned a in $t-1$			-0.209***	-0.207***
			(0.028)	(0.029)
Played D and earned b in $t-1$			-0.0866***	-0.0882***
			(0.032)	(0.032)
N	3780	3780	3528	3528

Marginal effects from probit regressions reported. Dependent variable: 1=cooperation, 0:defection.

Standard errors are clustered at session level and reported in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.3 Reinforcement Learning Results

Figure 2 shows that how cooperation rates change over time across treatments. To understand whether these conditions stem from different learning paths across treatments, we estimate reinforcement learning (RL) model by Roth and Erev (1998) for each treatment.

Learning models proposed in the literature (reinforcement learning, belief learning, and experience weighted attraction) differ depending on how attractions are updated. In reinforcement learning, player i forms his set of attractions $A_i{}^j(t)$ according to the payoffs received in the previous periods. In particular the updating rule for each attraction variable is:

Table 6: Estimates for reinforcement learning in the assortative treatments.

	Low assortative			Medium assortative			High assortative		
	$c = 16$	$c = 20$	$c = 24$	$c = 16$	$c = 20$	$c = 24$	$c = 16$	$c = 20$	$c = 24$
ϕ	0.98*** (0.03)	0.98*** (0.03)	0.85*** (0.05)	0.73*** (0.08)	0.83*** (0.06)	0.65*** (0.05)	0.79*** (0.04)	0.62*** (0.05)	0.80*** (0.06)
λ	0.01*** (0.004)	0.01*** (0.004)	0.02*** (0.004)	0.03*** (0.005)	0.02*** (0.004)	0.04*** (0.005)	0.03*** (0.004)	0.05*** (0.005)	0.02*** (0.006)
${A_1}^0(0)$	-40.70*** (13.88)	-74.20*** (22.38)	-19.24* (10.54)	-7.41 (10.14)	7.21 (11.05)	26.25** (11.17)	27.55*** (10.78)	9.07 (6.73)	47.65*** (17.52)
${A_2}^0(0)$	-38.97*** (13.96)	-71.27*** (22.60)	-17.96* (10.09)	-7.21 (10.11)	7.41 (10.92)	25.95** (11.05)	27.49*** (10.68)	9.34 (6.72)	46.41*** (17.12)
#Obs	480	480	480	390	420	360	390	420	360
LL	-530.54	-424.40	-564.54	-496.52	-523.78	-351.23	-451.82	-421.13	-294.83

$$A_i^{\,j}(t) = \phi A_i^{\,j}(t-1) + I(s_i(t) = s_i^{\,j})\pi(s_i^{\,j}, s_{-i}(t))$$

for $i = 1, 2$ $j = 0, 1$ $t = 1, ..., 15$ where $I(.)$ is the indicator function, takes value 1 if $s_i(t) = s_i^{\,j}$ and 0 otherwise, $s_i(t)$ is the strategy chosen by player i in round t. Each player has two possible strategies in every round, $s_i^{\,0}$ and $s_i^{\,1}$ which are defect and cooperate respectively. The parameter ϕ represents the speed at which past payoffs are forgotten. If ϕ is 0, only the most recent payoff is remembered, if ϕ is 1 all payoffs have equal weight in the current decision.

Players' relevant experiences before the start of the game are represented by the prior values, $A_i^{\,j}(0)$, which are also named as initial attractions. For identification, ${A_1}^1(0)$ and ${A_2}^1(0)$ are normalized to 0 and other two initial attractions ${A_1}^0(0)$ and ${A_2}^0(0)$ are estimated.

The choice probabilities in any period are determined by the attractions in the previous period.

$$P_i^{\,j}(t) = \frac{e^{\lambda * A_i^{\,j}(t-1)}}{e^{\lambda * A_i^{\,1}(t-1)} + e^{\lambda * A_i^{\,0}(t-1)}}$$

for $i = 1, 2$, $j = 0, 1$, $t = 1, ..., 15$, and the parameter λ represents sensitivity to attractions.

We estimate reinforcement learning model for all assortative treatments via maximum likelihood (Moffatt, 2016)[16]. Estimated parameters are summarized in Table 6.

In all treatments ϕ is strongly significant. For the low assortative treatment (except high cooperative payoff), ϕ is close to 1, meaning all past payoffs have equal weight. In the medium and high assortative treatments ϕ gets away from 1, meaning past payoffs are forgotten fairly quickly. Similarly in all

[16] We updated the code in Moffatt's book for our experimental data.

treatments, the estimate of the sensitivity parameter, λ is positive and strongly significant, i.e., players are influenced by attractions.

In all treatments ${A_1}^0(0)$ and ${A_2}^0(0)$ are very similar since the game is symmetric. The estimates for ${A_1}^0(0)$ and ${A_2}^0(0)$ are negative and significant in the low assortative treatments, indicating a preference for players to start with defect. In the medium and high assortative treatments, these values are positive and often strongly significant (except for cooperation payoffss $c = 16$ and $c = 20$ in the medium assortativity treatment and for $c = 20$ in the high assortativity treatment). The positive values for ${A_1}^0(0)$ and ${A_2}^0(0)$ indicate a preference for choosing cooperate for players' initial choices.

Observation 7: *Depending on assortativity and payoff levels, there are differences in learning paths according to the reinforcement learning estimates.*

6 Explanations

Behavioral data points out three main observations.

1. Under high and medium index of assortativity, there are defectors, i.e., all subjects do not cooperate; under low assortativity there are cooperators, i.e., all subjects do not defect.
2. Given the index of assortativity, F, the frequency of cooperation, p is (weakly) increasing with the payoffs from joint cooperation, c.
3. Given c, p is increasing with the index of assortativity, F.

We will next show that these observations can be organized using two different theoretical models: The first one of these is based on Quantal Response Equilibrium (QRE) (McKelvey and Palfrey, 1995, 1998) and the second one is a model which is built on the magical thinking model by Daley & Sadowski (2017), and employs a subjective interpretation of assortativity .

6.1 QRE

In section 3, we have shown that the expected payoffs from cooperation and defection are $E(C) = (c - a)(F + p - Fp) + a$ and $E(D) = (b - d)(p - Fp) + d$, respectively. We observe that, increasing c or increasing F, while keeping other terms constant will result in the ratio of two expected payoffs, $E(C)/E(D)$ to increase. This implies that the probability of cooperation increases in the experimental data as $E(C)/E(D)$ increases.

This observation is in line with the main implications of Quantal Response Equilibrium (QRE), a widely used model of structural thinking in game theory. In this section, we lay out the implications of QRE for our setting. More formally, take the 2-person PD game between players $\{1, 2\}$, and let $A_i = \{C, D\}$ be the strategy space for player i and $\sigma_i(C)$ be the probability that player i plays C, and $\sigma_i(D)$=1-$\sigma_i(C)$ be the probability that player i plays D. The mixed strategy profile for player i is then $\sigma_i =$

$(\sigma_i(C), \sigma_i(D))$. Let $\pi_i(C)$ denote the expected payoff for player i from playing her pure strategy C, given σ_j, where $j = 3 - i$. Then, for any given $\lambda \geq 0$, the mixed strategy for player i at the corresponding Logit Equilibrium is given by

$$\sigma_i(C) = \frac{e^{\lambda * \pi_i(C)}}{e^{\lambda * \pi_i(C)} + e^{\lambda * \pi_i(D)}} \tag{10}$$

$$\sigma_i(D) = \frac{e^{\lambda * \pi_i(D)}}{e^{\lambda * \pi_i(C)} + e^{\lambda * \pi_i(D)}} \tag{11}$$

We solve for a symmetric Logit Equilibrium where $\sigma_i(C) = \sigma_j(C) = p$, $\pi_i(C) = \pi_j(C) = E(C)$, and $\pi_i(D) = \pi_j(D) = E(D)$, where $E(C)$ and $E(D)$ are defined as before. In Figure 3, we take $\lambda = 0.5$ and plot the cooperation probability p against the index of assortativity, F, for three different values of cooperation payoff, c. As it can be seen from the Figure 3, given the index of assortativity, the probability of cooperation is higher as the cooperation payoff increases, consistent with observation 2 and observation 6. In addition, the probability of cooperation increases with index of assortativity, and this holds for all levels of the cooperation payoff, c, consistent with observation 3.

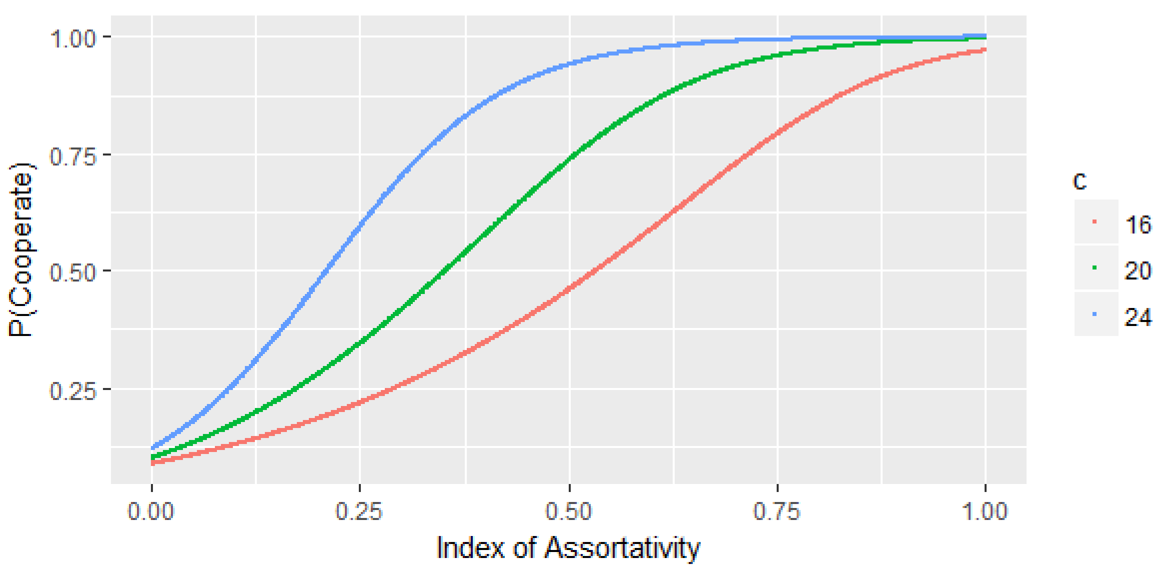

Figure 3: QRE across different cooperation payoffs and assortativity levels

6.2 A Model with Subjective Interpretation of Assortativity

In the magical thinking model of Daley & Sadowski (2017), players form beliefs as if their choices are going to affect their opponent's behavior. When applied to a PD game, this corresponds to a player choosing C to believe with a certain probability α that her opponent is going to do the same and otherwise she will be matched with someone randomly from the general pool. Similar to this notion, here we assume that each player i assigns a subjective probability to being assigned to the assortative pool and the non-assortative pool. For player i, these probabilities are denoted as α_i and $1 - \alpha_i$ respectively. We also assume that for each distinct assortative matching protocol with uniform assortativity F, a different distribution of α_i's is realized for the population and that the cumulative density of this distribution is given by $Q(\alpha) = \alpha^F$. Here $\alpha \in [0, 1]$ represents the subjective probability for being assigned to the assortative pool. For the game in Figure 1, player i's expected payoffs of choosing C and D are as follows:

$$V_i(C) = \alpha_i c + (1 - \alpha_i)(P_i a + (1 - P_i)c) \tag{12}$$

$$V_i(D) = \alpha_i d + (1 - \alpha_i)(P_i d + (1 - P_i)b) \tag{13}$$

where P_i represents the agent i's beliefs for being defected conditional on his belief for being in the nonassortative pool.

The strategy for player i specifies the action choice probabilities given her belief of being assigned to the assortative pool, α_i, and it is denoted by $\sigma_i = (\sigma_i(C|\alpha_i), \sigma_i(D|\alpha_i))$. For a given cumulative distribution function $Q(.)$, Daley & Sadowski (2017) define the equilibrium (σ, p) as a pair that satisfies the following conditions: 1) $\sigma_i = \sigma$ for all i (symmetry), 2) given a player's type and beliefs, it assigns positive probability to the strategy with the highest expected payoff, and 3) any player's belief conditional on his belief of being in the common (nonassortative) pool is consistent with his opponent's equilibrium strategy, that is, $P_i = P = \int_0^1 \sigma(D|\alpha) dQ(\alpha)$. Then, for any $P_i = P$, $V_i(C) - V_i(D)$ is strictly increasing in α_i and according to Daley and Sadowski (2017), it is possible to find a unique cutoff equilibrium, α^*, of the form $\sigma(D|\alpha)$=1 if $\alpha < \alpha^*$ and $\sigma(D|\alpha)$=0 if $\alpha > \alpha^*$ for some $\alpha^* \in [0, 1]$.[17] In equilibrium, the cutoff type α^*'s belief for being defected is $P_i = P = Q(\alpha^*) = \alpha^{*F}$ and he is indifferent between choosing C and D, i.e., $V_i(C) = V_i(D)$. Then we have

$$V_i(C) - V_i(D) = \alpha^*(c - d) + (1 - \alpha^*)(1 - \alpha^{*F})(c - b) + (1 - \alpha^*)\alpha^{*F}(a - d) = 0 \tag{14}$$

So with the existence of such an interior cutoff value, α^* it is possible to observe both defectors or cooperators for different F and c values which is consistent with the first observation we have. As stated at the beginning of this section, our data also shows that aggregate cooperation frequency increases with the payoff from mutual cooperation, c and with the index of assortativity, F. Now, we will show that the model we described above satisfies these conditions, hence can be used to explain our data.

1.Cooperation rates increase with c

Rearranging the equation 14, we have

$$V_i(C) - V_i(D) = \frac{\alpha^*}{(1 - \alpha^*)\alpha^{*F}}(c - d) + \frac{(1 - \alpha^{*F})}{\alpha^{*F}}(c - b) + a - d = 0 \tag{15}$$

In the above equation there are two terms with α^*, namely the coefficients in front of $(c - d)$ and $(c - b)$. We will now show that the first one of these increases with α^* whereas the second one decreases with α^*. For the first one of these terms, we have

[17] Proposition 1 shows the existence and Proposition 2 shows the uniqueness of the equilibrium in their paper. We chose $Q(\alpha) = \alpha^F$ as CDF for two reasons: 1) To satisfy type distribution condition for uniqueness that is described in Proposition 2 in Daley and Sadowski (2017). 2) To have an agent's belief for the population's α increases with F.

$$\frac{d}{d\alpha}\left(\frac{\alpha^*}{(1-\alpha^*)\alpha^{*F}}\right) = \frac{(1-\alpha^*)\alpha^{*F} - (F\alpha^{*F-1}(1-\alpha^*) - \alpha^{*F})\alpha^*}{\left((1-\alpha^*)\alpha^{*F}\right)^2} \tag{16}$$

$$= \frac{\alpha^{*F} - \alpha(1-\alpha^*)F\alpha^{*F-1}}{\left((1-\alpha^*)\alpha^{*F}\right)^2} \tag{17}$$

$$= \frac{\alpha^{*F}(1-(1-\alpha^*)F)}{\left((1-\alpha^*)\alpha^{*F}\right)^2} \tag{18}$$

$$\tag{19}$$

Since both $\alpha^* \leq$1 and $F \leq 1$, we know that the numerator is greater than 0. With the squared term at the denominator, we can then conclude that

$$\frac{d}{d\alpha}\left(\frac{\alpha^*}{(1-\alpha^*)\alpha^{*F}}\right) \geq 0 \tag{20}$$

Coming back to the coefficient in front of $(c-b)$, since α^{*F} is an increasing function of α^*, we have

$$\frac{d}{d\alpha}\left(\frac{1-\alpha^{*F}}{\alpha^{*F}}\right) \leq 0 \tag{21}$$

Given the last two inequalities above and the fact that $(c-d) > 0$ and $(c-b) < 0$, we conclude that when c increases, to keep $V_i(C) - V_i(D) = 0$, the cutoff value α^* has to go down. This means that for a given F, the frequency of defectors, α^{*F}, will decrease and the frequency of cooperators, $1-\alpha^{*F}$, will increase.

2. Cooperation rates increase with F

We start with rewriting equation 14 as follows:

$$V_i(C) - V_i(D) = \alpha^*(c-d) + (1-\alpha^*)((1-\alpha^{*F})(c-b) + \alpha^{*F}(a-d)) = 0 \tag{22}$$

Now suppose that F increases and the fraction of defectors, α^{*F}, also increases at the new equilibrium. Since F is less than 1 and α^* takes values between 0 and 1, we must have α^* increase, as well. We will now check what this implies for $V_i(C) - V_i(D)$ in equation 22 above. We can see that since $(c-d) > 0$ and given that α^* has increased, the term $\alpha^*(c-d)$ must have also increased. Next, since $(c-b) \leq (a-d)$ for our experiments and given that $(c-b) < 0$, $(a-d) < 0$, and α^{*F} has increased, we can conclude that $((1-\alpha^{*F})(c-b) + \alpha^{*F}(a-d))$ has also increased. Since this last term is negative and α^* has increased, it must be that $(1-\alpha^*)((1-\alpha^{*F})(c-b) + \alpha^{*F}(a-d))$ must have increased. Together with the fact that $\alpha^*(c-d)$ must have increased, we conclude that $V_i(C) - V_i(D) > 0$ at the new equilibrium, a contradiction. Therefore, our initial supposition is wrong and if F increases, the fraction of defectors, α^{*F}, must have decreased. Hence $1-\alpha^{*F}$, will increase as F increases.

7 Summary and Discussion

Exploring the factors that facilitate cooperation in social dilemma situtations is a major research programme involving various disciplines like anthropology, biology, game theory, political science, and psychology (Axelrod 1984; Boyd and Richerson, 1985; Bowles and Gintis, 2011; Hamilton, 1964) Previous studies have pointed out the role of assortative matching as one such factor (Eshel and Cavalli-Sforza, 1981). In this paper, we study a particular form of assortative matching, termed as action-assortative matching. We follow the definition of Nax et al. (2014), which is based on the theoretical framework provided in Bergstrom (2003), and we experimentially test the effect of action-assortative matching on the cooperation levels in Prisoner's Dilemma game. The matching algorithm we use assigns the players either to an assortative pool or to a common pool according to their action choices, and then matches players with another player from the same pool. We implemented a 3x3 design by changing the probability of being assigned to the assortative pool (i.e., low, medium, and high assortativity) and the payoff from mutual cooperation (low, medium, high) across treatments.

We observe that cooperation rates increase both with the level of assortativity and with the payoffs from mutual cooperation. In addition, it is possible to find both cooperators and defectors under all three levels of assortativity. Although we observe a decay in cooperation over time in the baseline and the low assortative treatments, it is possible to sustain cooperation under medium and high assortativity. Estimations of a reinforcement learning model supports these findings and demonstrate that the initial attractions and the relative weights assigned to recent payoffs differ across assortativity levels, leading to different temporal dynamics of action choices. We demonstrate that our behavioral findings could be explained with a model based on Quanta lResponse Equilibrium or a model with subjective interpretation of assortativity

Designing institutions (exogenously or endogenously determined) to increase cooperation is a prominent objective as societies try to tackle many problems like global security, environmental protection, infectious diseases, climate change, etc. Although the effect of institutions are studied extensively for public good games[18], the papers studying institutions to increase cooperation in prisoner's dilemma games are scarce.[19] Further, most of them are related to endogenous selection of the institutions.[20] Here, we contribute to this literature by analyzing the effect of action-assortative matching on cooperation rates in PD games.

[18]Punishment opportunities (Falk et al ., 2005; Fehr and Gächter, 2000; Masclet et al . 2003; Nikiforakis, 2008; Nikiforakis and Normann, 2008; or Yamagishi, 1986), communication (Bochet et al., 2006; Brosig et al., 2003; Cason and Khan, 1999; Isaac and Walker, 1988; or Ostrom et al., 1994), leadership (Arbak and Villeval, 2013; Güth et al., 2007; Potters et al., 2005; or Vesterlund, 2003), reputation opportunities (Milinski et al., 2006; or Sommerfeld et al., 2007), and ostracism (Cinyabuguma et al ., 2005; Güth et al., 2007; or Maier-Rigaud et al., 2005.) can be considered as the institutions that are used to increase cooperation in public good games.

[19]See Dannenberg and Gallier, 2019 for an extensive review.

[20]Dalbo et al., 2010, 2018; Barrett and Dannenberg, 2017; Grimm and Mengel, 2009, 2011.

References

Andreoni, J. and Miller, J. (1993). Rational cooperation in the finitely repeated prisoner's dilemma:experimental evidence, *Economic Journal*, vol. 103(418), pp. 570–85

Arbak, E. and M.-C. Villeval (2013), Voluntary Leadership: Motivation and Influence, *Social Choice and Welfare*, 40, 635–662.

Axelrod R (1984), The evolution of cooperation. Basic books, New York

Barrett, S., Dannenberg, A. (2017). Tipping versus cooperating to supply a public good. *Journal of the European Economic Association*, 15(4), 910–941.

Bergstrom, T.C. (2003). The algebra of assortative encounters and the evolution of cooperation *International Game Theory Review*, 5 (3) (2003), pp. 211-228.

Bergstrom, T.C. (2013). Measures of assortativity, *Biological Theory*, 8 (2) (2013), pp. 133-141

Bilancini E., Boncinelli L., Wu J. (2018). The interplay of cultural intolerance and action-assortativity for the emergence of cooperation and homophily, *European Economic Review*, 102, pp. 1-18

Bochet, O., T. Page and L. Putterman (2006), Communication and Punishment in Voluntary Contribution Experiments, *Journal of Economic Behavior and Organization*, 60, 11–26.

Bohnet, I. and Kuebler, D. (2005). Compensating the cooperators: is sorting in the prisoner's dilemma possible?, *Journal of Economic Behavior and Organization*, vol. 56(1), pp. 61–76.

Bowles S, Gintis H (2011), A cooperative species: human reciprocity and its evolution. *Princeton University Press*, Princeton

Boyd R, Richerson PJ (1985), Culture and the evolutionary process. *University of Chicago Press*, Chicago.

Brosig, J., J. Weimann and A. Ockenfels (2003), The Effect of Communication Media on Cooperation, *German Economic Review*, 4, 217–243.

Cason, T. N. and F. U. Khan (1999), A Laboratory Study of Voluntary Public Goods Provision with Imperfect Monitoring and Communication, *Journal of Development Economics*, 58, 533–552.

Charness, G., Rigotti, L., Rustichini, A. (2015), Social surplus and beliefs determine cooperation rates in the one-shot Prisoner's Dilemma, *Games and Economic Behavior*, 100, pp. 113-124.

Cinyabuguma, M., T. Page and L. Putterman (2005), Cooperation under the Threat of Expulsion in a Public Goods Experiment, *Journal of Public Economics* 89, 1421–1435

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

Dal Bó, E., Dal Bó, P., Eyster, E. (2018). The demand for bad policy when voters under appreciate equilibrium effects. *Review of Economic Studies*, 85(2), 964–998.

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

Daley, B., Sadowski, P. (2017). Magical thinking: A representation result. *Theoretical Economics*, 12(2), pp. 909-956.

Dannenberg, A., Gallier, C., (2019). The choice of institutions to solve cooperation problems: a survey of experimental research, *Experimental Economics* https://doi.org/10.1007/s10683-019-09629-8

Dawes, R. and Thaler, R. (1988). Anomalies: cooperation, *Journal of Economic Perspectives*, vol. 2(3), pp. 187–197.

Falk, A., E. Fehr and U. Fischbacher (2005), Driving Forces behind Informal Sanctions, Econometrica 73, 2017–2030.

Fehr, E. and S. Gächter (2000), Cooperation and Punishment in Public Goods Experiments, American Economic Review 90, 980–994.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments, *Experimental Economics*, 10(2), pp. 171–178.

Grimm, V. and Mengel, F. (2009). Cooperation in viscous populations–experimental evidence,*Games and Economic Behavior*, vol. 66(1), pp. 202–20.

Grimm, V., Mengel, F. (2011). Matching technology and the choice of punishment institutions in a prisoner's dilemma game. *Journal of Economic Behavior and Organization*, 78, 333–348.

Gunnthorsdöttir, A., Vragov, R., Seifert, S., McCabe, K.(2010). Near-efficient equilibria in contribution-based competitive grouping *Journal of Public Economics*, pp. 987-994

Güth, W., M. V. Levati, M. Sutter and E. van der Heijden (2007), Leading by Example with and without Exclusion Power in Voluntary Contribution Experiments, *Journal of Public Economics*, 91, 1023–1042.

Hamilton WD (1964), The genetical evolution of social behaviour. II. J Theor Biol 7(1):17–52.

Hamilton, I.M., Taborsky, M. (2005a). Contingent movement and cooperation evolve under generalized reciprocity, *Proc. R. Soc. B: Biol. Sci.*, 272 , pp. 2259-2267.

Hamilton, I.M., Taborsky, M. (2005b). Unrelated helpers will not fully compensate for costs imposed on breeders when they pay to stay, *Proc. R. Soc. B: Biol. Sci.*, 272 (2005), pp. 445-454.

Isaac, M. R. and J. M. Walker (1988), ‘Communication and Free-Riding Behavior: The Voluntary Contributions Mechanism’, Economic Inquiry 26, 585–608.

Knack, S. Keefer, P. (1997). Does social capital have an economic payoff? A cross-country investigation. Q. J. Econ. 112, pp. 1251–1288

Ledyard J.O. (1995). Public Goods: A Survey of Experimental Research Kagel J., Roth A. (Eds.), The Handbook of Experimental Economics, Princeton University Press, Princeton.

Maier-Rigaud, F. P., P. Martinsson and G. Staffiero (2005), Ostracism and the Provision of a Public Good – Experimental Evidence, *Journal of Economic Behavior and Organization* 73, 387–395.

Masclet, D., C. Noussair, S. Tucker and M.-C. Villeval (2003), Monetary and Nonmonetary Punishment in the Voluntary Contributions Mechanism, *American Economic Review*, 93, 366–380.

McKelvey, Richard D. and Thomas R. Palfrey (1995) Quantal Response Equilibria for Normal FormGames, *Games and Economic Behavior*, 10, pp. 6-38.

Mengel, F. (2018), Risk and Temptation: A Meta-study on Prisoner’s Dilemma Games, *The Economic Journal*, 128(616), pp. 3182–3209.

Milinski, M., D. Semmann and H.-J. Krambeck (2002), Reputation Helps Solve the ‘Tragedy of the Commons, *Nature*, 415, 424–426.

Moffatt, P. (2016), Experimetrics: Econometrics for Experimental Economics, *Palgrave Macmillan*.

Nax HH, Murphy RO, Helbing D. Stability and Welfare of ’Merit-Based’ Group-Matching Mechanisms in Voluntary Contribution Game. Available at SSRN 2404280. 2014.

Nax, H.H and A. Rigos (2016), Assortativity Evolving from Social Dilemmas,*Journal of Theoretical Biology*, 395, pp. 194-203.

Nikiforakis, N. (2008), Punishment and Counter-Punishment in Public Good Games: Can We Really Govern Ourselves?, *Journal of Public Economics*, 92, 91–112.

Nikiforakis, N. and H.-T. Normann (2008), ‘A Comparative Statics Analysis of Punishment in Public-Good Experiments’, Experimental Economics 11, 358–369.

Potters, J., M. Sefton and L. Vesterlund (2005), ‘After You-Endogenous Sequencing in Voluntary Contribution Games’, Journal of Public Economics 89, 1399–1419.

Rapoport, A. (1967), “A note on the” index of cooperation” for prisoner’s dilemma”, *Journal of Conflict Resolution*, 11(1), pp. 100–103.

Schmidt, D., Shupp, R., Walker, J., Ahn, T.K. and Ostrom, E. (2001). Dilemma games: game parameters and matching protocols, *Journal of Economic Behavior and Organization*, vol. 46(1), pp. 357–77.

Snijders, C. and Keren, G. (1999). Determinants of trust, in (I. E. D. V. Budescu and R. Zwick, eds.), *Games and Behavior: Essays in Honor of Amnon Rapoport*, pp. 355–83. Amsterdam: Lawrence Erlbaum Associates.

Sommerfeld, R. D., H.-J. Krambeck, D. Semmann and M. Milinski (2007), Gossip as an Alternative for Direct Observation in Games of Indirect Reciprocity, *Proceedings of the National Academy of Sciences*, 104, 17435–17440.

Vesterlund, L. (2003), The Informational Value of Sequential Fundraising, *Journal of Public Economics*, 87, 627–657.

Yamagishi, T. (1986), The Provision of a Sanctioning System as a Public Good, *Journal of Personality and Social Psychology Review*, 51, 110–116.

Yang, C. L., Yue, C. S. J., Yu, I. T. (2007). The rise of cooperation in correlated matching prisoners dilemma: An experiment, *Experimental Economics*, 10(1), pp. 3–20.

Yang, C. L., Yue, C. S. J. (2019). Cooperation in an Assortative Matching prisoners Dilemma experiment with pro-Social Dummies, *Scientific reports*, 9(1), pp. 1-8.

A Control Variables

Variable	Range	Definition
Age	$[18, 30]$	age of the subject
Male	$0/1$	1 if subject indicates gender as male.
Number of Siblings	$[0, 9]$	number of siblings of the subject
Number of Economics Courses	$[0, 4]$	number of economics courses taken by the subject, censored at 4.
Number of Mathematics Courses	$[0, 4]$	number of economics courses taken by the subject, censored at 4.
Risk Appetite	$[1, 9]$	response to general risk question: "How willing are you to take risks in general?" (0 lowest -10 highest)
Trust	$0/1$	response to general trust question: "Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?"
Civic 1	$[1, 5]$	claiming government benefits to which you are not entitled
Civic 2	$[1, 5]$	avoiding a fare on public transport
Civic 3	$[1, 5]$	cheating on taxes if you have the chance
Civic 4	$[1, 5]$	keeping money that you have found
Civic 5	$[1, 5]$	failing to report damage you have done accidentially to a parked vehicle
Friends	$[0, 14]$	number of people known in the session

Table 7: Ranges and definitions for control variables used in the regressions.

B English translation of experimental instructions

The instructions below are translated from those used in **low assortativity** and **low payoff** treatment. The instructions from the other treatments are similar and available upon request from the corresponding author.

General Information

Welcome and thank you for your participation.

This study aims to understand how people make decision in specific conditions. From now on, participants cannot talk among each other. If you violate this rule, we will have to terminate the experiment. If you have any questions, please raise your hand and ask your question.

The experiment will be over computers and you will convey all of your decisions about the experiment through the computer. All participants will earn some amount of money during the experiment. The money you earn might be different from the other participants' earnings. This amount is dependent on your decisions as well as the decisions of other participants. In addition to this earning you will earn money due to your participation.

There are two parts in the experiment.

OK

Now, we start explaining the first part.

There are 15 periods in this part of the experiment and you will play the game you see in the following table with another participant. Now we will explain the game you will play in detail. Please do not hesitate to ask if you have any question.

	If other player chooses A	If other player chooses B
If you choose A	Your earning **16** Other player's earning **16**	Your earnings **5** Other player's earning **28**
If you choose B	Your earning **28** Other player's earning **5**	Your earning **9** Other player's earning **9**

Both you and the other player will choose between A and B options. You will make your decisions without knowing the other person' s choice. Your earning in that period will depend on both your and the other player' s choice.
If both of you choose option A, both of you will earn 16 TL.
f you choose A and the other person choose B, then you will earn 5 TL, while the other person will earn 28 TL.
If both of you choose option B, both of you will earn TL 9 TL.If you choose B and the other person choose A, then you will earn28 TL, while the other person will earn 5 TL.

OK

Matchings:

Although it is possible to be matched with the same person, in general you will be matched with a different person from period to period. You will not learn the identity of your match in any of the periods. Similarly, your match will not learn you identity in any of the periods.

Earnings:

Your earnings from this part of the experiment will be determined according to your earnings in arandomly chosen period in 15 periods.

Total earnings from the experiment will be "Your Earnings in the first Part" + "Your Earnings in the second Part" + "Participation fee".

OK

Round

1 / 15

Remaining time [seconds]: 50

Payoffs of the game are given in the following Table.

	If other player chooses A	If other player chooses B
If you choose A	Your earnings **16** Other player's earning **16**	Your earnings **5** Other player's earning **28**
If you choose B	Your earning **28** Other player's earning **5**	Your earning **9** Other player's earning **9**

Which option do you choose? ○ A
○ B

OK

Round

1 / 15

Remaining time [seconds]: 56

Payoffs of the game are given in the following Table.

	If other player chooses A	If other player chooses B
If you choose A	Your earning **16** Other player's earning **16**	Your earning **5** Other player's earning **28**
If you choose B	Your earning **28** Other player's earning **5**	Your earning **9** Other player's earning **9**

Your choice: B

Other player's choice: B

You earned **9** TL in this period.

OK

Now, we start explaining the second part.

There are 15 periods in this part of the experiment and you will play the game you played in the firstpart with another participant and earn money according to the outcome of the game.

	If other player chooses A	If other player chooses B
If you choose A	Your earning **16** Other player's earning **16**	Your earning **5** Other player's earning **28**
If you choose B	Your earning **28** Other player's earning **5**	Your earning **9** Other player's earning **9**

But different than the previous part, matchings for the game will be realized after you make your choice. Hence, your matching will both depend on your choice and the luck.

We will explain the details of the matching in the next screen.

OK

Matchings:

In this part, computer will calculate a matching score for you as follows:

- If you choose option A, his score will be randomly drawn between 15 and 100.
- If you choose option B, this score will be randomly drawn between 0 and 85.

Then the computer will rank all participants of the experiment according to this score. If there are two participants with the same score, the computer will randomly rank one of them higher.

After the computer makes ranking of each participant according to the score:

- It will match two participant with the first and second-highest matching score together.
- It will match two participant with the third and fourth-highest matching score together.
- It will match two participant with the fifth and sixth-highest matching score together.

and it will complete whole matching procedure like this.

Earnings:

Your earnings from this part of the experiment will be determined according to your earnings in a randomly chosen period in 15 periods.

Total earnings from the experiment will be, "Your earnings from Part 1" + "Your earnings from Part 2" + "Participation fee".

OK

Round

1 / 15

Remaining time [seconds]: 56

Payoffs in the game are given in the following table:

	If other player chooses A	If other player chooses B
If you choose A	Your earning **16** Other player's earning**16**	Your earning **5** Other players earning **28**
If you choose B	Your earning **28** Other players earning **5**	Your earning **9** Other players earning **9**

Which action do you choose? ○ A
○ B

OK

Round

1 / 15

Remaining time [seconds]: 57

Payoffs in the game are given in the following table:

	If other player chooses A	If other player chooses B
If you choose A	Your earning **16** Other player's earning **16**	Your earning **5** Other player's earning **28**
If you choose B	Your earning **28** Other player's earning **5**	Your earning **9** Other player's earning **9**

Your matching score 42

Now the computer is first ranking you according to this score and then match you.

Your earnings will be determined according to your and your match's choices and you will see your earnings in the next screen.

OK