A behavioral study of "noise" in coordination games

Maes, Michael; Nax, Heinrich H.

Published in:
Journal of Economic Theory

DOI:
10.1016/j.jet.2015.12.010

Publication date:
2016

Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 162 (2016) 195–208

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

A behavioral study of "noise" in coordination games ☆

Michael Mäs [a], Heinrich H. Nax [b,*]

[a] *Sociology & ICS, University of Groningen, Grote Rozenstraat 31, 9712 TG Groningen, The Netherlands*
[b] *Social Sciences, ETH Zürich, Clausiusstrasse 37, 8092 Zürich, Switzerland*

Received 27 July 2015; final version received 16 December 2015; accepted 23 December 2015

Available online 31 December 2015

Abstract

'Noise' in this study, in the sense of evolutionary game theory, refers to deviations from prevailing behavioral rules. Analyzing data from a laboratory experiment on coordination in networks, we tested 'what kind of noise' is supported by behavioral evidence. This empirical analysis complements a growing theoretical literature on 'how noise matters' for equilibrium selection. We find that the vast majority of decisions (96%) constitute myopic best responses, but deviations continue to occur with probabilities that are sensitive to their costs, that is, less frequent when implying larger payoff losses relative to the myopic best response. In addition, deviation rates vary with patterns of realized payoffs that are related to trial-and-error behavior. While there is little evidence that deviations are clustered in time or space, there is evidence of individual heterogeneity.

JEL classification: C73; C91; C92

Keywords: Behavioral game theory; Discrete choice; Evolution; Learning; Logit response; Stochastic stability; Trial-and-error

☆ We acknowledge support by the European Commission through the ERC Advanced Investigator Grant "Momentum" (Grant 324247). We thank Peyton Young, Erik Mohlin, Tom Norman, Bary Pradelski, Dirk Helbing, as well as seminar participants at the 25th Game Theory Festival (Stony Brook), at the Norms and Networks Cluster at the University of Groningen (ICS), at the LGN Discussion Group (Nuffield College), at the Micro Theory Workshop (Cambridge-INET Institute), and an associate editor and two anonymous referees for very helpful comments.

* Corresponding author.

http://dx.doi.org/10.1016/j.jet.2015.12.010

1. Introduction

Individuals occasionally deviate from their prevailing behavioral rules because of, for instance, mistakes, misperceptions, inertia, or trial-and-error experiments. Evolutionary game theory demonstrates that the exact nature of such individual-level deviations can crucially influence equilibrium selection (Kandori et al., 1993; Young, 1993, 1998, 2011a; Kandori and Rob, 1995; Blume, 2003; Montaneri and Saberi, 2010; Sandholm, 2010; Newton, 2012a; Bergin and Lipman, 1996). In this paper, we therefore tested competing assumptions regarding the nature of deviations, analyzing data from a laboratory experiment on coordination games played in fixed networks. Using classical discrete-choice estimation techniques (McFadden, 1974), we tested, in particular, the assumptions that deviations (*i*) occur with a constant probability (Kandori et al., 1993; Young, 1993), (*ii*) depend on the costs of deviating (Blume, 1993), and (*iii*) vary with payoff patterns related to trial-and-error behavior (Young, 2009; Pradelski and Young, 2012).

To study deviations, it is fundamental to justify the behavioral model relative to which a decision is considered a deviation. Most evolutionary game theory models focus on variants on myopic best-response (MBR) behavior, which assumes that agents maximize their individual payoffs in the current period by best-responding to others' actions as previously sampled. The concept of myopic best response dates back to the Nash equilibrium (Nash, 1950b; Young, 2011b) but the section on myopic best-response dynamics from Nash's PhD thesis (Nash, 1950a) was unfortunately omitted in its published version (Nash, 1951). Based on our experimental data, statistical tests revealed that MBR (confined to a minimal memory length of one) accurately describes 96% of subjects' decisions, leading us to define deviations as the remaining decisions. It is noteworthy that best-response models with longer memories and also simple models of reinforcement learning (Bush and Mosteller, 1955; Suppes and Atkinson, 1959; Harley, 1981; Cross, 1983; Roth and Erev, 1995; Erev and Roth, 1998) make virtually identical predictions in our experiment. Hence, while we cannot be certain as to which precise underlying decision rule the subjects applied, we can be certain that decisions identified as deviations indeed deviated from the decision rule, whichever was applied by the subject.

Alternative terminologies for deviations from an underlying rule are "noise", "errors", "trembles", "experiments" or "mistakes" (Kandori et al., 1993; Young, 1993). We shall prefer the terminology of "deviation", a more neutral word with regard to causality. There is a related but separate literature on noise in static situations of stochastic choice under risk (for example, Wilcox, 2008; Butler et al., 2012). These contributions study the effects of random perturbations of utility functions, rather than deviations from dynamic strategy protocols. Related static notions are the trembling-hand perfect (Selten, 1975), quantal response (McKelvey and Palfrey, 1995, 1998) and proper (Myerson, 1978) equilibrium concepts, which are static analogues of various dynamics. Our paper complements this line of research with the study of evolutionary dynamics.

2. Competing deviation assumptions

We tested three competing deviation assumptions as used in theoretical models. First, we tested the assumption that deviations occur with constant probability. That is, agents play, for instance, MBR most of the time but occasionally deviate with some constant rate (Kandori et al., 1993; Young, 1993).[1] In two-by-two coordination games, MBR dynamics that are perturbed by constant errors select, in the sense of "stochastic stability" (Foster and Young, 1990), the risk-dominant (Harsanyi and Selten, 1988) Nash equilibrium as errors vanish (Young, 1993; Blume, 1996; Peski, 2010).

Second, we tested whether deviation rates are decreasing in their costliness vis-à-vis the prevalent choice. The most prominent implementation of this assumption is the *logit response* model (Blume, 1993). Except under certain symmetry/regularity conditions (Blume, 2003), cost-sensitive deviations have been shown to imply different convergence predictions than constant errors (Young, 1998; Blume, 2003; Myatt and Wallace, 2004; Alós-Ferrer and Netzer, 2010), particularly with regard to convergence times (Ellison, 1993; Young, 1998, 2011a).[2]

A third set of deviation assumptions stems from payoff-based learning models (Hart and Mas-Colell, 2003, 2006; Foster and Young, 2006; Young, 2009), which share the assumption that agents adjust their behavior based on trial-and-error heuristics rather than best-response considerations. Consequently, deviations are assumed to depend on agents' past experiences. Deviation rates have been assumed to increase when players enter an experimental state of mind, which may be due to payoff losses or recent action changes (Young, 2009; Pradelski and Young, 2012). Related trial-and-error dynamics (Thuijsman et al., 1995; Nax et al., 2013) have been shown to select Nash equilibria (Young, 2009), even welfare-maximizing Nash equilibria (Pradelski and Young, 2012).[3]

Although the empirical literature on the three deviation assumptions is very limited, there are related strands of literature. There is empirical research on evolutionary dynamics related to replicator dynamics (Taylor and Jonker, 1978; Weibull, 1995). Examples include the evolutionary interpretation by Crawford (1991) of the seminal study by Van Huyck et al. (1990), as well as more recent experiments by Friedman and Ostrov (2013), Xu et al. (2013), Cason et al. (2014), Hoffman et al. (2015). These studies show that replication makes macro-predictions that correspond to observed system behavior such as cyclical patterns in rock-paper-scissor games. However, in contrast to our study, these studies test macro-predictions that follow from competing deviation assumptions and were not designed to directly test the behavioral assumptions about micro-level deviations.[4]

The empirical study that comes closest to ours is a recent experiment on two-by-two coordination games under random re-matching by Lim and Neary (2014). These games have the advantage that most of the theoretical literature, beginning with Kandori et al. (1993) and Young (1993), addresses related macro-effects. However, two-by-two coordination games imply little variance in the deviation costs that players face, which makes it difficult, if not impossible, to

[1] The subtle difference between Kandori et al. (1993) and Young (1993) is that of mistakes from "imitation" (Kandori et al., 1993) versus MBR (Young, 1993; Kandori and Rob, 1995). We focus on the latter, because our subjects were not informed about others' payoffs, which precluded imitation.

[2] Generalizations with coalitional recontracting have been proposed by Newton (2012a) (constant errors) and Newton (2012b), Sawa (2014) (logit response).

[3] A dynamic is "uncoupled" when not dependent on information about others' utility functions (Hart and Mas-Colell, 2003), strengthened to "completely uncoupled" when only information about own actions and own payoffs is considered (Foster and Young, 2006). Uncoupled dynamics, in general, do not lead to Nash equilibrium (Hart and Mas-Colell, 2003, 2006). However, under certain genericity conditions, many completely uncoupled dynamics do select Nash equilibrium (Foster and Young, 2006; Germano and Lugosi, 2007; Marden et al. 2009, 2014; Pradelski and Young, 2012; Skyrms, 2012; Huttegger and Skyrms, 2012; Huttegger et al., 2014).

[4] A companion paper (Maes and Helbing, 2014) uses our data to test competing macro predictions of deterministic and noisy micro-models.

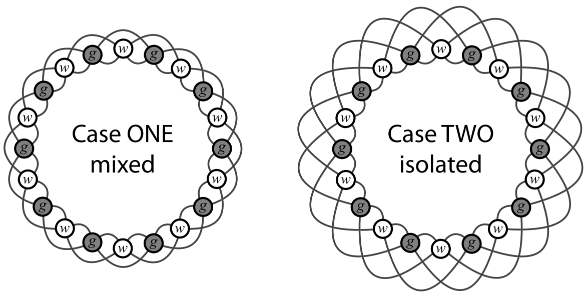

Fig. 1. Network types: Illustration of the two network structures studied in the experiment. Letters and colors depict the option for which the assigned participant earned payoff P. In the left-hand-side network nodes have mixed neighborhoods (Case ONE), whereas nodes are isolated in the right-hand-side network (Case TWO).

test individual-level phenomena such as cost-sensitive deviation rates. In our experiment, in contrast, subjects played coordination games with multiple network partners à la Ellison (1993). Thus, both the nature and "cost" of deviations varied depending on the proportion of interaction partners taking either action, allowing us to test these features.

3. Experiment

We analyzed data from an experiment on coordination in fixed networks that was conducted at ETH's Decision Science Laboratory (https://www.descil.ethz.ch). For each of the 13 experimental sessions, we invited 20 subjects who played 150-times repeated coordination games, yielding a total of 39,000 decisions from 260 subjects.

At the beginning of the experiment, subjects were assigned random positions in a circle network with either so-called "mixed" (Case ONE) or "isolated" (Case TWO) neighborhoods (see Fig. 1). Both network types were perfectly symmetric as in Ellison (1993), and every node was linked to four neighbors, two on each side. The position on the circle determined the node's type, which we denote here as w (white) and g (grey). Subjects retained their position and type for the whole experiment. In mixed neighborhoods, nodes were linked to the four closest w-nodes and g-nodes. In the isolated neighborhoods, nodes were linked to the four closest nodes of the other type; hence all w-nodes were linked to four g-nodes and vice versa.

Subjects played 150 rounds of the following game (see Section 2 of online appendix for instructions). In each round, subjects had to choose between two options, white and grey.[5] Subjects on a white node earned a payoff of P money units (MU) for choosing option white, and likewise for grey subjects choosing grey. Furthermore, subjects earned an additional MU for each network neighbor who chose the same action as them.

Subjects were not informed about the structure of the network or types of their network neighbors. In fact, the instructions did not even mention the existence of different types. In each round, subjects were informed about the choices of their interaction partners in the previous round and

[5] In the actual experiment, actions were labeled "yellow" and "blue". We use different labels here only for graphical reasons.

Table 1
Payoffs for an agent of type w as a function of others' actions.

$\#^W_{-i}(a_{-i})$	0	1	2	3	4
$\phi_i(W; a_{-i})$	P	$P+1$	$P+2$	$P+3$	$P+4$
$\phi_i(G; a_{-i})$	4	3	2	1	0

about their own payoff.[6] Importantly, subjects received no information about the payoffs of their neighbors.

The experiment comprised four treatments. Although we do not test here any hypotheses about treatment effects on deviation patterns, these experimental manipulations are important for the present study, because they generated variance with respect to the payoff differences between the two actions and the costs from not choosing the MBR. The four experimental treatments resulted from the manipulation of two variables. First, there were two types of networks; with mixed or isolated neighborhoods (see Fig. 1). Second, parameter P, that is, the payoff that subjects received for choosing the option that corresponded to their type, was either one or three providing different incentives to play the own type action.

Subjects knew that the experiment lasted roughly one hour and that the game was repeated. However, they did not know precisely how often the game was going to be repeated. At the end of the experiment, the computer program randomly picked three of the 150 periods and paid all subjects of the session according to their payoffs in these three rounds. On average, subjects earned 33 Swiss Francs (>30 US$) for spending just over one hour in the lab.

3.1. Subjects' decision problem

The decision problem that each subject faced was fully determined by the payoff rules and the color choices of the four network neighbors. More formally, each subject i, of type $\Theta \in \{g, w\}$ chose between actions $A \in \{G, W\}$, representing white and grey, respectively. w-nodes received payoff P for choosing action W, and g-nodes had the same preference for action G. W.l.o.g. suppose that i is of type w. Denote by a_{-i} an action profile excluding i and by $\#^W_{-i}(a_{-i})$ the number of neighbors of i choosing action W under a_{-i}.[7] Then, i's payoff, $\phi_i(A; a_{-i})$, depends on his choice of action A and on a_{-i} via $\#^W_{-i}(a_{-i})$ as specified in Table 1.

The best reply depends (i) on P, the 'bonus' associated with playing the color corresponding to one's own type, and (ii) on $\#^W_{-i}(a_{-i})$, the number of neighbors choosing action W under a_{-i}. For parameter values of $P = 1$ or $= 3$, subjects were never indifferent between the two actions, and neither one was a dominant strategy. Supposing (w.l.o.g.) that i is of type w, best replies summarize as follows. Given $P = 1$, i's best reply is W if two or more neighbors play W ($\#^W_{-i}(a_{-i}) \geq 2$), but three or four neighbors playing G ($\#^W_{-i}(a_{-i}) < 2$) switches i's best reply to G. Given $P = 3$, W is the best reply for i if at least one neighbor plays W ($\#^W_{-i}(a_{-i}) \geq 1$), and G is the best reply only if all four neighbors play G ($\#^W_{-i}(a_{-i}) = 0$).

Section 1 of the online appendix summarizes the structure of the stage-game Nash equilibria of the different conditions and games.

6 This information was provided in the form of colored boxes. The location of a box corresponding to a given interaction partner was chosen at random but did not change. See Section 2 of the online appendix for more detail.

7 For any feasible a_{-i}, $\#^W_{-i}(a_{-i})$ can be either 0, 1, 2, 3, or 4, with the number of choices of G, $\#^G_{-i}(a_{-i})$, being equal to $4 - \#^W_{-i}(a_{-i})$.

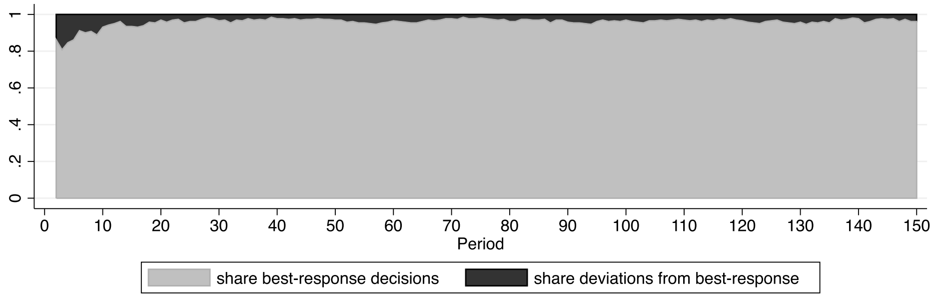

Fig. 2. Share of best responses and deviations over time.

4. Results

4.1. Test of the behavioral model

During the experiment, subjects were always provided with the information needed to choose the MBR. Information regarding others' payoffs or payoff functions was not provided in order to preclude imitation. Likewise, we did not provide any information about the network structure and the number of interaction periods, which made it virtually impossible to apply belief-based decision principles based on expectations about the behavior of others that are more complex than the best-response model.

Accordingly, it is not surprising that 96% of subjects' decisions were MBR (similar deviation rates were observed by Lim and Neary, 2014). 93 subjects (35.8%) never deviated from the MBR model. Another 47 (18.1%) subjects deviated exactly once, and 27 subjects (10.4%) deviated twice. Only 25 subjects (9.6%) had a deviation rate that exceeded 10 percent. Fig. 2 shows the evolution of the share of decisions that were the best response and the share of deviations in all sessions of the experiment. In Period 2, the deviation rate was 0.14. By Period 25 the deviation rate had dropped to .03. Overall, the deviation rate was only .04.

Despite the overwhelming support for the MBR model, it cannot be excluded that subjects formed decisions based on alternative, more complex choice rules. We tested two alternative decision models (for details see Section 3 of the online appendix). First, we tested whether subjects did not only respond to others' behavior in the previous period but also considered earlier rounds. Our results indicate that subjects did look back up to three rounds more than the MBR model assumes and that the relative impact of a given round decreased over time (see Figure A4 of online appendix). However, when we compared the choices that follow from the MBR rule and calibrated best-response models with longer memories, we found that the models make virtually the same predictions (see Tables A3 and A4 of online appendix). In other words, given the setting of our experiment and the behavior that subjects observed in their interaction partners, an actor that applied the MBR rule would have made virtually identical choices to an actor who responded to longer memories. As a consequence, a choice that deviates from the MBR is also a deviation from best-response rules with longer memory. Figure A5 of the online appendix furthermore shows that also the deviation costs of the alternative best-response models were very highly correlated.

Our tests of reinforcement-learning models led to a similar outcome. While our statistical analyses provided support for reinforcement learning (see Table A6 of online appendix), a simple learning model that is supported by our tests makes virtually identical predictions as the MBR model (see Table A7 of online appendix), making it impossible to distinguish between MBR and the learning model. Interestingly, however, there were about 600 decisions where learners would have been indifferent between the two behavioral options. It turned out that 86 percent of these decisions were myopic best-responses, which supports the MBR model.

In sum, although 96% of all observed decisions were MBR, our experiment does not allow to exclude that subjects used alternative decision rules. However, in our experiment the different decision models make virtually identical predictions about individuals' decisions. Therefore, even if (some) subjects used more complex best-response rules or a learning heuristic, we can be certain that decisions identified as deviations indeed deviated from the decision rule applied by the subject.

Nevertheless, we replicated our analyses by assuming best-response models with longer memory and a simple reinforcement learning model, rather than the MBR model. Section 3 of the online appendix shows that these analyses lead to the same conclusions (see Tables A5 and A9 of online appendix).

4.2. Statistical approach; discrete-choice estimation

In order to test competing deviation models, we estimated discrete-choice models with logistic regressions. The dependent variable adopted the value zero when a subject played MBR and the value one whenever a decision deviated. Given a^t_{-i}, the actions taken by actor i's network neighbors, we defined b_i^{t+1} as the *best response* and n_i^{t+1} as a *deviation*. The experimental design precluded situations where subjects were indifferent between the two actions. Period 1 was not included in the analyses, because necessary information for MBR was not available. For all regression results, we report robust standard errors clustered on subjects, to account for the nestedness of decisions in subjects. In addition, all results could be replicated also with random-intercept multi-level models (see Table A11 of the online appendix).

We included five independent variables to test the competing deviation assumptions. Finding a significant effect of any of these variables would challenge the assumption that deviations occur with a constant probability. As the first independent variable we included deviation costs, $C(n_i^{t+1})$, testing the assumption of cost-sensitivity of deviations (as in the logit-response model). The deviation cost is the difference between the expected payoff of the best response and the expected payoff of a deviation, taking period-t actions of others as given. Formally, $C(n_i^{t+1}) = \phi^{t+1}(b_i^{t+1}; a^t_{-i}) - \phi^{t+1}(n_i^{t+1}; a^t_{-i})$. A negative and significant coefficient would support the assumption that deviations occur with higher probabilities when they imply lower costs.

To test the deviation assumptions of learning models, we included four dummy variables that capture all relevant additional information from the previous period available to the subject that is not contained in the deviation cost. The first dummy, labeled "payoff increase $t-1$", adopts the value one when the subject experienced a prior payoff increase ($\phi_i^t > \phi_i^{t-1}$). According to theory, we expected a negative effect, which would show that deviations occur less frequently after subjects had experienced increasing payoffs. Second, the dummy labeled "payoff decrease $t-1$" is set to one whenever the subject experienced decreasing payoff in the previous period ($\phi_i^t < \phi_i^{t-1}$). We expected more deviations after a payoff decrease. Third, dummy variable "neighbor(s) switched $t-1$" adopts the value one when at least one of the four neighbors of i had previously changed her action ($a^t_{-i} \neq a^{t-1}_{-i}$). The fourth dummy, labeled "actor switched $t-1$", adopts the value one when the subject had previously changed her own action ($a_i^t \neq a_i^{t-1}$). We included dummies three and four to test whether deviation rates depend on information about others when

Table 2
Test of competing models of deviations from the best-reply rule.

	Constant error	Logit error	Full model	Final model	Control variables
Constant	−3.164		−0.934	−0.937	−1.080
	(0.185)**		(0.399)*	(0.376)*	(0.272)**
Deviation cost		−1.498	−1.122	−1.122	−0.928
		(0.121)**	(0.153)**	(0.150)**	(0.132)**
Payoff increase $t-1$			−0.111		
			(0.664)		
Payoff decrease $t-1$			0.706	0.745	1.088
			(0.197)**	(0.295)*	(0.318)**
Neighbor(s) switched $t-1$			0.022		
			(0.244)		
Actor switched $t-1$			1.985	1.928	1.906
			(0.177)**	(0.378)**	(0.426)**
Control variables					
Period t					−0.001
					(0.002)
subject type Θ					0.432
					(0.358)
MBR was color of type					−1.253
					(0.332)**
Log pseudolikelihood	−6570.85	−5262.16	−4865.20	−4865.49	−4693.30
Number decisions	38,740	38,740	38,740	38,740	38,740
Number subjects	260	260	260	260	260

* $p < 0.05$; ** $p < 0.01$; robust standard errors in parentheses.

it is available (dummy three), and whether changing one's own action incites a certain deviation proneness (dummy four). Note that the definitions of the dummies imply that all four dummies adopt the value zero for all decisions before period three.

4.3. Test of the deviation assumptions

Table 2 summarizes the analyses. The baseline regression model contained only the constant. The estimated constant of −3.16 translates into an overall deviation rate of 4%. In absolute numbers, 1570 of the 38,740 decisions deviated from the MBR principle. The second regression model contained only the deviation costs. In line with logit response, the coefficient is significant and negative. Thus, deviations occurred less often when they implied higher costs.

In the full model, we included the constant, the deviation costs, and the four trial-and-error dummies. The effect of the deviation costs remained negative and significant, providing further support for the assumption of cost-sensitive deviations. Furthermore, deviation rates were significantly lower after payoff increases than after payoff decreases. However, while payoff rates were not significantly affected when subjects observed other players' action changes, there was a strong and significant increase in deviations when subjects had themselves changed action.

The final model of Table 2 resulted from dropping the insignificant variables from the full model in a stepwise process. It supports cost-sensitive deviation models and the two trial-and-error components payoff decrease $t-1$, and actor switched $t-1$. Furthermore, the constants of the full model and the final model have a negative and significant parameter estimate. The nega-

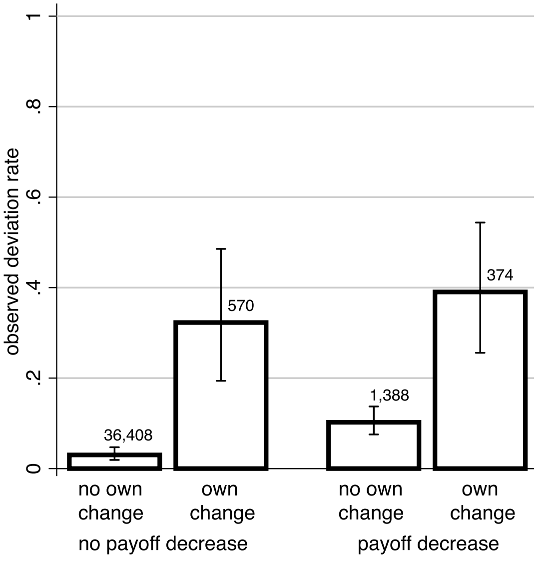

Fig. 3. Trial-and-error components: Observed deviation rates depending on whether or not subject changed color in previous round and whether or not subject's payoff decreased in the previous period. Error bars show confidence intervals with robust standard errors clustered on subjects. Numbers next to bars indicate how many decisions were made in the respective setting.

tive parameter estimate suggests that a model based only on the remaining independent variables overestimates the deviation rate, particularly when deviation costs are very low.

To facilitate the interpretation of our logistic regressions, Figs. 3 and 4 illustrate the deviation rates that follow from the significant independent variables of the final model of Table 2. First, the bar graph in Fig. 3 reports observed (rather than estimated) deviation rates depending on whether subjects had experienced a payoff decrease and whether subjects had changed action. The numbers next to the bars indicate how often a subject faced the respective situation.[8] When subjects had not changed action and had not experienced a payoff decrease, the observed deviation rate was 3%. However, the deviation rate adopted a significantly higher value of 32% when subjects had changed action in the previous period. Likewise, deviation rates increased significantly when subjects had experienced a payoff decrease. For instance, when subjects had not changed action but experienced a payoff decrease, subjects deviated from the best-response rule in 10% of the 1388 cases, which is also significantly above the deviation rate when subjects' payoffs had not decreased.

Obviously, the effects shown in Fig. 3 challenge both the assumption of constant deviations and of models where deviations depend only on payoff costs (as in the logit-response model), as both models neglect that deviation rates might (strongly) depend on trial-and-error patterns. However, Fig. 4 shows that the final model of Table 2, which is a generalized cost-dependency

[8] In total, subjects faced 2922 times a decision with deviation costs of seven MU. In all but one of these decisions, subjects had not changed action and had not experienced a payoff decrease in the previous period. We did not observe a single deviation from the BR rule when facing costs of seven MU. We therefore exclude cases where deviation costs summed up to seven MU. However, the observation that there were no deviations when they implied high costs challenges the constant-deviation model, and provides support for models where deviation rates decrease with payoff costs.

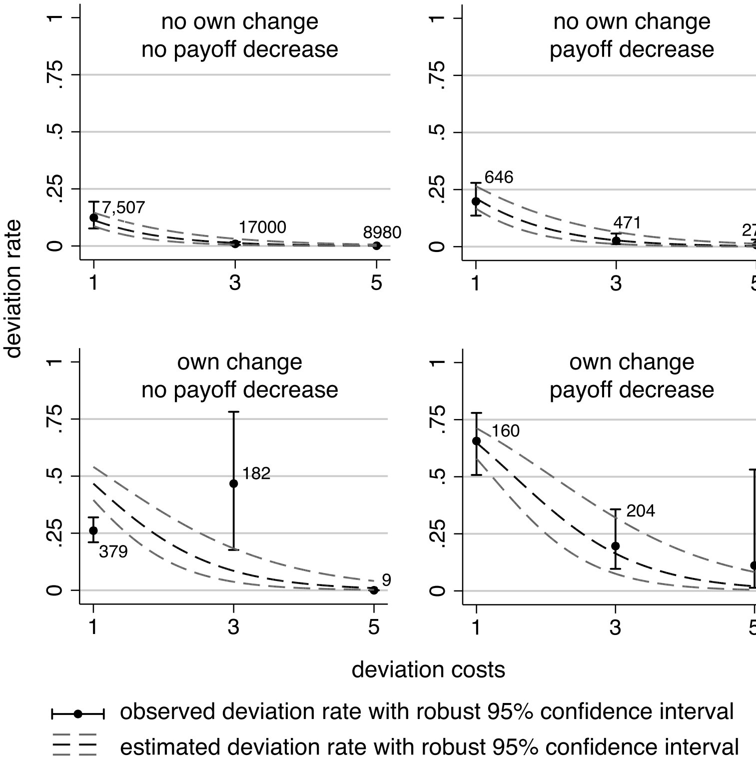

Fig. 4. Deviation cost and trial-and-error: Observed and predicted deviation rates as functions of deviation costs, categorized by whether subjects had experienced a payoff increase and/or had switched action in the previous period. Numbers next to bars indicate how many decisions were made in the respective setting.

(here, logit-response) model that also contains a constant and different deviation rates under the four scenarios presented by payoff and action changes, fits the data very well. In Fig. 4, the dots show the observed deviation rates for the four cases that were also compared in Fig. 3 categorized by the deviation costs. The dark dashed lines show the estimated deviation rates that follow from the final model of Table 2. Fig. 4 demonstrates that deviation rates always depended on deviation costs, independent of whether a subject had changed action or experienced a payoff drop. Moreover, past payoff losses and action changes unambiguously lead to higher deviation rates.

4.4. Robustness tests

As the first robustness test, we added three control variables to the final model of Table 2: the type Θ of the subject (w or g), the decision period t, and a dummy measuring whether (coded 1) or not (coded 0) the MBR was the same color as the type of the subject and, thus, resulted in payoff P. Demonstrating the robustness of our findings, including the three control variables did not affect the findings from the final model. Subject type Θ and period t did not have significant effects, but deviation rates were significantly lower when subjects faced a decision where the MBR was the subject's preferred option, which lends support to the assumption that deviations can be directed (Naidu et al., 2010; Hwang and Newton, 2014). When the MBR was the subject's preferred option, the overall deviation rate was 1%. Otherwise, subjects deviate with a rate of about 9%.

The period effect from the control-variable model from Table 2 is small and insignificant, which shows that the changes in deviation rates over time depicted in Fig. 2 are explained by the independent variables of the model and the evolutionary dynamics they imply (related to stochastic stability, Foster and Young, 1990). For example, decreasing deviation rates result from decreases in the exposure to situations with low deviation costs. Over time, the outcomes of the game imply higher deviation costs and fewer changes in own actions, which is precisely the prediction of stochastic stability under, for instance, logit response.

We tested the robustness of our findings across sessions, time, and subjects (see online appendix). First, the effects found in the final model of Table 2 were found also when models were estimated separately for each of the 13 experimental sessions (online appendix, Figure A7). Second, we could replicate the results of the final model also with separate models for subsets of 10 consecutive periods (Figure A8 of online appendix). However, these analyses suggest that the effect of payoff losses obtained manly during the first half of the experiment. Third, we replicated the final model of Table 2, estimating random-effects models with random intercepts on the level of subjects and found the same effects (Table A11 of online appendix).

As an additional illustration of the robustness of our findings, the scatter plots in Fig. 5 compare each subject's overall deviation rate with the deviation rate in those periods that followed a payoff loss and/or a change in action. In this figure, each cross informs about one subject's overall deviation rate during the experiment (x-axis) and her rate of deviation in the respective setting (y-axis). The figure does not show those subjects in the respective plot that never faced the respective situation, and also not those subjects (of which there were 93) with an overall deviation rate of zero.

If deviations were adequately described by the constant-deviation model, the crosses should be placed on or very close to the cross-diagonal of the scatter plots. However, Fig. 5 clearly shows that this was not the case. The top-left panel compares the overall deviation rates with the deviation rates in situations where the subject did not change action in the previous period and did not experience a payoff decrease. 94% of all decisions were made in this situation. Accordingly, it is not surprising that the overall deviation rates "match" the deviation rates in this specific setting (i.e. lie close to the cross-diagonal). Nevertheless, the average difference between subjects' overall deviation rates and their deviation rates in these situations is positive ($t = 12.69$; $p < 0.01$), indicating that the deviation rates in this situation were, on average, smaller than the overall deviation rates. This is in line with the trial-and-error components. For the other three situations in Fig. 5, we see that many subjects deviated substantially more often. When subjects had stuck to their previous action but experienced a decreasing payoff (see top-right panel of Fig. 5), the average difference between the overall deviation rate and the deviation rate was -0.19 ($t = -7.47$; $p < 0.01$). Likewise, the average difference in deviation rates was -0.17 ($t = -6.08$; $p < 0.01$) when subjects had changed action but did not experience a payoff decrease in the previous period (see bottom-left panel).

The biggest difference in deviation rates was observed for the case when subjects had changed action in the previous period and had experienced a payoff drop. In this situation, deviation rates were on average 37% higher than the overall deviation rate ($t = -9.66$; $p < 0.01$). In particular, the large number of crosses in the top left part of the plots draws attention. These subjects have positive but small overall deviation rates, which shows that they did not deviate very often. However, when they had changed action in the previous period and had experienced a payoff drop, they deviated with a very high rate or even with probability one.

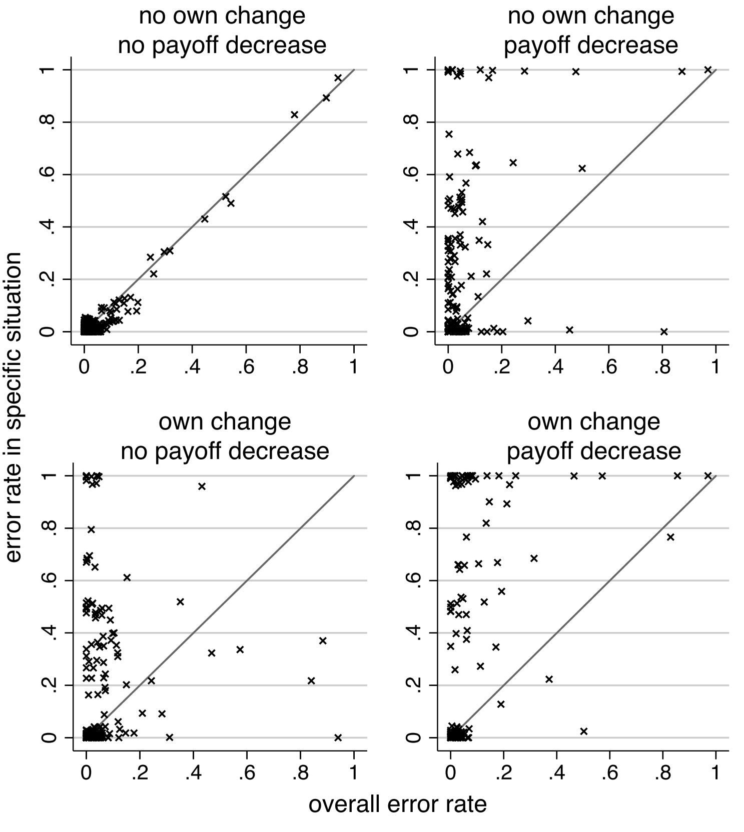

Fig. 5. Individual heterogeneity: Comparison of subjects' overall deviation rates and their individual rates when they had experienced a payoff drop in the previous round and/or changed action. Crosses below the diagonal show that a subject deviated less likely in the given setting than in the whole experiment. Crosses above the diagonal depict increased deviation rates.

5. Summary

Our analyses lead to two main results. First, we found that deviations were more frequent when they implied smaller deviation costs. Second, we found significantly increased deviation rates when subjects had changed action and/or experienced a payoff decrease in the previous period.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.jet.2015.12.010.

References

Alós-Ferrer, C., Netzer, N., 2010. The logit-response dynamics. Games Econ. Behav. 68, 413–427.
Bergin, J., Lipman, B.L., 1996. Evolution with state-dependent mutations. Econometrica 64, 943–956.
Blume, L.E., 1993. The statistical mechanics of strategic interaction. Games Econ. Behav. 5, 387–424.

Blume, L.E., 1996. Population games. Working paper no. 04 (022). Santa Fe Institute.
Blume, L.E., 2003. How noise matters. Games Econ. Behav. 44, 251–271.
Bush, R., Mosteller, F., 1955. Stochastic Models of Learning. Wiley.
Butler, D., Isoni, A., Loomes, G., 2012. Testing the 'standard' model of stochastic choice under risk. J. Risk Uncertain. 45, 191–213.
Cason, T., Friedman, D., Hopkins, E., 2014. Cycles and instability in a rock–paper–scissors population game: a continuous time experiment. Rev. Econ. Stud. 81, 112–136.
Crawford, V.P., 1991. An evolutionary interpretation of van Huyck, Battalio, and Beil's experimental results on coordination. Games Econ. Behav. 3, 25–59.
Cross, J.G., 1983. A Theory of Adaptive Economic Behavior. Cambridge University Press.
Ellison, G., 1993. Learning, local interaction, and coordination. Econometrica 61 (5), 1047–1071.
Erev, I., Roth, A.E., 1998. Prediction how people play games: reinforcement learning in experimental games with unique, mixed strategy equilibria. Am. Econ. Rev., 848–881.
Foster, D., Young, H.P., 1990. Stochastic evolutionary game dynamics. Theor. Popul. Biol. 38, 219–232.
Foster, D.P., Young, H.P., 2006. Regret testing: learning to play Nash equilibrium without knowing you have an opponent. Theor. Econ. 1, 341–367.
Friedman, D., Ostrov, D., 2013. Evolutionary dynamics over continuous action space for population games that arise from symmetric two-player games. J. Econ. Theory 148, 743–777.
Germano, F., Lugosi, G., 2007. Global Nash convergence of Foster and Young's regret testing. Games Econ. Behav. 60, 135–154.
Harley, C.B., 1981. Learning the evolutionarily stable strategy. J. Theor. Biol. 89 (4), 611–633.
Harsanyi, J.C., Selten, R., 1988. A General Theory of Equilibrium Selection in Games. MIT Press.
Hart, S., Mas-Colell, A., 2003. Uncoupled dynamics do not lead to Nash equilibrium. Am. Econ. Rev. 93, 1830–1836.
Hart, S., Mas-Colell, A., 2006. Stochastic uncoupled dynamics and Nash equilibrium. Games Econ. Behav. 57, 286–303.
Hoffman, M., Suetens, S., Gneezy, U., Nowak, M.A., 2015. An experimental investigation of evolutionary dynamics in the rock–paper–scissors game. Sci. Rep. 5, 8817.
Huttegger, S.M., Skyrms, B., 2012. Emergence of a signaling network with "probe and adjust". In: Signaling, Commitment, and Emotion. MIT Press, Cambridge, MA.
Huttegger, S.M., Skyrms, B., Zollman, K.J.S., 2014. Probe and adjust in information transfer games. Erkenntnis 79 (4), 835–853.
Van Huyck, J.B., Battalio, R.C., Beil, R.O., 1990. Tacit coordination games, strategic uncertainty, and coordination failure. Am. Econ. Rev. 80, 234–248.
Hwang, S.H., Newton, J., 2014. A classification of bargaining solutions by evolutionary origin. Working paper no. 2. University of Sydney, School of Economics.
Kandori, M., Rob, R., 1995. Evolution of equilibria in the long run: a general theory and applications. J. Econ. Theory 65, 383–414.
Kandori, M., Mailath, G.J., Rob, R., 1993. Learning, mutation, and long run equilibria in games. Econometrica 61, 29–56.
Lim, W., Neary, P., 2014. An experimental investigation of stochastic stability. Working paper.
Maes, M., Helbing, D., 2014. Noise in behavioral models can improve macro-predictions when micro-theories fail. Mimeo.
Marden, J.R., Young, H.P., Arslan, G., Shamma, J., 2009. Payoff-based dynamics for multi-player weakly acyclic games. SIAM J. Control Optim. 48, 373–396.
Marden, J.R., Young, H.P., Pao, L.Y., 2014. Achieving Pareto optimality through distributed learning. SIAM J. Control Optim. 52 (5), 2753–2770.
McFadden, D., 1974. Conditional logit analysis of qualitative choice behavior. In: Zarembka, P. (Ed.), Frontiers in Econometrics, pp. 105–142.
McKelvey, R., Palfrey, T., 1998. Quantal response equilibria for extensive form games. Exp. Econ. 1, 9–41.
McKelvey, R.D., Palfrey, T.R., 1995. Quantal response equilibria for normal form games. Games Econ. Behav. 10, 6–38.
Montaneri, A., Saberi, A., 2010. The spread of innovations in social networks. Proc. Natl. Acad. Sci. USA 107 (47), 20196–20201.
Myatt, D.P., Wallace, C.C., 2004. Adaptive play by idiosyncratic agents. Games Econ. Behav. 48, 124–138.
Myerson, R., 1978. Refinements of the Nash equilibrium concept. Int. J. Game Theory 7, 73–80.
Naidu, S., Hwang, S.H., Bowles, S., 2010. Evolutionary bargaining with intentional idiosyncratic play. Econ. Lett. 109, 31–33.
Nash, J., 1950a. Non-cooperative games. PhD thesis. Princeton University.

Nash, J., 1950b. Equilibrium points in n-person games. Proc. Natl. Acad. Sci. USA 36 (1), 48–49.
Nash, J., 1951. Non-cooperative games. Ann. Math. 54 (2), 286–295.
Nax, H.H., Burton-Chellew, M.N., West, S.A., Young, H.P., 2013. Learning in a black box. Department of Economics WP 653. University of Oxford.
Newton, J., 2012a. Recontracting and stochastic stability in cooperative games. J. Econ. Theory 147 (1), 364–381.
Newton, J., 2012b. Coalitional stochastic stability. Games Econ. Behav. 75, 842–854.
Peski, M., 2010. Generalized risk-dominance and asymmetric dynamics. J. Econ. Theory 145, 216–248.
Pradelski, B.S.R., Young, H.P., 2012. Learning efficient Nash equilibria in distributed systems. Games Econ. Behav. 75, 882–897.
Roth, A.E., Erev, I., 1995. Learning in extensive-form games—experimental data and simple dynamic models in the intermediate term. Games Econ. Behav. 8, 164–212.
Sandholm, W.H., 2010. Population Games and Evolutionary Dynamics. The MIT Press.
Sawa, R., 2014. Coalitional stochastic stability in games, networks and markets. Games Econ. Behav. 88, 90–111.
Selten, R., 1975. Reexamination of the perfectness concept for equilibrium points in extensive games. Int. J. Game Theory 4, 25–55.
Skyrms, B., 2012. Learning to signal with probe and adjust. Episteme 9 (02), 139–150.
Suppes, P., Atkinson, A.R., 1959. Markov Learning Models for Multiperson Situations. Stanford University Press.
Taylor, P.D., Jonker, L., 1978. Evolutionary stable strategies and game dynamics. Math. Biosci. 40, 145–156.
Thuijsman, F., Peleg, B., Amitai, M., Shmida, A., 1995. Automata, matching and foraging behavior of bees. J. Theor. Biol. 175 (3), 305–316.
Weibull, J., 1995. Evolutionary Game Theory. The MIT Press.
Wilcox, N.T., 2008. Stochastic models for binary discrete choice under risk: a critical primer and econometric comparison. In: Cox, J.C., Harrison, G.W. (Eds.), Risk Aversion in Experiments, vol. 12, pp. 197–292.
Xu, B., Zhou, H.-J., Wang, Z., 2013. Cycle frequency in standard rock–paper–scissors games: evidence from experimental economics. Physica A 392 (20), 4997–5005.
Young, H.P., 1993. The evolution of conventions. Econometrica 61, 57–84.
Young, H.P., 1998. Individual Strategy and Social Structure: An Evolutionary Theory of Institutions. Princeton University Press.
Young, H.P., 2009. Learning by trial and error. Games Econ. Behav. 65, 626–643.
Young, H.P., 2011a. The dynamics of social innovation. Proc. Natl. Acad. Sci. USA 108 (4), 21285–21291.
Young, H.P., 2011b. Commentary: John Nash and evolutionary game theory. Games Econ. Behav. 71 (1), 12–13.