Volume 30, Issue 1

Assortative Matching, Information and Cooperation: An Experiment

Chun-Lei Yang
Academia Sinica

Jack Yue
National Chengchi University

Abstract

Yang et al.(2007) show that assortative matching mechanisms can induce a high level of cooperation to prevail in repeated prisoners' dilemma experiments, but societies may also head to an all-defection state rapidly. Here, we additionally provide information on income distribution and show that cooperation significantly increases overall. Moreover, the effect can be traced back to a higher cooperation rate for people in the lowest income classes.

1. Introduction

Assortative matching within the population, i.e. when individuals of similar behavioral traits have a higher-than-chance likelihood to be matched with one another, has been implicated as a general framework of dynamic interaction that induces proliferation of cooperation in prisoners' dilemma game.[1] Sober and Wilson (1998) discuss many biological models and examples that show that the "general principle of group/multilevel selection" can generate assortative matching effects to solve the cooperation problem within species.

In the economic literature, Frank (1988) argues that, true to the proverbial "you are what you do," frequent cooperation may induce non-imitable, hard-wired emotional expressions that signal trustworthiness as trading partner. Yamagishi et al. (2003), Frank et al. (1993), Brosig (2002), and Orbell and Dawes (1993) show experimental evidence for such signal recognition at work. Consequently, people of similar cooperative attitudes and similar action histories naturally end up being matched together to play the one-shot PD. Indirectly, acting is also like investing in the chance of encountering people of similar action histories,[2]

Note that signal recognition is hardly controllable experimentally. Thus, Yang et al. (2007) introduce matching mechanisms that get people into trading matches according to their past actions known only to the system operator, as if they already solved Frank's signal recognition process. Could cooperation prevail under those mechanisms? Data hint that the answer depends on group dynamics, which might yield great success for cooperation as well as its faster decline in societies. There is great hope for cooperation but it is not a sure thing. In the current study, we show that adding minimal income-distribution information can solve this problem and avoid societies heading into all-defection dead-ends without rescue.

2. Game, Matching Procedures, and Experiment Design

Table 1 shows the payoffs in NT (1 USD = ca. 32NT) for the row player in the basic prisoners' dilemma. C and D stand for the strategies of cooperation and defection respectively. The numbers are chosen so that the incentive to play defection is strong.

Table 1: Payoff Matrix

	C	D
C	8	1
D	12	3

Yang et al. (2007) introduced the *weighted-history assortative-matching* (WH) procedure where subjects are matched according to weighted history of previous actions. In round t, one's sorting score is defined as $T(t) = 5a(t-1) + 3a(t-2) + 2a(t-3) + 1a(t-4) + 1a(t-5)$, where $a(s) =$

---

[1] Bergstrom (2003) discusses abstract models on assortative matching. Bowles and Gintis (2002) survey further related works.
[2] See Robson (1990) and Amann and Yang (1998) for further analyses.

0 if his action in period s is defection, and = 1 otherwise. For completeness, $T(t) = a(t) = 0$ for all $t \leq 0$. In each round, all subjects are sorted after their current T-scores and matched in pairs from top down. Ties are broken randomly. An additional test ensures that subjects understand the matching rule correctly. Each subject only knows his own payoffs and the opponents' actions from past matches, but nothing about others' T-scores.

In this study, we introduce the novel *weighted-history with payoff information* (WHP) treatment, which modifies the WH treatment by providing additional information. More precisely, each round we calculate subjects' current 5-round average payoffs and sort them into 3, 4, 4, 3 person subgroups from bottom up. Each subject is informed of the averages in those subgroups, next to his own 5-round average, to make him aware of his gross payoff ranking in society.

We have three treatments, RM, WH, and WHP. Every treatment has 5 groups of 14 participants each, who were students at National Chengchi University in Taiwan. A typical session consists of 3 separate games, where the treatment game of 25 rounds is sandwiched by 5 rounds of random-matching, no-feedback PD games as illustrated in Table 2. Subjects were unaware of what game was to come next. Average payoff for participants was about 350 NT (about 11 US dollars), including a 50 NT show-up fee, in sessions of around 75 minutes.

Table 2: Treatment summary

Treatment	Game 1	Game 2	Game 3
1: RM	RM-5	RM-25	RM-5
2: WH	RM-5	WH-25	RM-5
3: WHP	RM-5	WHP-25	RM-5

3. Data Analysis

Game 1 is designed to control for sampling bias regarding the initial inclination to cooperation. A Kruskal-Wallis rank sum test reveals no such bias across all treatments, $p = 0.978$. Also, cooperation in game 3 of RM/WH is significantly lower than that in game 1, $p < 0.0001$. Yet with $p = 0.465$, WHP seems to resist this decline.

To investigate the treatment effects, Table 3 summarizes the most relevant group data in game 2, limited to rounds 6-23 to avoid potential initial learning and end-game effects.[3] p2(c) denotes the rate of cooperation; p2(cc|c) the likelihood that a C-player encounters another C in the game, which measures the assortativeness endogenously induced; and avg-r is the average payoff that measures how efficient the group is at solving the problem of cooperation. Pair-wise KW-tests show that WH and RM are not significantly different in any of those variables.

---

[3] Aggregate data display a clear drop in cooperation in the last two rounds for all treatments.

Table 3: Group level Game-2 summary over rounds 6-23

		p2(c)	p2(cc|c)	avg-r	avg. sw	avg. sw (w/o 0)	#sw=0	#sw≥6
RM	avg	0.202	0.217	4.319	3.528	5.512	4.8	4.4
	std	*0.045*	*0.118*	*0.270*	*0.931*	*0.665*	*3.271*	*1.140*
WH	avg	0.273	0.430	4.606	3.771	5.082	4.4	4.2
	std	*0.154*	*0.297*	*0.815*	*1.780*	*1.301*	*2.881*	*2.864*
WHP	avg	0.367	0.386	4.864	5.814	6.217	1	7.8
	std	*0.124*	*0.096*	*0.360*	*0.708*	*1.078*	*1*	*1.789*
KW-test	RM-WH	0.602	0.251	0.602	0.917	0.465	0.832	0.832
	RM-WHP	0.009	0.029	0.009	0.017	0.347	0.071	0.019
	WH-WHP	0.465	0.917	0.465	0.059	0.117	0.070	0.045

However, the picture changes *dramatically* when information is added to WH. WHP is significantly different from RM across all relevant variables. This suggests that assortative matching with group-income information indeed induces the desired change regarding cooperation in the PD framework.

With a closer look, however, the WH treatment conspicuously displays overall much higher standard deviations than RM. The Ansari-Bradley equal-dispersion test shows that WH and RM indeed differ significantly, for both p2(c) and avg-r with $p = 0.011$.

Thus, WH has the potential to improve cooperation over RM, but without further institutional aids, it may also be amplifying the group dynamic towards more defection, as exemplified in the bifurcation phenomenon in the data (see Yang et al., 2007 for elaboration). With some gross information provided, however, the pitfall of quick convergence towards defection can apparently be avoided, so that improved cooperation can be achieved across the map. Not surprisingly, assortative matching manages to positively align average payoffs and average rate of cooperation ($\rho = 0.7705/0.5771$ for WH/WHP), while RM displays the typical negative correlation $\rho = -0.5717$.

Given the bifurcation nature of WH, it is no surprise that it cannot be distinguished statistically from WHP, in variables discussed so far. To further investigate the effect of information, Table 3 also summarizes the switching behavior in more detail. Apparently, RM and WH are not significantly different regarding switching behavior, while WHP displays a significantly higher rate of switching than both. Moreover, the origin for this increment can be traced back to the smaller number of subjects who never changed actions (#sw = 0), and to those who switched more than 33% of the time (#sw $\geq$ 6). Thus, the income-class information seems to motivate people to change their actions more often.

From Figure 2, where payoff position 1 refers to the lowest income class, it is obvious that WHP increases the switching behavior uniformly over all 5 income classes, i.e. generally induces more motion. From Figure 1, WHP appears to induce a much higher increment of cooperation rate in the lower two than in the higher three income classes, compared to WH. For

RM, the most cooperative subjects are always in the lowest income class, due to negative correlation between payoff and cooperation.
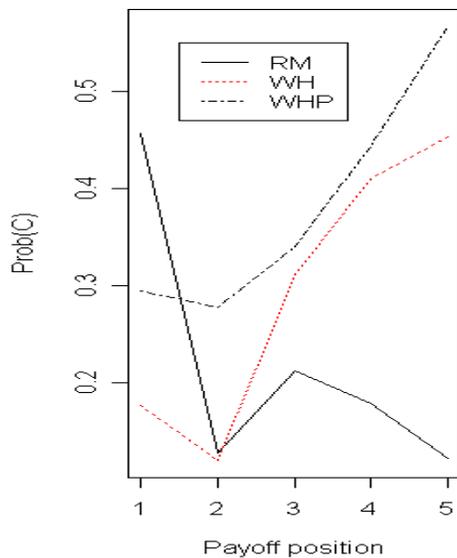
Figure 1. Payoff position vs. P(C)   Figure 2. Payoff position vs. P(switch)
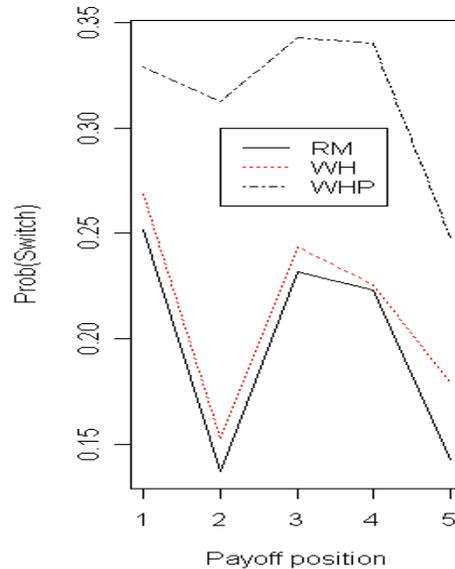
With these insights, we conduct logistic regressions that help to demonstrate the treatment effect in general, and to illustrate the role of the income-class information in particular.

Model 1 in Table 4 is adapted from Yang et al. (2007), applied only to data in RM/WH. The result indicates a strong inertia effect, i.e. the oftener one has been playing cooperation recently, the more likely he will do the same in the next round. The treatment and group dummies seem to aptly reflect the bifurcation phenomenon in WH.

Table 4: Regressions

	Model 1 (RM/WH)	*Model 2 (RM/WH)*	*Model 3 (RM/WH/WHP)*	*Model 4 (RM/WH/WHP)*
Intercept	-2.5279	-3.1827	-3.0029	-3.1000
T	0.8737	0.8762	0.6943	0.6893
T^2	-0.1643	-0.1592	-0.1115	-0.1113
T^3	0.0106	0.0102	0.0069	0.0069
Groups Dummy	-0.7035			
WH Treatment Dummy	0.3243 (.011)			
Avg. Payoff Information Dummy		0.1412	0.1633	0.1780 0.2723 (.034)
Log (likelihood)	-992.616	-991.365	-1733.115	-1730.921
Goodness-of-Fit	.700	.244	.155	.248
Concordant	79.8%	80.6%	78.5%	78.8%

Note: All estimations are significant with $p < 0.0001$ unless noted otherwise in parentheses.

4

After numerous error-and-trials in this study, Model 2 turns out to be a better fit by replacing the two dummies with one single variable, the 5-round average payoff of the subject, for the same data. The intuition is that individual payoffs are positively correlated with inclination for cooperation in assortative matching. The dummies in Model 1 are to capture the more cooperative and better-earning subjects and thus turn out to be well-aligned with the former. Extending to data of all three treatments, Model 3 also yields quite satisfactory estimations.

Model 4, however, shows that adding the Information dummy can indeed further improve the regression in a significant way (dummy value = 1, if first it is in WHP, second it is in the lowest income classes 1 and 2, and third [my 5-round avg. payoff] – [avg. payoff of the highest 3 subjects] < -2; and zero otherwise). Intuitively, in WHP, the knowledge of currently belonging to the lowest income classes might induce an additional inclination to play C if my current payoff is sufficiently lower (by at least 2 NT) than the best in the society. In other words, this dummy serves as a reminder that better-rewarding actions might be still out there. Due to a general tendency toward a positive correlation between payoff and cooperation rate in WHP, frequent defectors are indeed more likely to find their way into the lowest income classes.

4. Conclusion

In this study, we observe that cooperation can indeed prevail in a setting of artificial assortative matching. In particular, we find that information about one's social rank of income is conducive to generating increased cooperation uniformly over all societies. Subjects tend to be more cooperative when they fall in the lower ranks in the society, compared to when they are kept in the dark about it. Without this information, society can indeed be drawn into faster convergence into the all-defection equilibrium.

References

Amann, E. and C.-L. Yang (1998) "Sophistication and the Persistence of Cooperation" *Journal of Economic Behavior and Organization* **37**, 91-105.

Bergstrom, T. (2003) "The Algebra of Assortative Encounters and the Evolution of Cooperation" *International Game Theory Review* **5**(3), 211-228.

Bowles, S. and H. Gintis (2002) "The Origins of Human Cooperation" in *The Genetic and Cultural Origins of Cooperation* by P. Hammerstein, Ed., MIT press: Cambridge.

Brosig, J. (2002) "Identifying cooperative behavior: Some experimental results in a prisoner's dilemma game" *Journal of Economic Behavior and Organization* **47**(3), 275-290.

Frank, R.H. (1988) *Passions within Reason: The Strategic Role of the Emotions*, Norton: New York.

Frank, R.H., T. Gilovich, and D.T. Regan (1993) "The evolution of one-shot Cooperation: An

Experiment" *Ethnology and Sociobiology* **14**, 247-256.

Orbell, J.M. and R.M. Dawes (1993) "Social Welfare, Cooperators' Advantage, and the Option of Not Playing the Game" *American Sociological Review* **58**, 787-800.

Robson, A.J. (1990) "Efficiency in Evolutionary Games: Darwin, Nash, and the Secret Handshake" *Journal of Theoretical Biology* **144**, 379-396.

Sober, E. and D.S. Wilson (1998) *Unto Others: The Evolution and Psychology of Unselfish Behavior*, Harvard University Press: Cambridge.

Yamagishi, T., S. Tanida, R. Mashima, E. Shimoma, and S. Kanazawa (2003) "You can judge a book by its cover: Evidence that cheaters may look different from cooperators" *Evolution and Human Behavior* **24**, 290-301.

Yang, C.-L., J.C. Yue, and I.-T. Yu (2007) "The Rise of Cooperation in Correlated Matching Prisoners Dilemma: An Experiment" *Experimental Economics* **10**, 3-20.