Kochi University of Technology

Social Design Engineering Series SDES-2014-12

Reexamination of teams play in mixed-strategy game experiment

Yoshitaka Okano
Kochi University of Technology
Research Center for Social Design Engineering, Kochi University of Technology

24th October, 2014

School of Economics and Management
Research Center for Social Design Engineering
Kochi University of Technology

Reexamination of Teams Play in Mixed-Strategy Game Experiment

Yoshitaka Okano*

October 23, 2014

Abstract

Palacios-Huerta and Volij (2008) found that the behavior of professional soccer players in two-person zero-sum games is consistent with minimax play, while Wooders (2010) reexamined their data and found inconsistencies in several respects. This study applies a similar analysis of Wooders (2010) to the experimental data in Okano (2013), which found that the behavior of teams of two student subjects conforms closely to minimax play, and addresses whether teams exhibit the same inconsistencies as professionals. Teams were found to have consistency with minimax play, with no tendencies similar to those of professionals.

Keywords: Experimental game, teams' decision making, minimax, zero-sum
JEL Classifications: C72, C92

1 Introduction

The experimental and field data in two-player zero-sum games with a unique Nash equilibrium in mixed strategies reveal a difference between the behavior of professionals who are presumably familiar with strategic situations requiring unpredictability and that of students who are unfamiliar with them. General laboratory observations indicate that student subjects do not consistently play with the equilibrium strategy, especially when looking at decision-maker-level data (O'Neill, 1987; Brown and Rosenthal, 1990; Rapoport and Boebel, 1992; Mookherjee and Sopher,

*School of Management, Kochi University of Technology, 185 Miyanokuchi, Tosayamada, Kami, Kochi 782-8502 JAPAN. E-mail: oknystk@gmail.com. Phone: +81-887-57-2401. Fax: +81-887-57-2124. Financial support from the Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (JSPS) and the Ministry of Education, Culture, Sports, Science and Technology (MEXT) is gratefully acknowledged. The author is responsible for any errors.

1994; Binmore, Swierzbinski, and Proulx, 2001; Shachat, 2002; Rosenthal, Shachat, and Walker, 2003). On the other hand, several papers find that the behavior of professional sports players on the field is consistent with the minimax hypothesis (Walker and Wooders, 2001; Chiappori, Levitt, and Groseclose, 2002; Palacios-Huerta, 2003; Hsu, Huang, and Tang, 2007). Palacios-Huerta and Volij (2008) find that the play of professional soccer players in the lab in the well-known O'Neill (1987) game and the penalty-kick game they introduced conforms closely to the behavior predicted by the theory, whereas the play of student subjects does not.

Okano (2013) examines the behavior of teams of two students where team members are freely allowed face-to-face discussion to reach a single decision. Okano (2013) finds that when teams play the O'Neill game against other teams, their behavior is consistent with the minimax hypothesis. An important question arises: do teams behave in the same manner as individual professionals? This is important because if teams behave like professionals, we can provide an explanation for why professionals do well in the field and in the lab such that individual professionals can reason, think, and process information as well as two or more amateurs. The present study reexamines Okano's (2013) data and further explores the characteristics of the behavior of teams. In particular, we focus on the observation found by Wooders (2010) about professionals' inconsistencies with the theory. Wooders (2010) reexamines the data in Palacios-Huerta and Volij (2008) and finds the following.

1. Concerning the overall data, since the action frequencies of many professionals are too close to the theoretically expected play, the distribution of action frequencies across professionals is far from the distribution implied by the minimax hypothesis.

2. When the data are partitioned into halves, the behavior of professionals is not consistent with the minimax hypothesis in several respects. The choice frequencies are far from those implied by the minimax hypothesis at more than the expected rate for both professionals and student subjects. Professionals tend to follow nonstationary mixtures, with action frequencies that are negatively correlated between the first and second halves of the experiment, whereas student subjects do not. In particular, professionals tend to switch between halves between under- and overplaying an action relative to its equilibrium frequency.

We apply a similar analysis to the experimental data in Okano (2013) and find that the behavior of teams is, consistent with the theory, different from that of individuals, and that it does not have similar tendencies to that of professionals. Concerning the overall data, the distribution of choice frequencies across teams is reasonably consistent with that implied by the minimax hypothesis, whereas that of individuals is not for some choices. When partitioning the data into halves, the number of teams that make choices far from minimax play is slightly less than the expected rate, while that of individuals is larger than the expected rate. The distribution of choice frequencies across teams conforms to the theoretically expected distribution, whereas that of individuals does not for some actions. The play of teams in the first and second halves is not correlated, and there is no tendency for teams to switch their choices between halves, while that of individuals is positively correlated for some actions.

2 Okano's Experiment

Okano (2013) compared the behavior of teams (of two subjects each) and individuals using the game developed by O'Neill (1987), which is shown in Table 1. In the

Table 1: The O'Neill Game

		Player 2			
		J	*A*	*D*	*T*
Player 1	*J*	1,0	0,1	0,1	0,1
	A	0,1	0,1	1,0	1,0
	D	0,1	1,0	0,1	1,0
	T	0,1	1,0	1,0	0,1

O'Neill game, two players choose one of four alternatives, *Joker*, *Ace*, *Deuce*, and *Trey* (*J*, *A*, *D*, and *T* for short), which determines whether they win or lose.[1] The minimax hypothesis requires each player to choose *J*, *A*, *D*, and *T* with probabilities 0.4, 0.2, 0.2, and 0.2, respectively. Subjects played 15 times for practice and 150 times for real money. Because of a computer problem in the 133rd round, Okano (2013) uses the data through the 132nd round in the analysis, as does this study. In each round, winning individuals earned 50 yen, while winning teams earned 100 yen divided equally between team members. Losers earned nothing. Hence, the per-subject monetary incentives were the same across teams and individuals. We refer to the team-versus-team experiment as the *team treatment*, and the individual-versus-individual experiment as the *individual treatment*. There were 36 teams (72 students) in the team treatment and 36 individuals (36 students) in the individual treatment.

3 Reexamination

Tables 2 and 3 show the observed choice frequencies of teams and individuals, respectively, in the first and second halves, as well as the results of the statistical

[1] In the experiment, *Green*, *Red*, *Brown*, and *Purple* were used for the action labels in order to avoid the *Ace* bias, as suggested in Shachat (2002), but the original action labels were used to describe the results.

Table 2: Team Treatment

		First Half					Second Half					Overall				
Pair	Player	*J*	*A*	*D*	*T*		*J*	*A*	*D*	*T*		*J*	*A*	*D*	*T*	
1	1	27	15	12	12		24	15	15	12						
	2	27	10	12	17		26	19*	11	10						
2	1	25	8	17	16		25	15	10	16						
	2	20	18	12	16		22	12	17	15		*				
3	1	33*	15	12	6**		30	11	12	13		*				
	2	33*	13	13	7*		30	10	18	8		*			**	‡
4	1	32	16	7*	11		31	14	11	10		*		*		
	2	26	15	13	12		24	17	12	13						
5	1	37**	5**	11	13	‡	39**	14	7*	6**	‡	**		*		‡
	2	29	10	9	18		36**	9	10	11		**				†
6	1	33*	10	12	11		27	11	17	11						
	2	28	14	10	14		27	9	15	15						
7	1	30	14	16	6**		22	16	12	16						
	2	29	17	15	5**	†	24	17	16	9			*		**	‡
8	1	28	12	14	12		34*	11	10	11						
	2	26	16	13	11		29	16	10	11						
9	1	25	10	16	15		33*	11	9	13						
	2	35**	9	8	14		31	8	18	9		**	**			†
10	1	25	15	15	11		32	11	13	10						
	2	26	16	12	12		35**	11	10	10						
11	1	22	17	16	11		17**	17	15	17		**	*			†
	2	25	16	16	9		33*	11	10	12						
12	1	25	14	14	13		25	13	12	16						
	2	24	12	15	15		29	10	17	10						
13	1	28	13	10	15		21	14	13	18						
	2	27	13	14	12		24	19*	10	13						
14	1	25	11	12	18		25	11	18	12						
	2	24	10	15	17		32	9	13	12						
15	1	30	12	13	11		26	12	13	15						
	2	28	13	15	10		33*	12	9	12						
16	1	27	12	15	12		26	16	12	12						
	2	28	14	14	10		31	9	14	12						
17	1	21	18	11	16		27	9	14	16						
	2	31	9	13	13		24	13	14	15						
18	1	25	14	10	17		33*	14	9	10						
	2	29	9	15	13		23	12	19*	12				*		

** and * denote rejection of the minimax binomial model for a given choice at the 5% and 10% levels, respectively. ‡ and † denote rejection of the minimax multinomial model at the 5% and 10% levels, respectively.

Table 3: Individual Treatment

		First Half					Second Half					Overall				
Pair	Player	*J*	*A*	*D*	*T*		*J*	*A*	*D*	*T*		*J*	*A*	*D*	*T*	
1	1	37**	14	5**	10	‡	30	14	13	9		**		*		‡
	2	29	11	14	12		25	10	14	17						
2	1	32	11	9	14		38**	9	10	9	‡	**				‡
	2	32	12	15	7*		30	14	8	14						
3	1	31	13	9	13		25	16	15	10						
	2	30	12	10	14		27	14	13	12						
4	1	36**	11	7*	12	†	32	11	11	12		**		*		‡
	2	40**	1**	15	10	‡	29	17	9	11		**	*			‡
5	1	31	13	12	10		24	13	10	19*						
	2	28	16	13	9		23	18	14	11			*			
6	1	31	9	12	14		34*	11	12	9		**				
	2	28	12	14	12		25	12	7*	22**	‡				*	
7	1	28	9	15	14		29	12	15	10						
	2	36**	9	7*	14	†	33*	12	12	9		**				‡
8	1	30	17	11	8		30	18	6**	12	†		*	**		‡
	2	32	6**	15	13		35**	5**	15	11	‡	**	**			‡
9	1	27	14	13	12		24	16	16	10						
	2	30	12	11	13		29	14	6**	17				**		
10	1	23	12	16	15		24	13	14	15						
	2	25	11	19*	11		29	8	14	15						
11	1	30	13	9	14		26	12	14	14						
	2	33*	14	8	11		34*	7*	16	9	†	**				†
12	1	22	14	15	15		17**	19*	13	17	†	**				†
	2	25	11	14	16		20	14	15	17						
13	1	30	10	13	13		27	11	13	15						
	2	30	10	15	11		19*	14	16	17						
14	1	28	12	15	11		22	17	14	13						
	2	24	14	13	15		24	13	17	12						
15	1	22	18	11	15		23	10	15	18						
	2	34*	9	15	8		30	12	12	12		**				
16	1	20	17	16	13		17**	22**	11	16	‡	**	**			‡
	2	17**	16	15	18		27	16	6**	17					*	†
17	1	25	16	13	12		19*	14	20**	13						
	2	26	13	11	16		31	12	10	13						
18	1	24	23**	10	9	‡	17**	13	17	19*	†	**	**			
	2	19*	18	12	17		22	14	16	14		**				

** and * denote rejection of the minimax binomial model for a given choice at the 5% and 10% levels, respectively. ‡ and † denote rejection of the minimax multinomial model at the 5% and 10% levels, respectively.

tests. Okano (2013) analyzed the overall data and found the actual play of teams to be consistent with the minimax hypothesis. The chi-square goodness-of-fit tests for the minimax multinomial model reject the null hypothesis at the 5% level for three teams and eight individuals (1.8 rejections are expected). The chi-square goodness-of-fit tests for the minimax binomial model for a given action reject the null hypothesis at the 5% level for seven instances for teams and 18 instances for individuals (7.2 rejections are expected).

The chi-square test investigates how close the observed choice frequencies are to the theoretically expected ones. According to the minimax hypothesis, each action should be a *random draw* from the multinomial distribution equal to the mixed strategy equilibrium. Random draws mean that the realized actions of the minimax play are not always exactly equal to the equilibrium frequencies. In other words, we must reject the minimax hypothesis, for example, for experimental data in which the choice frequencies of *all* decision makers are *exactly* equal to the equilibrium ones. For the chi-square goodness-of-fit test of the minimax multinomial model, minimax plays with 36 decision makers imply that 36 test statistics are realizations of 36 random draws from the chi-square distribution with three degrees of freedom. Equivalently, the p-values associated with the realized test statistics should be 36 draws from the uniform distribution $U[0,1]$. Table 4 shows the results of the Kolmogorov–Smirnov test to the empirical distribution of 36 p-values from the tests for the minimax multinomial model and those for the minimax binomial model for each action. For teams, the Kolmogorov–Smirnov tests cannot reject the null hypothesis at the 5% level for any action or for the four actions jointly. On the other hand, the distributions of the four actions jointly, *Joker*, and *Ace* for individuals are far from those implied by the minimax hypothesis. These rejections stem from the fact that each empirical cumulative distribution function has many small p-values.

Table 4: P-values of the Kolmogorov–Smirnov Test of Conformity to $U[0,1]$

	Teams	Individuals
All	0.5777	0.0266
J	0.2813	0.0004
A	0.0781	0.0123
D	0.1595	0.3688
T	0.3116	0.1893

Next, we simply partition the experimental data into two groups of 66 rounds each. The minimax hypothesis requires the players to follow the stationary multinomial distribution over actions with parameters equal to the mixed strategy equilibrium, which implies that the choice frequencies should exhibit similar conformity to the equilibrium ones in each half.

In both halves, the null hypothesis that a player chooses an action according to the minimax multinomial model is rejected at a lower rate than expected for teams. On the other hand, it is rejected at higher rate for individuals, even though tests based on only half data have less power. It is rejected at the 5% level for one team and three individuals in the first half and for one team and four individuals in the second half (1.8 rejections are expected under the null). This is also the case for the null hypothesis that a player chooses a given action according to the minimax binomial model. It is rejected at the 5% level in six instances for teams and in nine instances for individuals in the first half and for five instances for teams and in 12 instances for individuals in the second half (7.2 rejections are expected under the null).

Table 5 shows the results of the Kolmogorov–Smirnov test of conformity to $U[0, 1]$ of the observed p-values from the chi-square goodness-of-fit test. As in the overall data, Kolmogorov–Smirnov tests for teams cannot reject the null hypothesis at the 5% significance level for any action or for the four actions jointly. On the other hand, the same null is rejected at the 5% level for *Joker* for individuals in both halves, with the empirical cumulative distribution function having many small p-values.

Table 5: P-values of the Kolmogorov–Smirnov Test of Conformity to $U[0, 1]$

	First Half		Second Half	
	Teams	Individuals	Teams	Individuals
All	0.1353	0.9617	0.2130	0.3109
J	0.0603	0.0071	0.0876	0.0128
A	0.7243	0.5481	0.2323	0.0675
D	0.0546	0.2514	0.1944	0.9461
T	0.6200	0.2514	0.1858	0.1944

Another implication of the minimax hypothesis is that the frequency with which a given action is played in the first half is statistically independent of the frequency with which it is played in the second half. This implies, for example, that a player should not switch between halves between under- and overplaying an action relative to its equilibrium frequency.

Table 6 shows the values of Spearman's rank correlation coefficient R for each action and the results of the statistical test of independence based on it. The independence hypothesis is not rejected for any action of teams at the 5% level. For individuals, on the other hand, independence is rejected for *Joker* and *Ace* at the 5% level, with the first- and second-half frequencies positively correlated.[2]

Of the 36 teams, 19 switched between halves between underplaying and overplaying *Joker* relative to its equilibrium frequency. The null hypothesis that the switching probability is 0.5 is not rejected at the 5% level ($p = 0.739$). For A, D, and

[2]Wooders (2010) also found that student subjects exhibit a positive correlation between halves for *Joker*.

Table 6: Spearman's Rank Correlation Coefficients

	Teams		Individuals	
	R	P-value	R	P-value
J	0.1143	0.5069	0.7070	0.0000
A	0.1530	0.3729	0.4194	0.0109
D	−0.0766	0.6571	−0.0079	0.9635
T	0.0951	0.5810	0.0877	0.6109

T, there are, respectively, 15, 20, and 13 switching teams, and the null hypothesis is not rejected at the 5% level for each action ($p = 0.317$ for A, $p = 0.505$ for D, and $p = 0.096$ for T). For the 36 individuals, there were 12, 12, 17, and 17 switches for J, A, D, and T, respectively, and the null hypothesis that the switching probability is 0.5 is rejected at the 5% level for J and A ($p = 0.046$ for J, $p = 0.046$ for A, $p = 0.739$ for D, and $p = 0.739$ for T). This is consistent with the finding of positive correlations for J and A between halves for individuals.

4 Conclusion

This study has found further evidence of teams' conformity to the minimax hypothesis in both overall data and partitioned data, whereas individuals exhibit some inconsistencies. It is important to address what causes the difference in behavior between teams and individuals. Okano (2013) addresses this question in additional experiments in which subjects play the hide-and-seek game developed in Rosenthal, Shachat, and Walker (2003) and team-versus-individual O'Neill game, but conclusive findings are not obtained. Moreover, the behavior of teams is also different from that of professionals. Teams do not exhibit the behaviors observed in Wooders (2010), such as the biased distribution of choice frequencies, the negative correlation of play between halves, and switching between halves between under- and over-playing an action relative to its equilibrium frequencies. This suggests that there may be some differences between teams and professionals in terms of reasoning, thinking, information processing, and other factors that shape their actions.

References

[1] Binmore, Ken, Joe Swierzbinski, and Chris Proulx (2001), "Does minimax work? an experimental study," *Economic Journal*, 111(473), 445-464.

[2] Brown, James N., and Robert W. Rosenthal (1990), "Testing the minimax hypothesis: a re-examination of O'Neill's game experiment," *Econometrica*, 58(5), 1065-1081.

[3] Chiappori, Pierre-Andre, Steve Levitt, and Timothy Groseclose (2002), "Testing mixed-strategy equilibria when players are heterogeneous: the case of penalty kicks in soccer," *American Economic Review*, 92(4), 1138-1151.

[4] Hsu, Shin-Hsun, Chen-Ying Huang, and Cheng-Tao Tang (2007), "Minimax play at Wimbledon: comment," *American Economic Review*, 97(1), 517-523.

[5] Mookherjee, Dilip, and Barry Sopher (1994), "Learning behavior in an experimental matching pennies game," *Games and Economic Behavior*, 7(1), 62-91.

[6] Okano, Yoshitaka (2013), "Minimax play by teams," *Games and Economic Behavior*, 77(1), 168-180.

[7] O'Neill, Barry (1987), "Nonmetric test of the minimax theory of two-person zerosum games," *Proceedings of the National Academy of Sciences*, USA, 84(7), 2106-2109.

[8] Palacios-Huerta, Ignacio (2003), "Professionals play minimax," *Review of Economic Studies*, 70(2), 395-415.

[9] Palacios-Huerta, Ignacio, and Oscar Volij (2008), "Experientia docet: professionals play minimax in laboratory experiments," *Econometrica*, 76(1), 71-115.

[10] Rapoport, Amnon, and Richard B. Boebel (1992), "Mixed strategies in strictly competitive games: a further test of the minimax hypothesis," *Games and Economic Behavior*, 4(2), 261-283.

[11] Rosenthal, Robert W., Jason Shachat, and Mark Walker (2003), "Hide and seek in Arizona," *International Journal of Game Theory*, 32(2), 273-293.

[12] Shachat, Jason M. (2002), "Mixed strategy play and the minimax hypothesis," *Journal of Economic Theory*, 104(1), 189-226.

[13] Walker, Mark, and John Wooders (2001), "Minimax play at Wimbledon," *American Economic Review*, 91(5), 1521-1538.

[14] Wooders, John (2010), "Does experience teach? professionals and minimax play in the lab," *Econometrica*, 78(3), 1143-1154.