Bonn Econ Discussion Papers

Discussion Paper 23/2004

Minority Game - Experiments and Simulations of Traffic Scenarios

by

Thorsten Chmura, Thomas Pitz

December 2004

Bonn Graduate School of Economics
Department of Economics
University of Bonn
Adenauerallee 24 - 42
D-53113 Bonn

The Bonn Graduate School of Economics is sponsored by the

Deutsche Post World Net

MAIL EXPRESS LOGISTICS FINANCE

Minority Game
–
Experiments and Simulations of Traffic Scenarios

T. Chmura[1], T. Pitz[1], and M. Schreckenberg[2]

[1] Laboratory of Experimental Economics – University of Bonn
Adenauerallee 24-42 - 53113 Bonn – Germany
chmura@uni-bonn.de
[2] Physics of Transport and Traffic –University Duisburg-Essen
47048 Duisburg – Germany

Abstract: This paper reports laboratory experiments and simulations on a *minority game*. The minority game is the most important example for a classic non-zero-sum-game. The game can be applied on different situations with social and economic contests. We chose an elementary traffic scenario, in which subjects had to choose between a road *A* and a road *B*. Nine subjects participated in each session. Subjects played 100 rounds and had to choose between one of the roads. The road which the minority of players chose got positive payoffs. We constructed an extended reinforcement model which fits the empirical data.

Introduction and Experimental Set-Up

The minority game is the most important example for a classic non-zero-sum-game and can be applied on different situations with social and economic contests. Imagine two big and famous gold fields in South Africa, near Cape Town and Johannesburg. The diggers heard that a big gold-nugget was found in Johannesburg. From now on every digger went to Johannesburg to dig gold, the city got overcrowded and there was not enough space for all of them, so the profit was very small. The diggers who stayed in Cape Town on the other hand had enough space for their claims. The profit in Cape

Town was very high for everybody. This is an example of the minority game, the people who choose the majority got no payoffs, but the people on the minority in Cape Town found enough gold for all of them, so everybody got a payoff.

The minority game which is also called the El Farol Bar Problem (EFPB), was introduced by Arthur [1]. The setup of the minority game is the following: a number of agents *n* have to choose in several periods whether to go in room A or B. Those agents who have chosen the less crowded room win, the others lose. Later on, the EFBP was put in a mathematical framework by Challet and Zhang [2], the so-called Minority Game (MG). An odd number *n* of players has to choose between two alternatives (e.g., yes or no, A or B, or simply 0 or 1). In the Literature are many examples, where the MG is discussed [2,3,4]. In this paper we transferred the minority problem into a route choice context. We did minority game experiments at the Laboratory of Experimental Economics (University of Bonn). In these Experiments subjects are told that in each of 100 periods they have to make a choice between a road *A* and road *B* for travelling from X to Y.

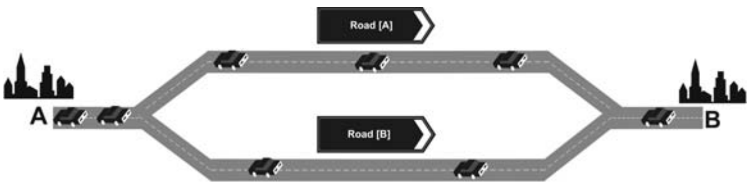

Fig. 1. Participants had to choose between a road [A] and a road [B].

The number of subjects in each session was 9. They were told the time t_A and t_B depends on the numbers n_A and n_B of participants choosing *A* and *B*, respectively:

$$t_A = 1,\ t_B = 0 \ \Leftrightarrow\ n_A < n_B$$

$$t_B = 1,\ t_A = 0 \ \Leftrightarrow\ n_A > n_B\ .$$

The period payoff was t_A if *A* was chosen and t_B if *B* was chosen. The total payoff of a subject was the sum of all 100 period payoffs converted to money payoffs in Euro [€] with a fixed exchange rate of 0.2 € for each experimental money unit (Taler).

Additionally every participant received a show-up fee of 3 €. One session took roughly one hour. There are no pure equilibria in this game. The pareto-optimum can be reached by 4 players on one road and 5 players on the other road. Two treatments have been investigated. In treatment I the subjects received information about: whether own last choice was in the minority or majority, the last chosen route, the payoff of the last period in Taler, the cumulated payoff in Taler and thenumber of the actual period. In treatment II additional feedback was provided about distribution on both-routes in the last period. Six sessions were run with treatment I and six with treatment II. No further information was given to the subjects.

Observed Behaviour

Number of Players on the Road *A*

Figure 2 shows the number of participants on the road *A* as a function of time for a typical session of treatment I. Fig. 3 shows the number of participants on the road *A* as a function of time for a typical session of treatment II. There are substantial fluctuations until the end of the session. The same is true for all sessions of both treatments. The mean number of players on the road *A* is 4.5 in treatment I and 4.43 in treatment II. This was expected, because of the experimental setup, there is no preference for one road. The fluctuations can be measured by the standard deviation of the number of participants choosing *A* per period. This standard deviation is between 0,67 and 1,5. In view of these numbers one can speak of substantial fluctuations in each of the 12 sessions. The fluctuations are obvious larger under treatment I than under treatment II. The effect is significant. The null-hypothesis is rejected by a Wilcoxon-Mann-Whitney-Test on the significance level of 1 % (one sided).

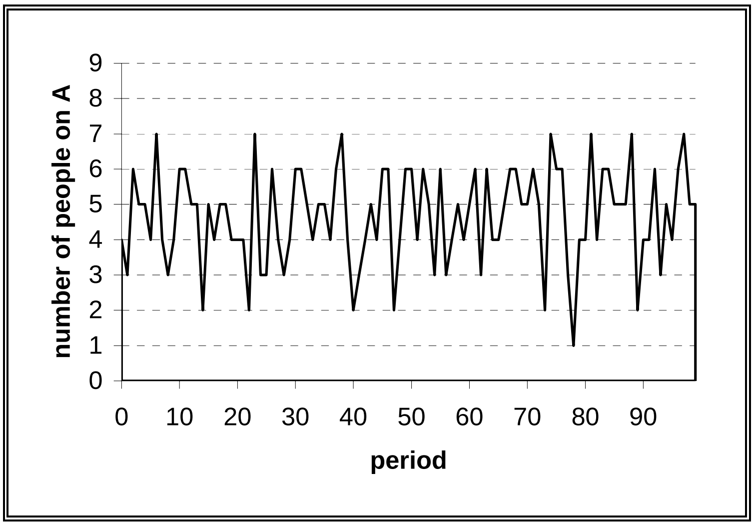

Fig. 2. Number of participants on *A* [a typical session of treatment I].

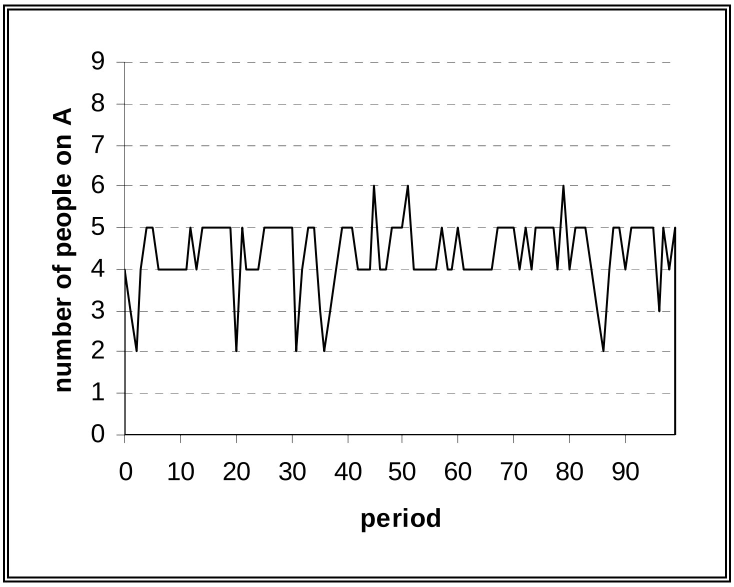

Fig. 3. Number of participants on *A* [a typical session of treatment II].

The non existence of pure strategy equilibria poses a coordination problem which may be one of the reasons for non-convergence and the persistence of fluctuations. Feedback on both travel times vs. feedback on only own travel time has a beneficial effect by the reduction of fluctuations. This effect is remarkable.

		number of players on *A*	
		mean	std. dev.
Treatment I	session I 01	4,33	1,36
	session I 02	4,74	1,50
	session I 03	4,41	1,50
	session I 04	4,40	1,31
	session I 05	4,65	1,33
	session I 06	4,44	1,28
	treatment I	4,50	1,38
Treatment II	session II 01	4,19	1,35
	session II 02	4,62	1,19
	session II 03	4,36	1,05
	session II 04	4,34	0,97
	session II 05	4,62	0,84
	session II 06	4,50	0,67
	treatment II	4,44	1,01

Tab. 1. Mean and standard deviation of participants on *A*.

Road Changes

Figure 4 shows an example of the number of road changes as a function of time for a typical session of treatment I. There was a negative trend in each session of treatment II. By comparison in treatment I there were four sessions with a positive and two with a negative trend. The fluctuations are connected to the total number of road changes within one session. The median number of road changes is significantly higher in treatment I. The null-hypothesis is rejected by the Wilcoxon-Mann-Whitney-Test on a level of 1% (one sided). The mean number of road changes under treatment I is also higher than under treatment II. A Wilcoxon-Mann-Whitney-Test rejects the null-hypothesis only on a significance level of 1% (one sided).

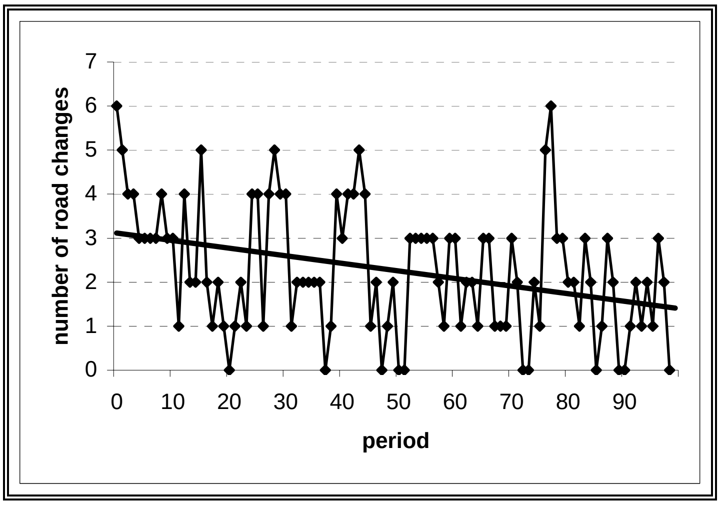

Fig. 4. Number of road changes [a typical session of treatment I].

Under treatment I subjects who mainly choose only one of the roads feel the need to travel on the other road from time to time in order to get information on both roads. Under treatment II there is no necessity for such information gathering. This seems to be the reason for the greater number of changes and maybe also for the stronger fluctuations under treatment I.

		number of road changes
		mean
Treatment I	session I 01	5,08
	session I 02	3,87
	session I 03	5,16
	session I 04	5,19
	session I 05	5,28
	session I 06	4,35
	treatment I	4,82
Treatment II	session II 01	3,99
	session II 02	3,68
	session II 03	3,67
	session II 04	5,19
	session II 05	4,67
	session II 06	4,44
	treatment II	4,27

Tab. 2. Mean and standard deviation number of road changes.

Payoffs and Road Changes

In all sessions except one session of treatment I the number of road changes of a subject is negatively correlated with the subject's payoff. Tab. 4 shows that the negative correlation between the payoff and number of road changes in treatment II is higher than in treatment I.

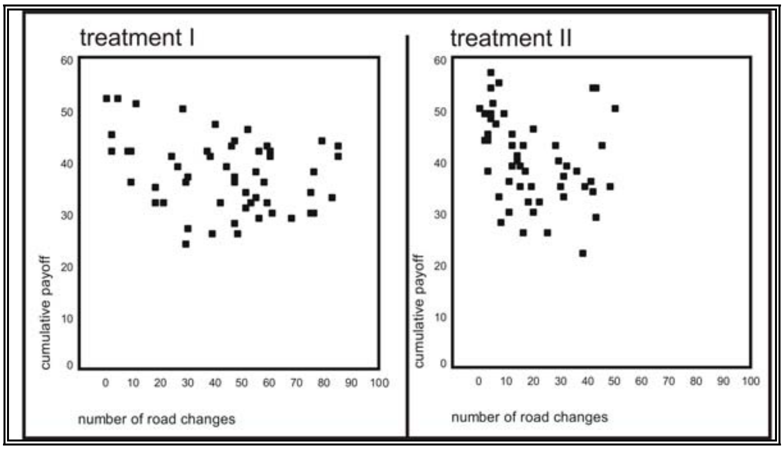

Fig. 5. Scatter diagram cumulative payoff/number of road changes for treatment I and II.

		Spearman rank correlation between cumulative payoffs and number of road changes
Treatment I	session I 01	-0,48
	session I 02	0,34
	session I 03	-0,44
	session I 04	-0,70
	session I 05	-0,18
	session I 06	-0,18
	treatment I	-0,27
Treatment II	session II 01	-0,51
	session II 02	-0,54
	session II 03	-0,30
	session II 04	-0,82
	session II 05	-0,27
	session II 06	-0,78
	treatment II	-0,54

Tab. 3. Spearman rank correlation between cumulative payoffs and number of road changes for treatment I and II.

In every eleven observations of both treatments the Spearman rank correlations between cumulative payoffs and the number of road changes is negative. The Spearman-correlation-coefficients in treatment II are lower than in treatment I. This effect is small but significant. A Wilcoxon-Mann-Whitney-Test rejects the null-hypothesis on a significance level of 10% (one sided). Even if subjects change roads in order to get higher payoffs, they do not succeed in doing this on the average. This suggests that it is difficult to use the information provided by the feedback to one’s advantage.

Response Mode

A participant who had no payoff on the road chosen may change his road in the next period in order to travel where it is less crowded. We call this the *direct* response mode. The *direct* response mode is the prevailing one but there is also a *contrarian* response mode. The contrarian participant expects that a positive payoff will attract many others and that therefore the road chosen will be crowded in the next period. For each subject let c_- (c_+) be the number of times in which a subject changes the roads when there was a payoff p=0 (p=1) in the period before. And for each subject let s_- (s_+) be the number of times in which a subject stays on the road when there was a payoff s=0 (s=1) in the period before.

	change	**stay**
p=0	c_-	s_-
p=1	c_+	s_+

Tab. 4. 2x2 table for the computation of Yule coefficients.

For each subject such a 2x2 table has been determined and a Yule coefficient Q has been computed as follows.

$$Q = \frac{c_- \cdot s_+ - c_+ \cdot s_-}{c_- \cdot s_+ + c_+ \cdot s_-}$$

The Yule coefficient has a range from –1 to +1. In our case a high Yule coefficient reflects a tendency towards direct responses and a low one a tendeThe mean and the standard deviation of the Yule coefficients are shown in Tab. 6. The mean Yule coefficients are significantly higher in treatment II. The null-hypothesis for both treatments is rejected by a Wilcoxon-Mann-Whitney-Test on the significance level of 1% (one sided). That means there are less contrarian response modes in treatment II.

		Yule coefficiants Q	
		mean	std. dev.
Treatment I	session I 01	0,14	0,62
	session I 02	0,15	0,43
	session I 03	0,27	0,76
	session I 04	0,01	0,47
	session I 05	0,11	0,75
	session I 06	-0,01	0,54
	treatment I	0,11	0,60
Treatment II	session II 01	0,21	0,75
	session II 02	0,42	0,39
	session II 03	0,48	0,61
	session II 04	0,72	0,40
	session II 05	0,68	0,66
	session II 06	0,87	0,33
	treatment II	0,56	0,52

Tab. 5. Mean and standard deviation of the Yule coefficients in both treatments.

Simulations

In order to get more insight into this theoretical significance of our result, we have run simulations based on a version of a well known reinforcement learning model, the payoff-sum model. This model already described by Harley (1981) [5] and later by Arthur (1991) [6] has been used extensively by Ereth and Roth [7,8] in the experimental economics literature. Here we used an extended payoff-sum model, which is already published in Selten et al. [9]. Table 7 explains the version underlying our simulations. We are looking at player *i* who has to choose among *n* strategies 1,…,*n* over a number proportional to its “propensity” $x_{i,j}^{t}$. In period 1 these propensities are exogenously determined parameters. Whenever the strategy *j* is used in period *t*, the resulting payoff a_i^t is added to the propensity if this payoff is positive. If all payoffs are positive, then the propensity is the sum of all previous payoffs for this strategy plus its initial propensity. Therefore one can think of a propensity as a payoff sum. In our simulations we chose the same conditions as in the experiments. For 100 periods 9 players (agents) interact with each other. Each player has four strategies:

1. *road A*: This strategy simply consists in taking the decision for the road *A*.

2. *road B*: This strategy consists in taking the road *B*.

3. *direct*: This strategy corresponds to the direct response mode. The payoff of a player is 1, then the player stays on the road last chosen. If his payoff is 0 the players changes the road.

4. *contrarian*: This strategy corresponds to the contrarian response mode. The payoff of a player is 1, then the player changes the road. If his payoff is 0 the players will stay on the road.

In the first period only strategy one and two were available to the simulated subjects since strategy three and four cannot be applied because there is no previous payoff. In the simulations we did not want to build in prejudices based on theoretical values. Our simulated players base their behaviour on initial propensities and observations only. Of course, it is assumed that as in the experiments the players get feedback about their own payoffs immediately after their choices. In the experimental treatment II additional feedback about the payoff on the route not chosen was given. The payoff sum model makes use of a player's own payoff only and therefore ignores the additional feedback of treatment II.

The differences between treatment I and treatment II cannot be explained by the payoff sum model since it does not process the additional feedback information given in treatment II. For the purposes of comparing our simulation data with the experimental data we ignore the differences between treatment I and II which are not big anyhow.
The difficulty arises that the initial propensities must be estimated from the data. We did this by varying the initial propensities for the strategies *road A* and *road B* over all integer values from 1 to 10 and the initial propensities for the strategies *direct* and *contrarian* over all integer values from 0 to 10.

We compared the simulation results with the six variables listed in Tab. 8. We aimed at simulation results which were between the minimum and maximum experimental results over all twelve sessions of treatment I and II. For each of the 12100 parameter combinations we have run 1000 simulations.

Initialisation: For each player i let $[x^1_{i,1},...,x^1_{i,n}]$ the initial propensity, where n is the number of strategies, which are used in the simulations.

1. period: Each player i chooses strategy k with probability $\frac{x^1_{i,j}}{\sum_k x^1_{i,k}}$.

t+1. period: For each player i, let a^t_i the payoff of player i in period t, and j the number of the chosen strategy in period t.

IF $a^t_i \geq 0$:

$$x^{t+1}_{i,j} := x^t_{i,j} + a^t_i$$

$$x^{t+1}_{i,k} := x^t_{i,k}\,,\; k \neq j$$

ELSE

$$x^{t+1}_{i,j} := x^t_{i,j}$$

$$x^{t+1}_{i,k} := x^t_{i,k} - a^t_i\,,\; k \neq j$$

Each player i chooses strategy j with the probability $\frac{x^{t+1}_{i,j}}{\sum_k x^{t+1}_{i,k}}$.

Tab. 6. The extended payoff-sum model.

There were three parameter combinations which satisfied the requirement of yielding means for the six variables between the minimal and maximal experimentally observed values. This was the parameter combination *(1,1,2,1)* and *(2,2,1,1)* and *(3,3,4,2)*. The numbers refer to *road A*, *road B*, *direct* and *contrarian* in this order. The parameter combination is a reasonable vector of initial propensities. There is no difference between road *A* and road *B*, so it is reasonable to have the same propensities for both roads. In two of the three vectors the propensity of the direct mode is greater than the others propensities. There were especially in treatment II as you see on the yule coefficients more direct response modes. This could be an explanation for the direct propensities in the simulations.

	Treatment I & Treatment II				
		Simulations			
	Minimum	{1,1,2,1}	{2,2,1,1}	{3,3,4,2}	Maximum
Player on road A [mean]	4,19	4,48	4,50	4,54	4,74
Player on Road A [standard deviation]	0,67	1,45	1,48	1,50	1,50
Road Changes [mean]	0,59	4,32	4,18	4,51	5,17
Last Roadchange [mean]	54,44	96,11	97,67	97,44	98,11
Yule Coefficiant [mean]	-0,01	0,10	0,04	0,14	0,87
Yule Coefficiant [mean for every player]	0,33	0,50	0,40	0,35	0,76

Tab. 7. Experiments and simulations with 9 players.

It is surprising that a very simple reinforcement model reproduces the experimental data as well as shown by Tab. 8. Even the mean Yule coefficient is in the experimentally observed range. In spite of the fact that at the beginning of the simulation the behaviour of all simulated players is exactly the same. It is not assumed that there are different types of players.

Conclusion

Fluctuations persist until the end of the sessions in both treatments. Feedback on both road times significantly reduces fluctuations in treatment II compared to treatment I. This effect is strong. There is a significant rank correlation between the total number of road changes and the size of fluctuations. In treatment I road changes may serve the purpose of information gathering. This motivation has no basis in treatment II. However, road changes may also be attempt to improve payoffs. The finding of a negative correlation between a subject’s payoff and number of road changes suggests that on the average such attempts are not successful. Two response modes can be found in the data, a *direct* one in which road changes follow bad payoffs and a *contrarian* one in which road changes follow good payoffs. One can understand these response modes as due to different views of the causal structure of the situation. If one expects that the road which is crowded today is likely to be crowded tomorrow one will be in the direct response mode but if one thinks that many people will change to the other road because it was crowded today one has reason to be in the contrarian response mode. We have presented statistical evidence for the importance of the two response modes. We have also run simulations based on a simple payoff sum reinforcement model. Simulated mean values of six variables have been compared with the experimentally observed minimal and maximal of these variables. The simulated means were always in this range. Only four parameters of the simulation model, the initial propensities, were estimated from the data. In view of the simplicity of the model it is surprising that one obtains a quite close fit to the experimental data. The response modes *direct* and *contrarian* also appear in the simulations as the result of an endogenous learning behaviour by which initially homogeneous subjects become differentiated over time.

References

1. W. B. Arthur. Inductive reasoning and bounded rationality. Am. Eco. Rev. 84, 406 (1994).
2. D. Challet and Y.-C. Zhang. Emergence of cooperation and organization in an evolutionary game. Physica A 246, 407–418 (1997).
3. D. Challet and Y.-C. Zhang. On the minority game: Analytical and numerical studies. Physica A 256, 514–532 (1998).
4. N. F. Johnson, S. Jarvis, R. Jonson, P. Cheung, Y. R. Kwong, and P. M. Hui. Volatility and agent adaptability in a self-organizing market. Physica A 258, 230–236 (1998).
5. Harley, C. B.: 1981, Learning in Evolutionary Stable Strategy, J. Teoret. Biol. 89, 611-633.
6. Arthur W. B. 1991: Designing economic agents that act like human agents: A behavioural approach to bounded rationality. Amer. Econ. Rev. Papers Proc. 81 May, 353-359.
8. Roth, A.E., Erev, I.: 1995, Learning in extensive form games: Experimental data and simple dynamic models in the intermediate term, Games and economic Behavior 8, 164 – 212.
9. Roth, A.E., Erev, I.: 1998, Predicting how people play games: Reinforcement learning in games with unique mixed strategy equilibrium, American economic review 88, 848 – 881.
10. Selten, R., M. Schreckenberg, T. Pitz, T. Chmura, J. Wahle (2003). Experiments on Route Choice Behaviour. In Interface and Transport Dynamics, Lectural Notes in Computational Science and Engeneering. Springer Verlag, Heidelberg 2003, Pages 217.221.