How to Efficiently Distribute Scarce Resources: Volatile Decision Behavior and Optimal Guidance through Information Services

Dirk Helbing,[1] Martin Schönhof,[1] and Daniel Kern[1]

The coordinated and efficient distribution of limited resources by individual decisions is a fundamental and unsolved problem. When individuals compete for road capacities, time, space, money, etc., they normally take decisions based on aggregate rather than complete information, such as TV news or stock market indices. In related experiments, we have observed a volatile decision dynamics and far-from-optimal return distributions. We have also identified ways of information presentation that can considerably improve the overall performance of the system. In order to determine optimal strategies of decision guidance by means of user-specific recommendations, we have developed a stochastic behavioural description. These strategies manage to increase the adaptibility to changing returns and to reduce the deviation from the time-dependent user equilibrium, thereby enhancing the average and individual payoffs. Hence, our guidance strategies can increase the performance of all users by reducing overreaction and stabilizing decision behaviour. Our results are highly significant for predicting decision behaviour, for reaching optimal behavioural distributions by decision support systems, and for information service providers.

Optimal route guidance strategies in overloaded traffic networks require reli-

[1]Institute for Economics and Traffic, Dresden University of Technology, D-01062 Dresden, Germany, helbing@trafficforum.org, www.helbing.org

able traffic forecasts. Driver reactions to route choice recommendations, however, change the traffic situation, which invalidates forecasts. The hope of some keen scientists is to solve this problem by means of an iteration scheme (*1-6*): If the driver reaction was known from experiments (*7-14*), the resulting traffic situation could be calculated, yielding improved route choice recommendations, etc. Given this iteration scheme converges, it would facilitate optimal recommendations and reliable traffic forecasts anticipating the driver reactions. Later on, we will even propose an efficient *one-step* procedure.

To determine the route choice behaviour, Schreckenberg *et al.* (*14*) have recently carried out a route choice game. N test persons had to repeatedly decide between two alternatives 1 and 2 (the routes) and should try to maximize their resulting payoffs (describing something like the speeds or inverse travel times). To reflect the competition for a limited resource (the road capacity), the received payoffs $P_1(N_1) = P_1^0 - P_1^1 N_1$ and $P_2(N_2) = P_2^0 - P_2^1 N_2$ went down with the numbers of test persons N_1 and $N_2 = N - N_1$ deciding for alternatives 1 and 2, respectively. The user equilibrium corresponding to equal payoffs for both alternative decisions is found for a fraction $F_1^{\rm eq} = N_1/N = P_2^1/(P_1^1 + P_2^1) + [(P_1^0 - P_2^0)/(P_1^1 + P_2^1)]/N$ of persons choosing alternative 1 and agrees with the system optimum in the limit $N \to \infty$ of many participants. Small groups were chosen to investigate the fluctuations in the system. Schreckenberg *et al.* found that the test groups managed well to adapt to the user equilibrium on average. However, although it appears reasonable to stick to the same decision once the equilibrium is reached, the standard deviation stayed at a finite level. This was not only observed in *"treatment" 1*, where all players knew only their own payoff, but also in *treatment 2*, where

Electronic copy available at: http://ssrn.com/abstract=2693458

the payoffs $P_1(N_1)$ and $P_2(N_2)$ for both, 1- *and* 2-decisions, were presented to all players. Nevertheless, treatment 2 could decrease the changing rate and increase the average payoffs.

To explain the mysterious persistence in the changing behaviour and explore possibilities to support it, we have repeated these experiments with more iterations and tested additional treatments. In the beginning, all treatments were consecutively applied to the same players in order to determine the response to different kinds of information (see Fig. 1). Afterwards, single treatments and variants of them have been repeatedly tested with different players to check our conclusions. In *treatment 3*, every test person was informed about the own payoff $P_1(N_1)$ [or $P_2(N_2)$] and the *potential* payoff $P_2(N - N_1 + \epsilon N) = P_2(N_2) - \epsilon N P_2^1$ [or $P_1(N - N_2 + \epsilon N) = P_1(N_1) - \epsilon N P_1^1$] he or she would have obtained, if a fraction ϵ of persons had additionally chosen the other alternative (here: $\epsilon N = 1$ person). Treatments 4 and 5 were variants of treatment 3, but some payoff parameters were changed in time to simulate varying environmental conditions. In *treatment 5*, each player additionally received an individual recommendation which alternative to choose.

The higher changing rate in treatment 1 compared to treatment 2 can be understood as effect of an exploration rate ν_1 required to find out which alternative performs better. It is also plausible that treatment 3 could further reduce the changing rate: In the user equilibrium with $P_1(N_1) = P_2(N_2)$, every player knew that he or she would *not* get *the same*, but a *reduced* payoff, if he or she would change the decision. Therefore, the new treatment 3 reached the best adaptation performance, which is reflected by a very low standard deviation and higher average payoffs. But why did players change their decision in the user equilibrium at all? Figure 2 shows some kind of intermittent behaviour, i.e. quiet periods without changes followed by turbulent periods with many changes. This is reminiscent of volatility clustering in stock market indices (*15,16,17,18*), where individuals also react to aggregate information reflecting all decisions (the trading transactions). Single players seem to change their decision to reach above-average payoffs. In fact, although the cumulative individual payoff is anticorrelated with the average changing rate, some players receive higher payoffs with larger changing rates than others, i.e. they profit from the overreaction in the system. Once the system is out of equilibrium, all players respond in one way or another. Typically, there are too many decision changes. The corresponding overcompensation, which had also been predicted by computer simulations (*2,12,19,20*), gives rise to turbulent periods.

To avoid overreaction, in treatment 5 we have recommended a number $F_1^{\rm eq}(t+1)N - N_1(t)$ of players to change their decision and the other ones to keep it. These user-specific recommendations helped the players to reach the smallest overreaction of all treatments and a very low standard deviation, although the payoffs were changing in time (see Fig. 3). Treatment 4 shows how the group performance was affected by the time-dependent user equilibrium: Even without recommendations, the group managed to adapt to the changing conditions surprisingly well, but the standard deviation and changing rate were approximately as high as in treatment 2 (see Fig. 1). This adaptability (the collective 'group intelligence') is based on complementary responses (*14*). That is, if some players do not react to the changing conditions, others will take the chance to earn additional payoff. This experimentally supports the behaviour assumed in the theory of efficient markets, but here the efficiency is limited by overreaction.

In most experiments, we found a constant and high compliance $C_S(t) \approx 0.92$ with recommendations to stay, but the compliance $C_M(t)$ with recommendations to change (to 'move') (*5,10,11*) turned out to vary in time. It decreased with the reliability of the recommendations (see Fig. 3C), which again dropped with the compliance rate. Based on this knowledge, we have developed a model, how the competition for limited resources (such as road capacity) can be *optimally* guided by means of information services. Let us assume we had $N_1(t)$ 1-decisions at time t, but the optimal number of 1-decision at time $t+1$ is calculated to be $F_1^{\rm eq}(t+1)N \ge N_1(t)$. (In the case $\Delta N_1(t+1) = F_1^{\rm eq}(t+1)N - N_1(t) < 0$, indices 1 and 2 just need to be exchanged.) We want to exactly balance the difference $\Delta N_1(t+1)$ by transitions from decision 2 to decision 1. Let us assume we give recommendations to fractions $I_1(t)$ and $I_2(t)$ of players who had chosen decision 1 and 2, respectively. The fraction of changing recommendations to previous 1-choosers shall be denoted by $R_1(t)$, and for previous 2-choosers by $R_2(t)$. Correspondingly, fractions of $[1 - R_1(t)]$ and $[1 - R_2(t)]$ receive a recommendation to stick to the previous decision. Moreover, $[1 - C_M(t)]$ is the refusal rate of recommendations to change, while $[1 - C_S(t)]$ is the refusal rate of recommendations to stay. Finally, we denote the spontaneous transition probability from decision 1 to 2 by $p_a(2|1, N_1; t)$ and the inverse transition probability by $p_a(1|2, N_1; t)$, in case a player does not receive any recommendation. This happens with probabilities $[1 - I_1(t)]$ and $[1 - I_2(t)]$, respectively. Both transition probabilities are functions of the number $N_1(t) = N - N_2(t)$ of previous 1-decisions. The index a allows us to reflect different strategies of players (e.g. direct or contrary responses (*14*)). The fraction of players pursuing strategy a is then denoted by $F_a(t)$. Applying methods summarized in Ref. (*21*), the expected change of N_1 is given by the balance equation

$$\begin{aligned}\Delta N_1(t+1) = & \sum_a p_a(1|2,N_1;t)F_a(t)[1-I_2(t)]N_2(t) \\ & - \sum_a p_a(2|1,N_1;t)F_a(t)[1-I_1(t)]N_1(t) \\ & + \{C_M(t)R_2(t) + [1-C_S(t)][1-R_2(t)]\}I_2(t)N_2(t) \\ & - \{C_M(t)R_1(t) + [1-C_S(t)][1-R_1(t)]\}I_1(t)N_1(t)\,. \end{aligned} \quad (1)$$

We have evaluated the overall transition probabilities $p(1|2,N_1;t) = \sum_a p_a(1|2,N_1;t)F_a(t)$ and $p(2|1,N_1;t) = \sum_a p(2|1,N_1;t)F_a(t)$. According to classical decision theories (*21,22,23*), we would expect that the transition probability $p(2|1,N_1;t)$ should be a monotonic function of the payoff $P_2(N-N_1(t))$, the payoff difference $P_2(N-N_1(t)) - P_1(N_1(t))$, the potential payoff $P_2(N-N_1(t)+\epsilon N)$, or the potential payoff gain $P_2(N-N_1(t)+\epsilon N) - P_1(N_1(t))$. All these quantities vary linearly with N_1, so that $p(2|1,N_1;t)$ should be a monotonic function of $N_1(t)$. A similar thing should apply to $p(1|2,N_1;t)$. Instead, the experimental data point to transition probabilities with a *minimum* at the user equilibrium (see Fig. 2C). That is, the players stick to a certain alternative for a longer time, when the system is close to the user equilibrium. This is a result of learning. In fact, we find a gradual change of the transition probabilities in time (see Fig. 2D). The corresponding 'learning curves' reflect the players' adaptation to the user equilibrium.

After the experimental determination of the transition probabilities $p(2|1,N_1;t)$, $p(1|2,N_1;t)$ and specification of the compliance rates $C_M(t)$, $C_S(t)$, we can guide the decision behaviour in the system via the levels $I_i(t)$

of information dissemination and the fractions $R_i(t)$ of recommendations to change ($i \in \{1, 2\}$). These four degrees of freedom allow us to apply a variety of guidance strategies depending on the respective information medium. For example, a guidance by radio news is limited by the fact that $I_1(t) = I_2(t)$ is given by the average percentage of radio users. Therefore, equation (1) cannot always be solved by variation of the fractions of changing recommendations $R_i(t)$. User-specific services have the greatest guidance potentials. Among the different guidance strategies fulfilling equation (1), the one with the minimal statistical variance will be the best. When *individual* recommendations cannot be broadcasted to the users, it will help to inform everyone about the *fractions* $R_i(t)$ of participants who should change their decision, as users can learn to respond with varying probabilities (see Figs. 2C, D).

We suggest to determine the compliance rates $C_j(t)$ with $j \in \{M, S\}$ (and the transition probabilities) on-line with an exponential smoothing procedure according to $C_j(t+1) = \alpha C'_j(t) + (1-\alpha)C_j(t)$ with $\alpha \approx 0.1$, where $C'_j(t)$ is the percentage of players who have followed their recommendation at time t. As the average payoff for decision changes is normally lower than for staying with the previous decision (see Fig. 3B), a high compliance rate C_M is hard to achieve. That is, individuals who follow recommendations to change normally pay for reaching the user equilibrium (because of the overreaction of the system). Hence, there are no good preconditions to charge the players for recommendations, as we did in another treatment. Consequently, only a few players requested recommendations, which reduced their reliability, so that the overall performance of the system went down. We conclude that information services are most successful when user-specific recommendations are given to many participants and for free.

In this study, we have explored different ways of information presentation that allow us to guide user decisions in the spirit of an optimal distribution of limited resources. The least standard deviations from the user equilibrium could be reached by presenting the own payoff and the potential payoff, if he or she (or a certain fraction of players) had additionally chosen the other alternative. The observed intermittent decision behaviour is reminiscent of volatility clusters in stock markets, where the individuals also react to aggregate information. It results from the desire to reach above-average payoffs, combined with the immanent overreaction in the system. Payoff losses due to a volatile decision dynamics (e.g., excess travel times) can be minimized by user-specific recommendations. A similar strategy may also facilitate to stabilize economic markets and to increase average profits. Optimal guidance strategies to reach the user equilibrium follow directly from the balance equation for decision changes derived above. The quantification of the transition probabilities required a novel stochastic description of the decision behaviour, which was not just driven by the potential (gains in) payoffs. To understand the findings, we had to consider individual learning.

References and Notes

1. M. Schreckenberg, R. Selten (eds.) *Human Behaviour and Traffic Networks* (Springer, Berlin, 2002), to appear.

2. W. Barfield, T. Dingus, *Human Factors in Intelligent Transportation Systems* (Erlbaum, Mahwah, NJ, 1998).

3. M. Ben-Akiva, J. Bottom, M. S. Ramming, Route guidance and information systems. *Int. J. Syst. Contr. Engin.* **215**, 317–324 (2001).

4. Articles in *Route Guidance and Driver Information*, *IEE Conference Publications*, Vol. **472** (IEE, London, 2000).

5. Articles in *Transportation Research Record* **1725** (Transportation Research Board, National Research Council, Washington, D.C., 2000).

6. J. Wahle, A. Bazzan, F. Klügl, M. Schreckenberg, Decision dynamics in a traffic scenario. *Physica A* **287**, 669–681 (2000).

7. H. S. Mahmassani, R.-C. Jou, Transferring insights into commuter behavior dynamics from laboratory experiments to field surveys. *Transportation Research A* **34**, 243–260 (2000).

8. P. Bonsall, P. Firmin, M. Anderson, I. Palmer, P. Balmforth, Validating the results of a route choice simulator. *Transportation Research C* **5**, 371–387 (1997).

9. Y. Iida, T. Akiyama, T. Uchida, Experimental analysis of dynamic route choice behavior. *Transportation Research B* **26**, 17–32 (1992).

10. P. S.-T. Chen, K. K. Srinivasan, H. S. Mahmassani, Effect of information quality on compliance behavior of commuters under real-time traffic information. *Transportation Research Record* **1676**, 53–60 (1999).

11. R. D. Kühne, K. Langbein-Euchner, M. Hilliges, N. Koch, Evaluation of compliance rates and travel time calculation for automatic alternative route guidance systems on freeways. *Transportation Research Record* **1554**, 153–161 (1996).

12. R. Hall, Route choice and advanced traveler information systems on a capacitated and dynamic network. *Transportation Research C* **4**, 289–306 (1996).

13. A. Khattak, A. Polydoropoulou, M. Ben-Akiva, Modeling revealed and stated pretrip travel response to advanced traveler information systems. *Transportation Research Record* **1537**, 46–54 (1996).

14. M. Schreckenberg, R. Selten, T. Chamura, T. Pitz, J. Wahle, Experiments on day-to-day route choice (and references therein). e-print www.trafficforum.org/01080701, submitted to *Cooper@tive Tr@nsport@tion Dyn@mics* (2001).

15. R. N. Mantegna, H. E. Stanley, *Introduction to Econophysics: Correlations and Complexity in Finance* (Cambridge University, Cambridge, England, 1999).

16. S. Ghashghaie, W. Breymann, J. Peinke, P. Talkner, Y. Dodge, Turbulent cascades in foreign exchange markets. *Nature* **381**, 767–770 (1996).

17. M. Youssefmir, B. A. Huberman, T. Hogg , Bubbles and market crashes. *Comp. Econ.* **12**, 97–114 (1998).

18. T. Lux, M. Marchesi, Scaling and criticality in a stochastic multi-agent model of a financial market. *Nature* **397**, 498–500 (1999).

19. Articles by M. Ben-Akiva *et al.*, H. S. Mahmassani *et al.*, and R. Arnott *et al.* in *Transportation Research A* **25**(5), 251ff+293ff (1991).

20. J. Wahle, A. L. C. Bazzan, F. Klügl, M. Schreckenberg, Anticipatory traffic forecast using multi-agent techniques. In D. Helbing, H. J. Herrmann, M. Schreckenberg, D. E. Wolf (eds.) *Traffic and Granular Flow '99* (Springer, Berlin, 2000), pp. 87–92.

21. D. Helbing, *Quantitative Sociodynamics* (and references therein) (Kluwer Academic, Dordrecht, 1995).

22. M. Ben-Akiva, S. R. Lerman, Discrete Choice Analysis: Theory and Application to Travel Demand (MIT Press, Cambridge, MA, 1997).

23. M. Ben-Akiva, D. McFadden *et al.* Extended framework for modeling choice behavior. *Marketing Letters* **10**, 187–203 (1999).

FIGURES

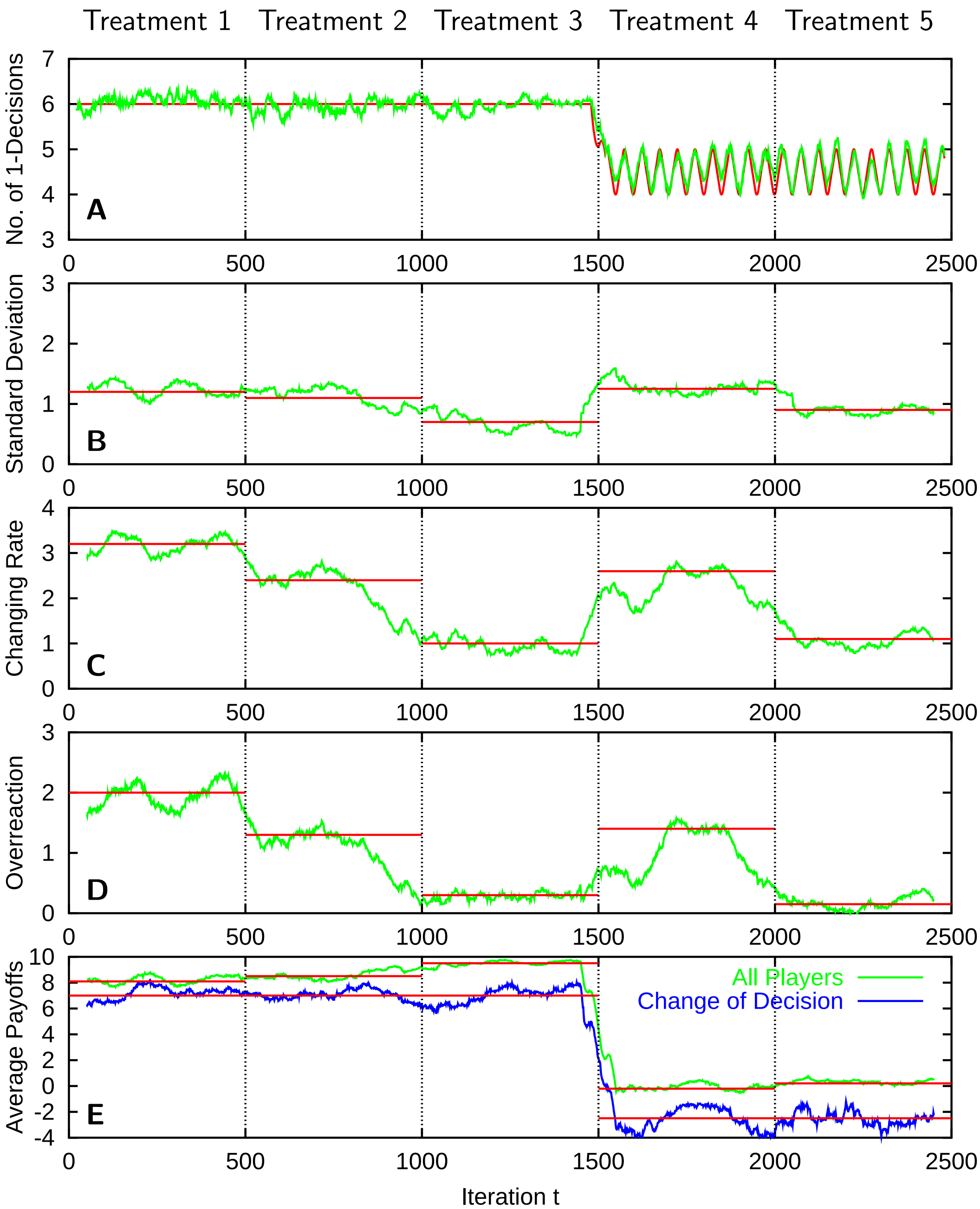
Treatment 1
Treatment 2
Treatment 3
Treatment 4
Treatment 5
No. of 1-Decisions
A
Standard Deviation
B
Changing Rate
C
Overreaction
D
Average Payoffs
E
All Players
Change of Decision
Iteration t
0
500
1000
1500
2000
2500

FIG. 1. Overview of treatments 1 to 5 (with $N = 9$ and payoff parameters $P_2^0 = 28$, $P_1^1 = 4$, $P_2^1 = 6$, and $P_1^0 = 34$ for $0 \leq t \leq 1500$, but a zick-zack-like variation between $P_1^0 = 44$ and $P_1^0 = -6$ with a period of 50 for $1501 \leq t \leq 2500$): **(A)** Average number of decisions for alternative 1 (green line) compared to the user equilibrium (red line), **(B)** standard deviation of the number of 1-decisions from the user equilibrium, **(C)** number of decision changes from one iteration to the next one, **(D)** overreaction, i.e., difference between the actual number of decision changes (changing rate) and the required one (standard deviation), **(E)** average payoff per iteration for players who have changed their decision and for all players. The latter increased with a reduction in the changing rate, but normally stayed below the payoff in the user equilibrium (which is 1 on average in treatments 4 and 5, otherwise 10). This payoff loss is caused by the overreaction in the system. The displayed moving time-averages [(A) over 40 iterations, (B)-(E) over 100 iterations] illustrate the systematic response to changes in the treatment every 500 iterations. Red lines in (B)-(E) are estimates of the stationary values after the transient period, while time periods around the dotted lines are not significant. Compared to treatment 1, treatment 3 managed to reduce the changing rate and to increase the average payoffs (even more than treatment 2 did). These changes were systematic for *all* players. In treatment 4, the changing rate and the standard deviation went up, since the user equilibrium changed in time. The user-specific recommendations in treatment 5 could almost fully compensate for this and managed to reach the minimum overreaction in the system. The above conclusions are also supported by additional experiments with single treatments.

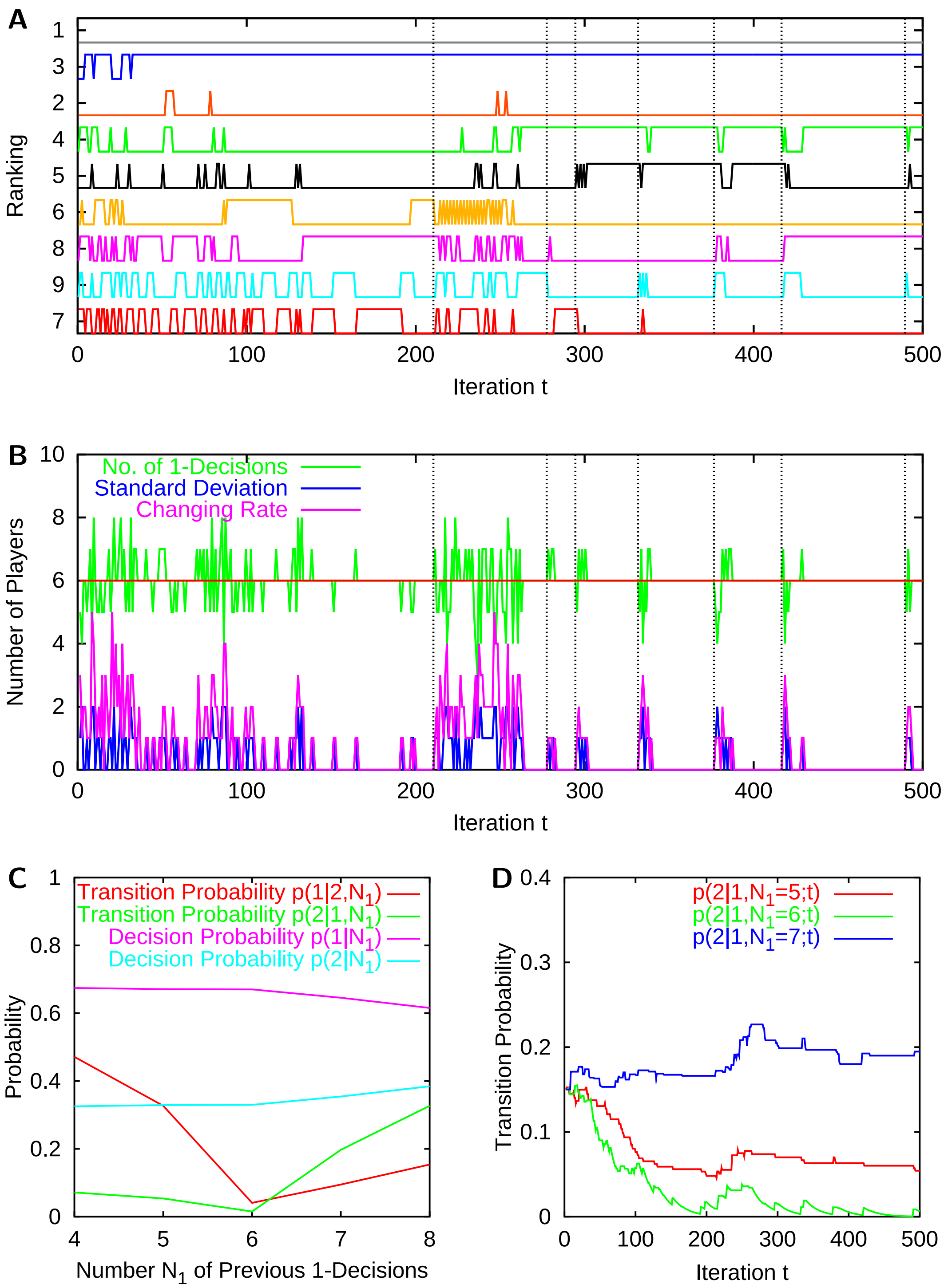
A
Ranking
1
3
2
4
5
6
8
9
7
Iteration t
B
Number of Players
No. of 1-Decisions
Standard Deviation
Changing Rate
Iteration t
C
Probability
Transition Probability p(1|2,N1)
Transition Probability p(2|1,N1)
Decision Probability p(1|N1)
Decision Probability p(2|N1)
Number N1 of Previous 1-Decisions
D
Transition Probability
p(2|1,N1=5;t)
p(2|1,N1=6;t)
p(2|1,N1=7;t)
Iteration t

FIG. 2. Illustration of typical results for treatment 3 (which was here the only treatment applied to the test persons, in contrast to Fig. 1). **(A)** Decisions of all 9 players. Players are displayed in the order of increasing changing rate. Although the ranking of the cumulative payoff and the changing rate are anticorrelated, the relation is not monotonic. Note that turbulent or volatile periods characterized by many decision changes are usually triggered by individual changes after quiet periods (dotted lines). **(B)** The changing rate is often larger than the (standard) deviation from the user equilibrium $N_1 = F_1^{\rm eq} N = 6$ (red line), indicating an overreaction in the system. **(C)** The probability $p(1|N_1)$ to choose alternative 1 was approximately 2/3, independently of the number N_1 of players who had previously chosen alternative 1. The probability $p(2|N_1)$ to choose alternative 2, given that N_1 players had chosen alternative 1, was always about 1/3. In contrast, the transition probability $p(1|2, N_1)$ describing decision changes from alternative 2 to 1 did depend on the number N_1 of players who had chosen decision 1. The same was true for the inverse transition probability $p(2|1, N_1)$ from decision 1 to decision 2. Remarkably enough, these transition rates are not monotonically increasing with with the payoff or the expected payoff gain, as they do not monotonically increase with N_1. Instead, the probability to change the decision shows a minimum at the user equilibrium $N_1 = F_1^{\rm eq} N = 6$. The reason for the different transition probabilities is an adaptation process in which the group players learn to take fewer changing decisions, when the user equilibrium is reached or close by, but more, when the user equilibrium is far away (see **(D)**: exponentially smoothed curves with $\alpha = 0.05$).

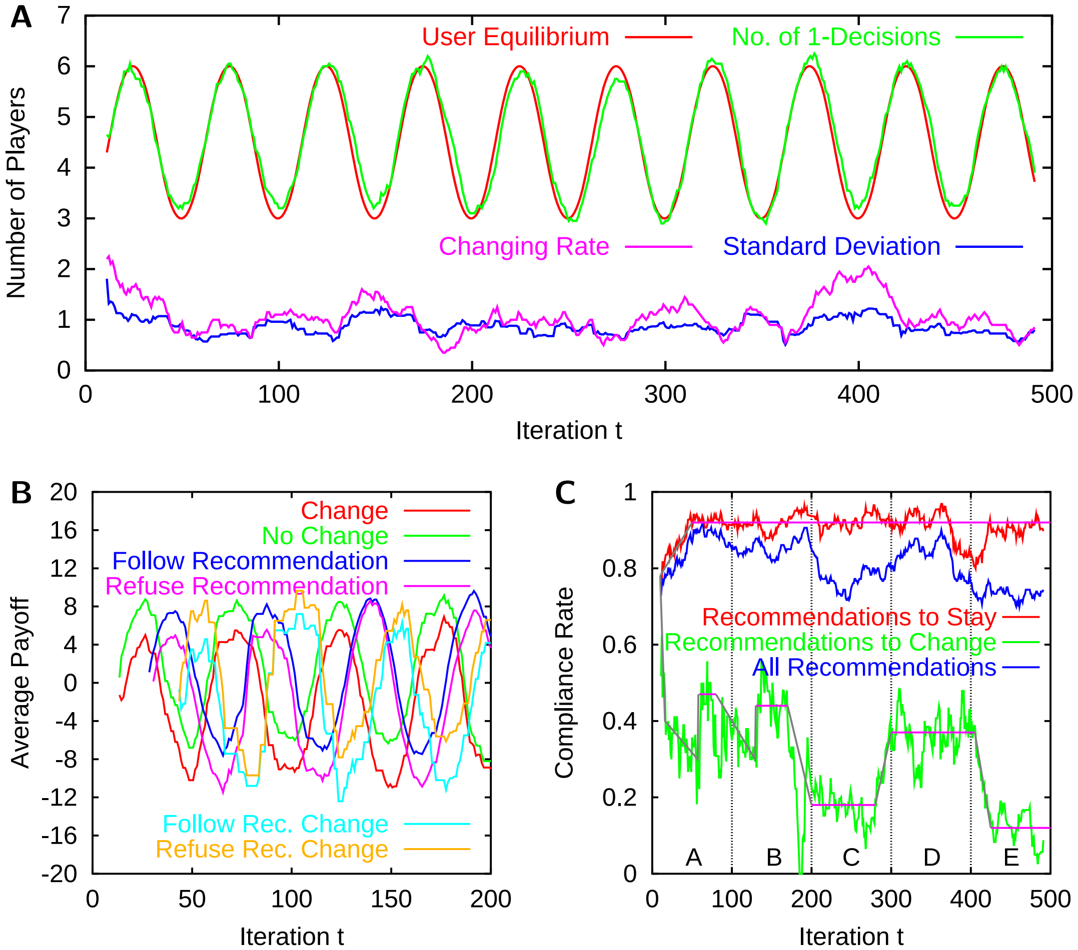
A
User Equilibrium
No. of 1-Decisions
Changing Rate
Standard Deviation
Number of Players
Iteration t
B
Change
No Change
Follow Recommendation
Refuse Recommendation
Follow Rec. Change
Refuse Rec. Change
Average Payoff
Iteration t
C
Recommendations to Stay
Recommendations to Change
All Recommendations
Compliance Rate
A
B
C
D
E
Iteration t

FIG. 3. Representative examples for treatments 4 and 5. The displayed curves are moving time-averages over 20 iterations. For illustrative reasons, blue and pink lines are shifted by 20, light blue and orange lines by 40. **(A)** Compared to treatment 4 (see Fig. 1), the user-specific recommendations in treatment 5 (assuming $C_M = C_S = 1$, $R_1 = 0$, $R_2 = \max([F_1^{\rm eq}(t+1)N + B(t+1) - N_1(t)]/N_2(t), 1)$, $I_1 = I_2 = 1$) could increase the group adaptability to the user equilibrium a lot, even if they had a systematic or random bias B, see (C). The standard deviation was reduced considerably and the changing rate even more. **(B)** The average payoffs varied largely with the decision behaviour. Players who changed their decision got much lower payoffs on average than those who kept their previous decision. Even recommendations could not overcome this difference: It stayed profitable not to change, although it was generally better to follow recommendations than to refuse them. **(C)** Consequenctly, the compliance to recommendations to change dropped considerably below the compliance to recommendations to stay. The compliance to changing recommendations was very sensitive to the degree of their reliability, i.e. participants followed recommendations just as much as they helped them to reach the user equilibrium. While during time interval I, the recommendations would have been perfect, if all players had followed them, in time interval II the user equilibrium was overestimated by $B = +1$, in III it was underestimated by $B = -2$, in IV it was randomly over- or underestimated by $B = \pm 1$, and in V by $B = \pm 2$. Obviously, a random error is more serious than a systematic one of the same amplitude. Pink lines illustrate the estimated stationary compliance levels, while transient periods are indicated by grey lines.