SFB 649 Discussion Paper 2008-026

Information and Beliefs in a Repeated Normal-form Game

Dietmar Fehr*
Dorothea Kübler*
David Danz*

* Technische Universität Berlin, Germany

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

http://sfb649.wiwi.hu-berlin.de
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin

SFB 649 ECONOMIC RISK BERLIN

Information and Beliefs in a Repeated Normal-form Game*

Dietmar Fehr
Technische Universität Berlin

Dorothea Kübler
Technische Universität Berlin & IZA

David Danz
Technische Universität Berlin

March 29, 2008

Abstract

We study beliefs and choices in a repeated normal-form game. In addition to a baseline treatment with common knowledge of the game structure and feedback about choices in the previous period, we run treatments (i) without feedback about previous play, (ii) with no information about the opponent's payoffs and (iii) with random matching. Using Stahl and Wilson's (1995) model of limited strategic reasoning, we classify behavior with regard to its strategic sophistication and consider its development over time. We use belief statements to track the consistency of subjects' actions and beliefs as well as the accuracy of their beliefs (relative to the opponent's true choice) over time. In the baseline treatment we observe more sophisticated play as well as more consistent and more accurate beliefs over time. We isolate feedback as the main driving force of such learning. In contrast, information about the opponent's payoffs has almost no effect on the learning path. While it has an impact on the average choice and belief structure aggregated over all periods, it does not alter the choices and the belief accuracy in their development over time.

Keywords: experiments, beliefs, strategic uncertainty, learning

JEL classification numbers: C72, C92, D84

*For helpful comments we thank Harald Uhlig, Georg Weizsäcker, Axel Werwatz and seminar participants at the Humboldt Universität zu Berlin, the European University Institute Florence, the SFB 649 Workshop (Motzen), the ESA World Meeting 2007 (Rome) and the IMEBE 08 (Alicante). We are indebted to Jana Stöver and Susi Thiel for research assistence and programming. Financial support from the Deutsche Forschungsgemeinschaft (DFG) through SFB 649 "Economic Risk" is gratefully acknowledged. Corresponding Author: Dorothea Kübler, Faculty of Economics and Management, H 50, Straße des 17. Juni 135, 10623 Berlin, Germany. Email: *d.kuebler@ww.tu-berlin.de*

1 Introduction

The literature on learning has opened the black box of how an equilibrium is reached. Numerous theoretical and experimental papers have studied learning over a large number of periods and have focused either on the convergence properties of the learning algorithms or on the evolution of observed behavior in experimental data. Here, we focus on the development of behavior in relatively few periods of play. The idea is to take a microscopic view of how beliefs and choices change over time, controlling for the impact of information on this process. Thus, our experiment aims at a better understanding of the development of strategic thinking.

We use a repeated two-person normal-form game with a unique Nash equilibrium of the stage game. In this relatively simple setup, we observe whether subjects learn to play the game in the sense that the Nash-equilibrium strategy is chosen more often in later than in earlier periods. A novel feature of the experiment is that we elicit the beliefs of a player about the action of the other player in every period. Thus, we can observe the joint development of beliefs and actions over time. This allows us to answer a number of questions in a dynamic setting that have up to now only been studied in one-shot games.

A widely used classification of behavior regarding the level of strategic thinking is the level-of-reasoning model of Stahl and Wilson (1995). We will rely on this approach to categorize behavior and in particular to study how behavior changes over time. Players can have various levels of strategic sophistication. The most important behavioral rules proposed by Stahl and Wilson are $L0$ which prescribes randomization over all possible actions, $L1$ which prescribes a best response to $L0$, $L2$ a best response to $L1$ etc. In addition, we will look for the Nash strategy, where a player chooses a best response to the belief that the other player chooses the Nash Equilibrium strategy.

The categorization of choices according to their strategic sophistication is complemented by the elicited beliefs. First, we can analyze whether players' actions are best responses to their beliefs more frequently in later periods than in the beginning of the experiment. The alternative hypothesis would be that best-response behavior is invariant over time, i.e. it is not learned in the limited time of an experimental session. In addition, we study whether beliefs become more accurate in predicting the opponents' actual behavior in later periods. In the fixed-pair matching protocol we employ, it can be expected that observing the other player's actions leads to more accurate beliefs over time.

In order to better understand the reasons for the development of actions and beliefs over time, we vary the information that is available to the players. Learning theories typically make use only of a limited amount of information. To be able to separate between different forms of learning, we run a baseline treatment with full information about the game and feedback about one's own payoff (and thereby the other's payoff and action) in the previous period. In addition, we run a treatment where subjects do not get any feedback about the outcome of play in the previous period and a treatment where subjects do not know the payoffs of the other player in the game, only their own payoffs. As we change only one aspect at a time, we can observe which kind of information is important for the learning process. Finally, we control for repeated game affects by running a treatment with random matching in every period.

In principle, two extreme learning patterns are possible and can be distinguished with our data. First, subjects can learn inductively, based on the history of play. Players look back to determine which strategy to choose in the next period. For example, belief learning and reinforcement learning fall into this category. Second, deductive reasoning implies that players analyze the game in order to understand its strategic properties and thereby form beliefs about the opponent's choice. This learning without feedback requires more sophistication of the players than the typical inductive learning algorithms. While both forms of learning have already been studied separately, we provide a unified framework to compare no-feedback learning with inductive learning. Using the level-k model, we characterize behavior as strategic or non-strategic and can then evaluate under which information conditions subjects learn faster to play strategically than in others.

We find an initially high level of non-strategic behavior in all treatments, i.e. subjects tend to neglect the incentives of their opponents. In the baseline treatment with full informtion about the game and feedback about past outcomes, this non-strategic behavior decreases in later rounds. Also, experience has a moderate positive impact on the accuracy of beliefs and on the best-response rates in the baseline treatment. The learning path crucially depends on the information available. At first, subjects seem to have only a limited understanding of the strategic properties of the game, showing a rather low level of strategic sophistication. Accordingly, information about the other player's payoff is of some importance for initial play, but not as much as expected. Behavior over time is very similar in treatments with and without information about the opponent's payoff. Our results clearly indicate the importance of feedback. Independent of whether subjects know the complete game or only their own payoffs, it is the experience through feedback which reduces non-strategic behavior. Analyzing the beliefs we conclude that the elicited beliefs are a better proxy for the underlying true beliefs than beliefs generated by established belief learning models. While both, feedback and complete payoff information, have a positive impact on the development of the best response rates, only feedback is needed to observe an increase of belief accuracy over time.

The literature related to this study can be organized into several groups of papers. First, the level-of-reasoning model by Stahl and Wilson (1995) has been applied to a number of data sets based on 3x3 one-shot normal-form games. Costa-Gomes, Crawford and Broseta (2001) study decision rules and use the mouselab technique to record how subjects use payoff information. Costa-Gomes and Weizsäcker (forthcoming) elicit the subjects' beliefs about the other player's choice and find that subjects perceive the game differently when asked for beliefs than when playing it themselves. Rey-Biel (forthcoming) focuses on constant-sum games to analyze the dependency of equilibrium predictions on the game characteristics. Finally, Ivanov (2006) combines the level-of-reasoning approach with risk aversion to explain observed behavior.

Repeated normal-form games with belief elicitation have been studied in two other papers. Nyarko and Schotter (2002) focus on the matching-pennies game to compare stated beliefs with Cournot and fictitious-play beliefs. Ehrblatt, Hyndman, Özbay and Schotter (2006) use two different normal-form games with a unique Pareto-efficient Nash Equilibrium in pure strategies to study convergence to the Nash Equilibrium. They focus on the mechanisms underlying the convergence process and on strategic teaching. Our experimental design is closest to the last paper. However, the Nash equilibrium in our game is not Pareto-efficient, leading to less convergence. We focus more broadly on learning how to play strategically and pay close attention to the development and nature of non-strategic play.

Another strand of the literature studies learning in normal-form games under different information conditions. Oechssler and Schipper (2003) and Gerber (2006) use normal-form games with incomplete information about opponents' payoff in order to study whether players can figure out which game they are playing. Subjects receive feedback about the strategy chosen by the other player and can thereby form a "subjective game" (Kalai and Lehrer, 1993). In contrast, Weber (2003) studies a repeated beauty-contest game without feedback and Weber and Rick (2008) focus on repeated normal-form games without feedback. Both studies observe some amount of no-feedback learning.

The paper is organized as follows. The next section introduces the design of the experiment and provides a detailed description of the Stahl and Wilson model applied to the normal-form game we used. Section 4 present the results, focusing first on choices and then on belief statements. Section 5 concludes.

	Left	Center	Right
Top	78, 68	72, 23	12, 20
Middle	67, 52	59, 63	78, 49
Bottom	21, 11	62, 89	89, 78

Table 1: Game

2 Experimental design

2.1 Procedures

In all of our experiments we used the asymmetric normal-form game presented in Table 1. The game has a unique Nash equilibrium in pure strategies in which the row player chooses Top and the column player chooses Left. This equilibrium can be found by applying iterative elimination of dominated strategies. Note that the Nash equilibrium of the stage game is not Pareto efficient. The strategy combination of Bottom and Right leads to higher payoffs for both players.[1] This outcome maximizes the payoff of the player that is least well off, and it also maximizes the sum of payoffs. The unique Nash equilibrium of the stage game is also the unique subgame perfect equilibrium of the repeated game.

To study the impact of information on choices and belief statements we implemented four treatments, the details of which are given in Table 2. Our main interest is in the baseline treatment, denoted by BASE. In this treatment subjects had all relevant information about the game, i.e. the set of players, the set of strategies and the payoff function of each player. In addition, after each period they received feedback about the payoff earned in this period. Every other treatment differs from BASE only in one respect. In the treatments NF (no feedback) and RM (random matching) subjects had common knowledge of the elements of the game, but we varied the available feedback after each period and the matching protocol. In treatment NF, subjects received no feedback at all. In treatment RM subjects received feedback about their payoff, but were randomly matched with another participant in each period. In treatment PI (partial information), subjects had incomplete information about some elements of the game. They only knew their own payoff function, but

[1] There is a Nash equilibrium of the finitely repeated game in which the players play this strategy combination (Bottom, Right) for a number of periods and then switch to the Nash Equilibrium (Top, Left). In case a player deviates in this equilibrium, she is minmaxed by the other player choosing Middle or Center, respectively, for the rest of the game.

Treatment	Payoff	Feedback	Matching	Periods	Sessions	# of subjects
BASE	own/opponent	own payoff	fixed	20	4	54
PI	own	own payoff	fixed	20	4	48
NF	own/opponent	none	fixed	20	4	50
RM	own/opponent	own payoff	random	20	3	40

Table 2: Treatments

not the payoff function of their opponent. But they received feedback after each period, just as in treatments BASE and RM, such that they could infer the choice of their opponent. In all treatments subjects did not receive any feedback about their payoffs from the belief elicitation task.[2]

In the beginning of all treatments, subjects were randomly assigned a player role (row player or column player), which they kept during the whole experiment. However, they made all their decisions from the perspective of the row player, i.e. for column players we used a transformation of the matrix game in Table 1. Before choosing an action, we asked subjects to indicate their beliefs regarding the behavior of their opponent.[3] In particular, we asked subjects to state the expected frequencies of play, i.e., they had to specify in how many out of 100 times they expect the column player to choose Left, Center and Right in the current period.[4] After the belief elicitation task, subjects had to make their choice by selecting one of the three possible actions (mixing was not possible).

Subjects were paid for both tasks. For the choice task we paid subjects according to the numbers in the payoff matrix, which were exchanged at the commonly known rate of 1 point = €0.15. To reward the belief elicitation task we used a quadratic scoring rule (QSR) which is incentive compatible given that subjects are risk-neutral money maximizers. At the end of the experiment, we randomly and independently selected one period to determine the payoffs for each of the two tasks.

The QSR is defined as follows. The payoff Π_{it}^{QSR} for player i in period t for a given action

[2] Nevertheless, they could infer their payoff from this task in treatments BASE, PI and RM. One reason for not showing the payoffs from the belief elicitation task was to change as few parameters as possible when going from BASE, PI and RM to NF.

[3] The same procedure was also applied by other studies, e.g. Costa-Gomes and Weizsäcker (forthcoming) or Rey-Biel (forthcoming).

[4] For simplicity we restricted the expected frequencies of play to integers. Therefore, we count any belief assigning a weight of 34 percent to one action and 33 percent to each of the remaining actions as a uniform belief statement.

a_{jt}^k with $k \in \{L, C, R\}$ of player j in period t and belief vector $b_{it} = \left(b_{it}^L, b_{it}^C, b_{it}^R\right) \in \Delta^2$ such that $\Delta^2 = \left\{b_{it} \in \mathbb{R}^3 | \sum_{k \in \{L,C,R\}} b_{it}^k = 1\right\}$ is

$$\prod\nolimits_{it}^{QSR} (b_{it}, a_{jt}) = A - b * \left(\sum\nolimits_{k \in \{L,C,R\}} \left(b_{it}^k - 1_{[a_{jt}^k]}\right)^2\right) \tag{1}$$

where $1_{[a_{jt}^k]}$ is an indicator function equal to 1 if a_{jt}^k is chosen in period t and 0 otherwise. While paying subjects for both tasks is necessary to ensure incentive compatibility, it allows subjects to engage in hedging. Subjects can for example coordinate on a cell of the payoff matrix that is not an equilibrium and become unwilling to move away from it in order to avoid losses in the belief elicitation task. To avoid such behavior we chose the two parameters A and b of the QSR such that the maximum payoff from the belief elicitation task is low compared to payoffs from choice task. In the experiment the parameters are set to $A = 1.5$ and $b = 0.75$. Thus, the maximum payoff from the belief elicitation task is €1.50, the minimum is €0. Note that subjects could guarantee themselves a payoff of €1 by stating uniform beliefs.[5] Note also that the Nash Equilibrium [Top, Left] would lead to a payoff of €11.7 and €10.2, respectively.

The experiments were conducted in the computer lab at Technische Universität Berlin using the software tool kit *z-Tree*, developed by Fischbacher (2007). Subjects were recruited via a mailing list, where they could voluntarily register for participating in decision experiments. Upon entering the lab, subjects received written instructions and were asked to read them carefully and take their time.[6] After everybody had finished reading the instructions, we distributed an understanding test that covered both the matrix game and the QSR. Only after all subjects had answered the questions correctly, we proceeded with the experiment. In total 192 students (106 males and 86 females) from various disciplines participated in the four treatments. Sessions lasted about one hour. Subjects' average earnings were about €12.80, including a show-up fee of €3 for arriving at the laboratory on time.

2.2 Strategies

Stahl and Wilson (1995) proposed a theory of boundedly rational types, based on a hierarchical model by Nagel (1993). Stahl and Wilson assume that players differ in their level of strategic sophistication. Their model classifies players into types according to their "level of reasoning",

[5] Although stating uniform beliefs can be an attractive choice for a risk-averse subject, we find no evidence for such behavior in our treatments. Only 7.5 percent of belief statements in our experiment assign no less than 30 and no more than 35 percent to all three of the opponent's actions. (BASE 5.8%, PI 5.9%, NF 12.1% and RM 6.3%)

[6] The instructions are available from the authors upon request.

hence the term level-k model. A level-0 type randomizes uniformly over his strategy space, whereas a level-k type best responds to level-$(k-1)$ behavior for $k \in \{1, 2, .., \infty\}$.[7]

The level-k model is a useful approach to track off-equilibrium behavior. It has been tested and extended by various other studies (e.g. Costa-Gomes et al, 2001, Costa-Gomes and Weizsäcker, forthcoming and Camerer et al, 2004). The model and its extensions are successful in organizing data, e.g. from normal-form and beauty contest games, but also from other games as recently shown by Crawford and Iriberri (2007a, 2007b) and Gneezy (2005). The most common types found are level-1 ($L1$), level-2 ($L2$) and Nash, but the distribution of types crucially depends on the set of games investigated.

All above mentioned normal-form game studies use data from one-shot interactions. In a repeated setting additional strategic considerations come into play, and learning becomes possible. The level-k model can accommodate learning in that subjects can learn to play higher-level strategies. Suppose a subject starts out by playing the $L1$ action, but then learn to best respond to $L1$ by playing $L2$ and so forth. Thus, subjects can learn by updating their beliefs in the course of the game, and we will investigate this on the basis of our data.

We use the level-k model to classify the strategy space of our game. The most prominent types of the level-k model ($L1$, $L2$ and Nash) can be classified into two broad types, namely strategic and non-strategic types. Strategic types form beliefs based on an analysis of what others do and best respond to these beliefs, whereas non-strategic types do not take into account the incentives of others. Given this definition, strategic types are $L2$ and Nash and the non-strategic type is $L1$.

We also introduce a Rawlsian rule, defined as choosing the action that maximizes the payoff of the player with the lower payoff, given the other player has the same objective and chooses accordingly. Remember that in the game we use, the Rawls strategy is the same as the Utilitarian strategy which maximizes the sum of payoffs. With our definition of strategic behavior, the Rawls action should be counted as strategic because the rule requires the belief that the other player has the same preferences and acts accordingly (the same reasoning holds for its interpretation as a Utilitarian rule). Previous studies did not explicitly explore Rawlsian or Utilitarian strategies, but some of them found behavior pointing in this direction (e.g. Costa-Gomes and Weizsäcker,

[7] The model contains also other types to capture behavior eventually more in line with traditional game theory. These are the naive Nash type who chooses the Nash equilibrium strategy, the wordly type who plays a best response to a subjective distribution of all other types and the rational expectation type who correctly anticipates the distribution of boundedly rational types and best responds to this distribution.

	Row player		Column player
Top	Nash(L2)	**Left**	Nash
Middle	L1	**Center**	L1(L2)
Bottom	Rawls	**Right**	Rawls

Table 3: Decision rules

forthcoming). In order to be able to separate between Nash play and play of the most efficient and/or fair outcome, it is necessary to allow for efficiency or fairness to lead to a separate outcome, which motivated the choice of our game.[8]

The main focus of this study is on the development of strategic and non-strategic behavior over time. We therefore designed the game for the experiment so as to identify strategic and non-strategic behavior as clearly as possible. In particular our interest was to achieve the best possible separation of the four rules of behavior ($L1$, $L2$, Rawls and Nash). We chose an asymmetric game for which the different rules overlap differently for the two player roles. Table 3 summarizes the decision rules implied by the possible actions in the game.

Only the $L2$ rule cannot be identified clearly for any of the two player roles. For the row player, it prescribes the same action as Nash and for the column player it is the same as $L1$. Assuming that there is a considerable proportion of $L2$ play, which is suggested by previous studies, we will overestimate the proportion of Nash play of the row player and the proportion of $L1$ play of the column player.[9] We will keep this in mind when interpreting the findings. However, our focus is on subjects learning to play strategically, and the $L2$ rule represents an intermediate level of strategic reasoning. We are mainly interested in the comparison between $L1$ and Nash behavior as the two extreme ends of the spectrum of strategic play.

Notice that we use the names $L1$, Nash and Rawls also for the three strategies in treatment PI even though a priori the subjects cannot reason about the other player's incentives and consequently cannot identify the Nash and Rawls strategy in this treatment.[10]

[8] In contrast, Ehrblatt et al. 2006 run a similar experiment based on a game where the Nash equilibrium coincides with the Rawlsian/ Utilitarian outcome.

[9] Indeed, we find about 10% more Nash play for the row than for the column player and about 15% more L1 play for the column than for the row player.

[10] Only if the subjective game constructed by the participants happens to be equivalent to the true game, the names of the strategies can be interpreted as decision rules. See Kalai and Lehrer (1993) for the theory of subjective games.

3 Results

In the first part of our analysis we examine the choices of our subjects without considering their stated beliefs. We begin this analysis with a focus on first period behavior and a comparison of these results to previous experiments. Afterwards we extend our analysis to all periods and focus on the development of behavior over time and considering the impact of the information available. In the second part of the data analysis, we focus on the elicited beliefs. After confirming that they outperform standard models of belief formation, we examine the consistency of the corresponding best responses and the observed actions of the players with respect to their development over time and the information conditions. Furthermore we check the accuracy of the stated beliefs in predicting the opponent's choice and the role of feedback information and payoff information for the formation of beliefs.

Note that unlike in most other studies on asymmetric one-shot games (e.g. Costa-Gomes and Weizsäcker, forthcoming), we do not pool the data over player roles. As we study only one specific game, we are able to consider the exact strategic situation of each player role. This differentiation would be lost by pooling the data. Thus, we run all statistical tests separately for row and for column players.

3.1 Choices

3.1.1 First-period choices

We first look at behavior in the first period only. This is of some stand-alone interest, since many experiments on behavior in one-shot 3x3 normal-form games have used similar games, and we can compare the results. First-period behavior in each treatment is presented in Figure 1. The figure shows the fraction of each strategy in a certain treatment for row players and column players, respectively.

In the first period subjects in treatments BASE, RM and NF all face the same strategic situation. Therefore we should not observe any differences in behavior. This is clearly the case, as can be taken from Figure 1. The frequency of chosen strategies of the row players (column players) in all three treatments is 19 (8) percent Nash, 43 (64) percent $L1$ and 38 (28) percent Rawls. We cannot reject the hypothesis that the frequency of strategies is the same using a χ^2-Test.[11] Our

[11] For both player roles we perform a pairwise comparison of BASE with NF and RM, respectively. The test yields no p-value smaller than 0.64 ($\chi^2_{(2)}$).

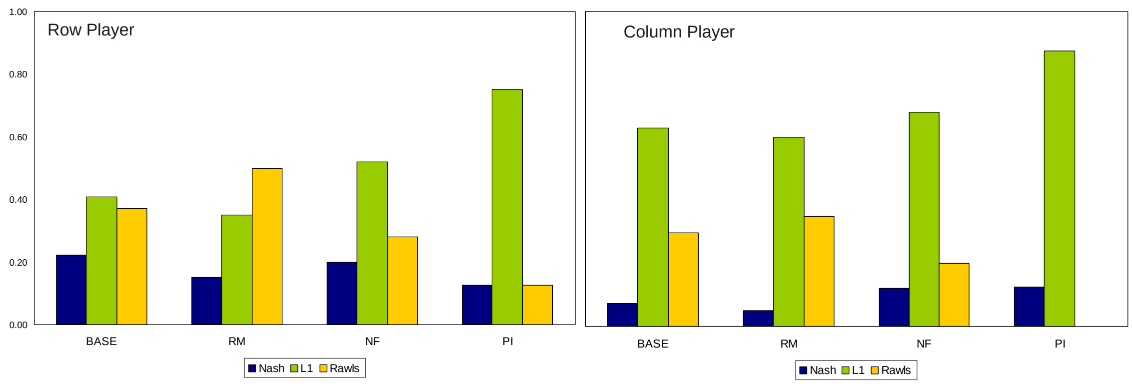

Figure 1: First period choices

findings of 64 percent $L1$ behavior in the first period in BASE, RM and NF are in line with previous studies.[12] For instance, Costa-Gomes and Weizsäcker (forthcoming) found approximately 60% $L1$ behavior, whereas Costa-Gomes et al. (2001) and Rey-Biel (forthcoming) found slightly lower rates of 45% and 50%, respectively.

Now, consider the decision situation of a subject in treatment PI. Subjects only know their own payoffs for their available strategies, such that they cannot base their decisions on strategic considerations. Hence, it is no surprise to see 39 out of 48 subjects (81 percent) choosing the $L1$ rule in period 1 in PI, which not only maximizes the minimum payoff, but also the expected payoff assuming that the opponent chooses her action randomly. We see a lot of violations of dominance by the column player in the other treatments where she knows the payoffs of the row player. It is remarkable that no column player in PI chooses Rawls in the first period, indicating that the choice of dominated actions in the other treatments is due to the payoff structure of the other player and not only to mistakes. The frequency of the three strategies in PI is significantly different from BASE in the first period for both player roles ($\chi^2_{(2)}$, $p = 0.043$ for row players and $p = 0.014$ for column players). We summarize the findings on choices in the first period in the following result.

Result 1 *(i) First-period behavior in BASE, RM and NF is statistically indistinguishable from each other and comparable to findings from one-shot experiments. (ii)* $L1$ *is the most frequently chosen strategy in the first period in all treatments and for both player roles. (iii) First-period play in treatment PI is significantly different from BASE.*

[12] For ease of comparison to the other studies, we pool $L1$ behavior over player roles.

	Row-Player			Column-Player		
	(1) **Nash**	(2) **L1**	(3) **Rawls**	(4) **Nash**	(5) **L1**	(6) **Rawls**
PI	0.310	0.818***	-1.039***	0.556**	0.651***	-1.199***
	(0.290)	(0.290)	(0.306)	(0.273)	(0.224)	(0.311)
NF	-0.331	0.849***	-0.419	0.104	0.644***	-0.662**
	(0.291)	(0.286)	(0.300)	(0.274)	(0.223)	(0.301)
RM	-0.008	0.428	-0.292	-0.241	0.693***	-0.464
	(0.306)	(0.305)	(0.318)	(0.295)	(0.237)	(0.318)
Constant	-0.590***	-1.252***	-0.018	-1.207***	-0.317**	-0.445**
	(0.201)	(0.204)	(0.207)	(0.196)	(0.154)	(0.207)
$\log\mathcal{L}$	-993.86	-925.47	-979.97	-821.49	-1160.56	-803.67
$\chi^2_{(3)}$	4.67	11.41***	12.02***	8.12**	13.24***	15.24***
N		1920			1920	

Notes: Random-effects probit regression, standard errors in parentheses, * Significant at 10-percent level; ** Significant at 5-percent level; *** Significant at 1-percent level.

Table 4: Regression: Decision rules

3.1.2 Choices over all periods

First, we are interested in average behavior over all 20 periods in the different treatments. For this purpose, we perform a separate regression for each strategy and player role combination. We regress the strategies on treatment dummies without controlling for time effects, which gives us a first indication of the influence of the different information conditions. To model the repeated decisions of the same subject in each treatment, we use random-effects regressions. Since subjects had to choose one out of three possible strategies, a probit model is employed where the dependent variable reflects the inclination to choose one strategy over the other two. All results reported in this paper are significant at the 5% level.

The results shown in Table 4 reveal the importance of the opponent's payoff and of the feedback on past choices. The coefficients of PI are significantly different from BASE for all strategies except for the Nash strategy of the row player. There is significantly more $L1$ play and less Rawls play in PI than in BASE. The lack of feedback in NF results in more $L1$ play than in BASE for both player roles and less Rawls play for column players. A similar but weaker effect can be

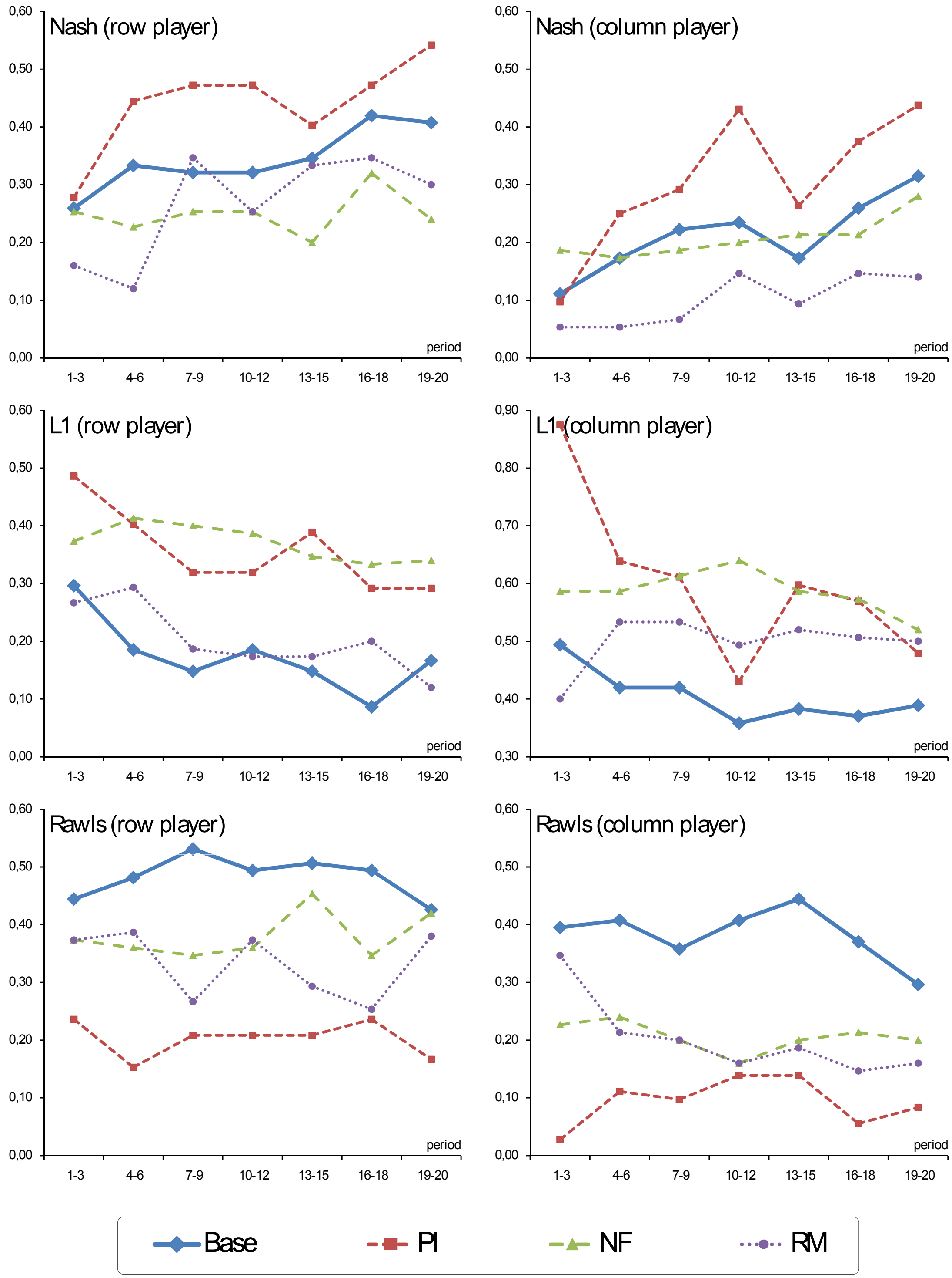

Figure 2: Decision rules over time

	Row Player			Column Player		
	(1) **Nash**	(2) **L1**	(3) **Rawls**	(4) **Nash**	(5) **L1**	(6) **Rawls**
PI	0.306	0.809**	-1.008***	0.443	0.987***	-1.413***
	(0.341)	(0.342)	(0.350)	(0.343)	(0.276)	(0.377)
NF	-0.083	0.559*	-0.477	0.354	0.532*	-0.683*
	(0.346)	(0.335)	(0.343)	(0.346)	(0.273)	(0.352)
RM	-0.209	0.411	-0.082	-0.204	0.371	-0.107
	(0.363)	(0.357)	(0.361)	(0.384)	(0.288)	(0.370)
Period	0.032***	-0.045***	0.001	0.042***	-0.018*	-0.013
	(0.012)	(0.013)	(0.011)	(0.013)	(0.010)	(0.012)
PI*Period	0.001	0.004	-0.003	0.011	-0.031**	0.020
	(0.016)	(0.018)	(0.017)	(0.018)	(0.015)	(0.020)
NF*Period	-0.024	0.032*	0.005	-0.023	0.011	0.002
	(0.017)	(0.017)	(0.016)	(0.018)	(0.015)	(0.017)
RM*Period	0.019	0.003	-0.021	-0.004	0.031**	-0.037**
	(0.017)	(0.019)	(0.017)	(0.021)	(0.016)	(0.018)
Constant	-0.929***	-0.831***	-0.030	-1.671***	-0.126	-0.310
	(0.240)	(0.239)	(0.236)	(0.251)	(0.187)	(0.242)
$\log\mathcal{L}$	-977.66	-908.25	-978.62	-801.85	-1147.71	-795.93
$\chi^2_{(7)}$	36.21***	43.73***	14.65***	45.70***	37.95***	29.82***
N		1920			1920	

Notes: Random-effects probit regression, standard errors in parentheses, * Significant at 10-percent level; ** Significant at 5-percent level; *** Significant at 1-percent level.

Table 5: Regression: Decision rules over time

observed in RM compared to BASE where $L1$ play is also more frequent.

We now turn to the question how behavior changes over time and how this change is influenced by the different information conditions. Choices over time are represented in Figure 2 which displays the proportion of the behavioral rules for each treatment. The figure shows averages over three periods in a given treatment for row players in the left panel and for column players in the right panel. As stated above we cannot clearly identify $L2$ behavior, since this rule overlaps with Nash for row players and $L1$ for column players (see also Table 3). This might be the reason why we observe on average less Nash and more $L1$ play of column players compared to row players over all treatments.

To investigate potential learning paths we extend our regressions from Table 4 by including a time trend and interaction terms for treatments with time. The results of these regressions are presented in Table 5. In these regressions the dummy variables are coded such that the corresponding coefficients represent the intercept and the development over time in each treatment relative to the baseline treatment. In order to assess the absolute time trends in each treatment directly, we additionally test the hypothesis that the sum of the coefficient for Period and the relevant coefficient for Treatment*Period is equal to zero.

First, let us focus on the development of the three strategies in BASE. The regression shows that in BASE subjects tend to choose the $L1$ strategy less often in later periods while Nash play increases and Rawlsian behavior is stable over time. We can now compare this learning path to the trends in treatment PI. The inclusion of time controls reveal that behavior in PI changes in a similar way as in BASE, with an even stronger decrease of $L1$ play for the column player. The average difference in the choices between the BASE and the PI treatment is therefore mainly due to the behavioral differences in the first period. Although the removal of feedback in treatment NF does not produce significant differences in the time trend compared to BASE, the time trends in NF are no longer significant, when tested directly. Finally, we compare the effect of random matching to fixed matching on the time trend. While we do not find differences between RM and BASE for row players, column players in RM choose Rawls less and $L1$ more often over time than in BASE. This may be due to the fact that reputation building is not possible and a deviation from Rawls to $L1$ which gives a higher payoff cannot be sanctioned effectively by the row player.

The findings based on the various regressions can be summarized as follows.

Result 2 *(i) In treatments PI and NF there is on average significantly more* $L1$ *and less Rawls play than in BASE. (ii) Over time the proportion of the Nash strategy increases in all treatments* *and player roles except in NF. (iii) The proportion of the $L1$ strategy decreases in BASE and PI over time. Again there is no similar time trend in NF. (iv) The proportion of Rawls choices is almost constant over time for all treatments and player roles (except for the column player in RM).*

Thus, in the sense of Stahl and Wilson we observe a trend towards more strategic play in all treatments with feedback in that there is an increase in Nash and a decrease in $L1$ play. In treatment PI the lower proportion of strategic behavior can be ascribed to the lack of information about the opponent's payoffs. However, the fact that players in PI can observe the choices of their opponent over time and react to these observations leads to a development of behavior away from the $L1$ rule, just as in BASE. In treatment NF behavior does not change over time. As the NF treatment is comparable to a repeated one-shot situation, this finding lends support to the frequently applied method of giving no feedback between different tasks in experiments in order to minimize learning effects. Finally, as our control for repeated game effects, treatment RM reveals no differences to BASE for the row player. But we see that the column player's behavior is affected by the matching protocol in that he chooses on average more non-strategic $L1$ play. Also, over time he is less likely to choose the dominated strategy (Rawls) in RM compared to BASE.

3.2 Belief formation

In standard equilibrium analysis it is assumed that subjects form beliefs over the behavior of the opponent and then best respond to these beliefs. Models of bounded rationality depart from this view either by positing that subjects best respond to their beliefs with noise, i.e. they make errors in best-responding to expectations (e.g. McKelvey and Palfrey, 1995) or that subjects differ in their strategic sophistication (e.g. Stahl and Wilson, 1995).

In this section we focus on the relationship between the elicited beliefs and the subjects' own as well as their opponents' actions. There are some caveats concerning the belief elicitation procedure. First, subjects need not hold beliefs about opponent's play at all. For example, they might choose some non-strategic decision rule in the first period and then condition play on received payoffs (reinforcement learning). Forcing them to state beliefs could alter play if these subjects move their decisions in the direction of belief-based play. However, our design is based on a comparison between treatments which all use belief elicitation. Unless, the effects of belief elicitation interact with our treatment variables, our results are immune to such problems.

Also, subjects might make mistakes when stating their beliefs, just as when taking decisions. We therefore propose that the belief statements should only be taken as a proxy of the true underlying beliefs of subjects.[13] Finally, even though we asked explicitly to state myopic beliefs, i.e. beliefs only for the current period, we cannot rule out that subjects follow repeated game strategies and state beliefs accordingly. This would open up the possibility that subjects exploit the repeated interaction structure of the game in order to achieve a cooperative outcome. As the choices that are part of repeated-game strategies are not necessarily best responses to myopic beliefs, we will take this into account when interpreting best-response rates. For this purpose, treatment RM is necessary to check for repeated-game effects.

3.2.1 Stated beliefs vs. models of belief formation

In this subsection we follow the approach used in Nyarko and Schotter (2002) and compare the explanatory power of elicited beliefs to standard belief learning models. The purpose of this comparison is to establish whether stated beliefs are a good measure of strategic uncertainty or whether stated beliefs are inferior to beliefs derived indirectly from the participants' choices.

Standard belief learning models assume that players update their beliefs based on the opponent's history of play and best-respond to these beliefs. The two most prominent models based on this assumption are the fictitious-play and the Cournot best response model. While in the Cournot model subjects best respond to the opponent's play in the very last period, players in a pure fictitious-play model best respond to beliefs based on all previous actions of the opponent. The γ-weighted fictitious-play model introduced by Cheung and Friedman (1997) contains Cournot best response and fictitious-play as special cases. In this model subject i's belief $b_{i,t+1}^k$ that subject j will choose action $a_{jt}^k, k \in \{L, C, R\}$ in period $t+1$ is defined as:

$$b_{i,t+1}^k = \frac{1_{[a_{jt}^k]} + \sum_{u=1}^{t-1} \gamma_i^u 1_{[a_{j,t-u}^k]}}{1 + \sum_{u=1}^{t-1} \gamma_i^u} \tag{2}$$

The parameter γ_i is the weight player i gives to the past actions of his opponent. It is obvious from (2) that $\gamma_i = 0$ leads to the Cournot best-response model and $\gamma_i = 1$ leads to fictitious-play, respectively. In accordance with the imperfect best-response behavior observed in the preceding subsections we use a standard logistic choice model in which subjects to choose their actions with some noise in response to their beliefs. Subject i chooses action k with probability

$$\Pr\left(a_{it}^k\right) = \frac{\exp\left(\lambda\pi[a_{it}^k, b_{it}]\right)}{\sum_{l \in \{L,C,R\}} \exp\left(\lambda\pi[a_{it}^l, b_{it}]\right)}, \tag{3}$$

[13] See Costa-Gomes and Weizsäcker (forthcoming) for a thorough analysis of belief statements.

where $\pi[a_{it}^k, b_{it}]$ is the expected payoff of player i when she chooses an action k given her beliefs b_{it} over the action set of her opponent. The parameter λ determines the impact of this expected payoff on her own choice probability and thus can be interpreted as a rationality parameter. A player with $\lambda = 0$ chooses all actions with equal probability disregarding the expected payoff of her choice. On the other hand if $\lambda \to \infty$ the player is unboundedly rational, i.e. she makes no errors in best responding to her beliefs.

We now turn to the estimation and probabilistic comparison of the choice model (3) based on the most flexible γ-weighted fictitious-play model (2) on the one hand and on the stated beliefs on the other hand. Since the belief-learning model assumes that subjects process information about their own payoff matrix and about the history of their opponent's play, only the data of treatments BASE and PI are used in the following analysis, while treatment NF is excluded. The data from treatment RM are also analyzed since the process described in (2) can be interpreted as the formation of beliefs over the average play of the population rather than over individual choices. The estimation results for each treatment and player role separately are presented in Table 6.[14]

		ML-Estimation of Model (3) using					**Model Selection Tests**			
		Fictitious-play (2)			**Stated Beliefs**		**Vuong's Test**		**Clarke's Test**	
Treatment	**Role**	λ	γ	$\log \mathcal{L}$	λ	$\log \mathcal{L}$	**Z**	**p-value**	**Z**	**p-value**
BASE	row	0.0575	0.7418	-484.01	0.1005	-422.12	-3.46	0.0005	-9.12	0.0000
	column	0.0373	0.6009	-492.96	0.0586	-421.93	-6.88	0.0000	-5.42	0.0000
PI	row	0.0442	0.6488	-487.68	0.0646	-451.16	-3.72	0.0002	-3.47	0.0005
	column	0.0571	0.6220	-413.59	0.1066	-307.98	-5.82	0.0000	-14.97	0.0000
RM	row	0.0233	0.5821	-427.04	0.0825	-372.34	-5.35	0.0000	-5.90	0.0000
	column	0.0729	0.9067	-350.21	0.0604	-334.25	-1.42	0.1548	-1.30	0.1936

Notes: p-values are two-sided. Clarke's corrected B has been by the approximated by the standard normal distribution.

Table 6: Model Estimation and Selection.

As a first result we observe that the stated beliefs play a significant role in explaining the behavior of our subjects, since appropriate likelihood-ratio tests reject the hypothesis that the rationality parameter λ is equal to zero ($p = 0.00$ for all treatments and player roles).

[14] With respect to the γ-weighted fictitious-play model we estimated γ and λ simultaneously. All ML-estimations and tests have been conducted with MATLAB and R.

Using tests for the selection between non-nested models introduced by Vuong (1989) and Clarke (2003)[15], the hypothesis of equal explanatory power of the models can be rejected at all usual significance levels for all treatments and player roles, the only exception being the column player in the random-matching treatment.[16] In our notation the negative signs of the test statistics reveal that the stated belief model is closer to the real data generating process than the beliefs generated by the belief-learning models.

To summarize, we extend the finding of Nyarko and Schotter (2002) from a matching-pennies game to our normal-form game with a unique Nash equilibrium in pure strategies. We find that stated beliefs are better at explaining observed choices than beliefs that are implied by the standard models of belief formation. Therefore, we use the stated beliefs for analyzing the impact of experience and information on the consistency and accuracy of beliefs.

3.2.2 Consistency of actions and stated beliefs

Both in the standard Nash equilibrium and in the level-k model it is assumed that subjects best respond to their beliefs. Therefore, we investigate the consistency of actions and stated beliefs, i.e. whether subjects best respond to their stated beliefs. In Figure 3 the proportion of players best responding to their stated beliefs is displayed for each player role and treatment separately. Again, the figure shows the average proportion over three periods. Obviously, the best response rates are low, but in line with previous findings ranging from 54 to 75 percent best responses. In order to compare our results to these studies, it is useful to look at aggregated best-response behavior over all subjects. Averaging over all treatments and player roles, subjects best-respond to their stated beliefs in 63 percent of the cases. The best-response rates found in similar studies are summarized in Table 7. In simple games like 2x2 games (Nyarko and Schotter, 2002) or constant-sum games (Rey-Biel, forthcoming) consistency rates are around 75 percent, whereas in more complicated games similar to the one we used (e.g. Costa-Gomes and Weizsäcker, forthcoming, Ehrblatt et al.,

[15] Vuong's test is based on the overall likelihood ratio of two rival models and is asymptotically normally distributed under the null. Clarke's test uses the number of single likelihood ratios being greater than 1 which is under the null binomial distributed with $\theta = 0.5$ and the number of observations in each data subset. Vuong's test is outperformed by the latter when the distribution of the single log-likelihood ratios is highly peaked. Both tests where calculated using corrections for the dimension of the models as proposed by Schwarz (1978) and Clarke (2007) respectively.

[16] The insignificance of test statistic for the column player in treatment RM may be due to the fact that the fictitious play model is sufficient to capture the formation of beliefs over the play of a population, since the beliefs in this model are just weighted averages of all opponents' historical actions that tend to stabilize over time.

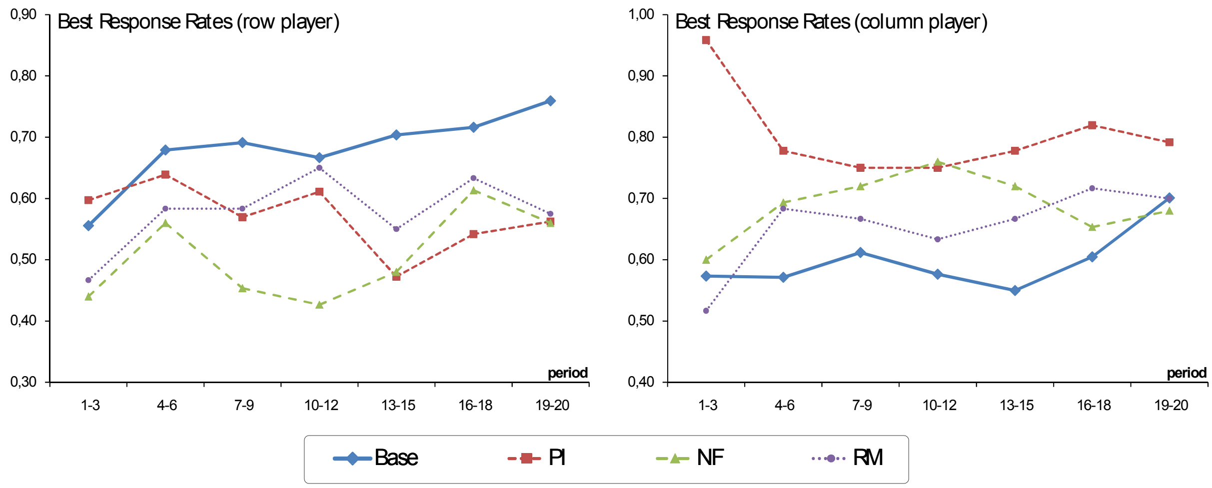

Figure 3: Best-response rates over time

	Costa-Gomes & Weizsäcker (forth.)	**Rey-Biel (forth.)**	**Ehrblatt et al (2006)**	**Nyarko & Schotter (2002)**	**our study**
Games	various 3x3	various 3x3	two 3x3	one 2x2	one 3x3
Interaction	one-shot	one-shot	repeated	repeated	repeated
⌀	54	73	54	75	63

Table 7: Best-response rates (in %) in various studies

2006) rates are about 54%.

For statistical evidence on differences between the treatments and the development of best response rates over time, we run random-effects panel regressions. As the dependent variable is either 0 (no best response) or 1 (best response), we use a probit model. Besides the constant the independent variables are dummies for PI, NF and RM, a linear time trend and interaction dummies for trend and treatment. The regression results are summarized in Table 8. Again we run direct tests of the absolute time trends in the control treatments.

Consider in the following the comparison of each control treatment with BASE separately and let us start with PI. For the row player, the average number of best responses in BASE is slightly higher than the number of best responses in PI. The opposite holds for the column player which is due to less Rawls play in PI (as the efficient or Rawls payoff combination cannot be identified as such), avoiding violation of dominance. Using direct tests for the time trends, there is no significant development over time for both player roles in PI. When comparing BASE to NF

	Best Response			
	row player		column player	
	(1)	(2)	(3)	(4)
PI	-0.356*	0.105	0.846***	1.262***
	(0.212)	(0.263)	(0.257)	(0.314)
NF	-0.517**	-0.387	0.499**	0.578*
	(0.209)	(0.260)	(0.253)	(0.302)
RM	0.326	-0.176	0.335	0.244
	(0.222)	(0.275)	(0.269)	(0.320)
Period		0.032***		0.018*
		(0.011)		(0.011)
PI*Period		-0.046***		-0.039**
		(0.015)		(0.017)
NF*Period		-0.014		-0.007
		(0.015)		(0.016)
RM*Period		-0.016		0.009
		(0.016)		(0.017)
Constant	0.557***	0.234	0.176	-0.012
	(0.146)	(0.181)	(0.174)	(0.208)
$\log\mathcal{L}$	-1160.74	-1152.70	-1008.56	-1002.94
$\chi^2_{(.)}$	$\chi^2_{(3)} = 6.48^{*}$	$\chi^2_{(7)} = 22.13^{***}$	$\chi^2_{(3)} = 11.22^{**}$	$\chi^2_{(7)} = 22.09^{***}$
N	1920		1920	

Notes: Random-effects probit regression, * Significant at 10-percent level; ** Significant at 5-percent level; *** Significant at 1-percent level

Table 8: Regression: Best-response rates over time

the overall level of best responses is again higher in BASE than in NF for the row player and lower for the column player (because the dominated Rawls strategy is played less often). As in PI the time trends in NF are not significant when tested directly. Besides the row players in BASE, only the column players in RM display higher best-response rates in later periods (the time trend being significant when using a direct test).

These findings raise the question why best-response rates of the row player are higher in BASE compared to treatments PI and NF. Internal consistency requires best responding to one's beliefs, independent of the information conditions. We can merely offer possible explanations of our observations, but further research is necessary to disentangle the causes of behavior more thoroughly. In NF, subjects might be doubtful about the accuracy of their beliefs, lacking any information about the other player's behavior. This might induce them to put less weight on their beliefs when choosing an action. In treatment PI where players can only learn the structure of the game over time, there is also no discernible increase in best-response behavior. Two possible explanations come to mind. First, the complexity of learning both the structure of the game and to best respond to beliefs at the same time is too high. Second, in treatment PI many subjects start with uniform beliefs and best respond to them. As the belief set of $L1$ is large and $L1$ is an attractive strategy initially, there is a high rate of consistency at the outset. This effect is absent in BASE and NF.

The focus of the preceding analysis was on myopic beliefs. However, our game also allows subjects to achieve a cooperative outcome, because in our repeated-game setting Folk-Theorem results are possible. If this is the case, column players choose their dominated action (Rawls) in response to Rawls play of row players. A necessary condition for a repeated game strategy is the observability of the (past) behavior such that subjects can condition their actions on opponents' play. To achieve a cooperative outcome a minimum of information flow is needed to make sanctions for deviations possible.[17] The fact that Rawls can never be a best response to any myopic belief statement, explains why we observe very low best-response rates for column players in BASE but not in NF and PI where less Rawls play is observed. If the low best-response rates are indeed a result of repeated-game strategies, we should observe significantly higher best-response rates in RM. The reason is that the finite time horizon and the random-matching protocol do not allow for cooperation based on Folk theorem results. But we observe a substantial proportion of Rawls play

[17] For instance, Ellison (1994) and Kandori (1992) have shown for infinitely repeated games with random matching that a cooperative outcome is possible through contagious sanctions.

also in RM in both player roles. Moreover, the regressions reveal no significant differences between BASE and RM neither for the extent of average Rawls play (see Table 4) nor for the average best response rates (see Table 8).

Theoretically, the insignificant difference of best response rates could be due to a higher number of failures to best respond to undominated actions in RM, which would push best response down in the direction of BASE. This is not the case. When considering only the best-response behavior to Nash and $L1$, we find best response rates of about 90 percent in all treatments. We can further support our result of equal best-response rates in BASE and RM by a Kolmogorov-Smirnov Test which compares the number of best responses of each subject. The test yields a p-value of $p > 0.7$.[18] For these reasons we consider the evidence for repeated-game strategies as weak.

Result 3 *(i) Row player: The best-response rates in PI and NF are on average significantly lower than in BASE. While the number of best responses increases over time in BASE, there is no significant time trend in the remaining treatments. (ii) Column player: Best-response rates are on average higher in treatments NF and PI than in BASE. This difference disappears when restricting attention to undominated actions. There is no significant time trend for any treatment except RM. (iii) For both player roles, treatments BASE and RM do not significantly differ from each other with respect to average best-response rates and time trends.*

3.2.3 Accuracy of stated beliefs

In the Nash equilibrium of the stage game, subjects hold accurate beliefs about their opponent's choice. In the level-k model, however, this is typically not the case as subjects' beliefs can be at odds with true behavior. In order to measure how well stated beliefs predict the opponent's play, we use the earnings from the quadratic scoring rule (QSR). Figure 4 shows the average earnings over three periods from the QSR for all treatments and for both player roles.[19]

A natural benchmark for the accuracy of belief statements is the payoff that subjects receive by stating uniform beliefs, which is 1€. However, the average across treatments for both player roles is about 1€. Remember that we do not observe many uniform belief statements (see also

[18] We use each column player as an independent observation and compare the number of best responses of each column player across BASE and RM.

[19] In principle the accuracy of predicting other's behavior should not depend on the player role. Indeed, we only find a weak significant difference between player roles in RM (Mann-Whitney test, $p = 0.098$). In all other treatments the same test yields p-values higher than 0.45.

	Beliefpay			
	row player		column player	
	(1)	(2)	(3)	(4)
PI	-0.017	-0.030	-0.069	-0.087
	(0.063)	(0.082)	(0.063)	(0.081)
NF	-0.144**	0.240	-0.219***	-0.039
	(0.063)	(0.081)	(0.062)	(0.080)
RM	-0.100	-0.053	-0.216***	-0.138
	(0.067)	(0.086)	(0.066)	(0.085)
Period		0.005		0.013***
		(0.003)		(0.003)
PI*Period		0.001		0.002
		(0.005)		(0.005)
NF*Period		-0.011**		-0.017***
		(0.005)		(0.005)
RM*Period		-0.005		-0.007
		(0.005)		(0.005)
Constant	1.056***	0.999***	1.080***	0.941***
	(0.044)	(0.056)	(0.043)	(0.055)
$\chi^2_{(.)}$	$\chi^2_{(3)} = 6.84^{*}$	$\chi^2_{(7)} = 15.60^{**}$	$\chi^2_{(3)} = 17.61^{***}$	$\chi^2_{(7)} = 55.56^{***}$
R^2	0.07	0.07	0.16	0.16
N	1920			

Notes: Random-effects regression, * Significant at 10-percent level; ** Significant at 5-percent level; *** Significant at 1-percent level.

Table 9: Regression: Accuracy of stated beliefs over time

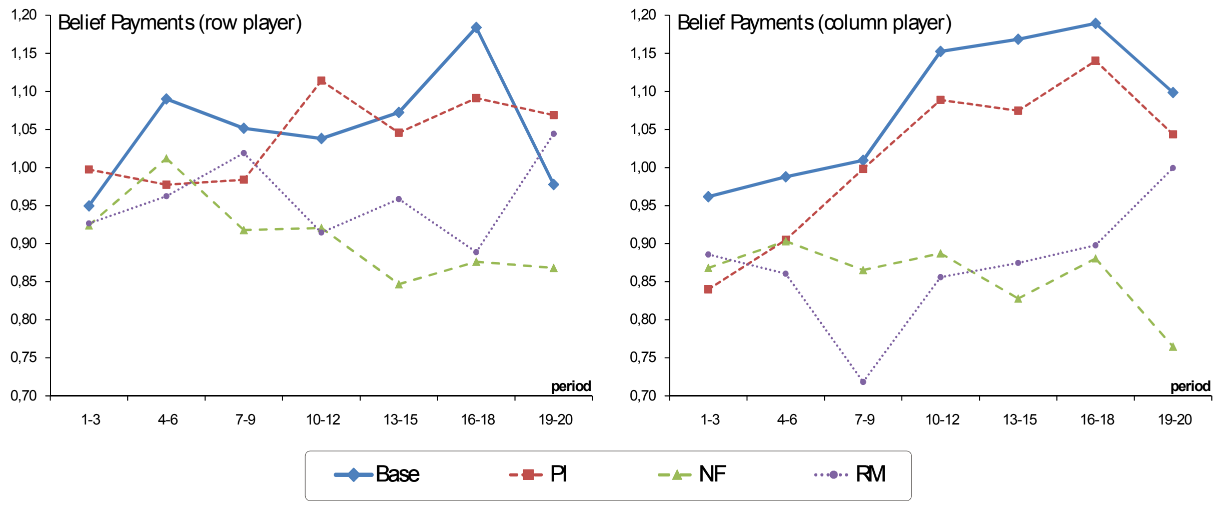

Figure 4: Accuracy of stated beliefs over time

footnote 5). Row players in all treatments earn slightly more than the benchmark of 1€, but we cannot reject the hypothesis of equal means at any conventional significance level (Wilcoxon Signed Rank test, p-values > 0.12). The same holds for column players in BASE and PI, whereas column players in NF and RM earn on average less than 1€ (Wilcoxon Signed Rank test, for NF and RM p-values < 0.01).

Although the accuracy of belief statements is low, we observe improvements in predicting the play of the opponent over time in treatments BASE and PI. Apparently, this is not the case for treatments NF and RM. We run a random-effects panel regression on the earnings from the belief elicitation task. The dependent variable is the payoff from the QSR. Besides the constant and an reference time trend for BASE, the regression includes treatment and time interaction dummies for the controls PI, NF and RM as independent variables in order to measure the corresponding performance relative to BASE. Again, direct tests of the absolute time trends in the control treatments where performed separately and are not reported.

On average beliefs are significantly less accurate in NF than in BASE while PI and BASE show the same accuracy of beliefs. When comparing BASE to RM, only the column players differ significantly from each other due to a lower accuracy of beliefs in RM than in BASE. Focusing on the development over time we observe some learning in BASE since the beliefs become more accurate over time for the column players. The row players show no significant learning path in any treatment. Whereas we observe the same pattern over time in PI and BASE, tests of the absolute time trends reveal that there is no learning at all in NF and RM. In addition, the difference between

BASE and NF is more pronounced than the difference between BASE and RM.

These results indicate that information about the behavior of previous opponents is more useful for subjects to predict the current opponent's behavior than information about the game structure. Further support for the relatively minor role of the opponent's payoffs comes from the fact that in the first rounds beliefs are not significantly less accurate in PI than in BASE. Our data also reveal that there is a positive effect of playing with the same opponent repeatedly on the accuracy of beliefs. We summarize these findings in the following result.

Result 4 *(i) The accuracy of stated beliefs is low, and only very limited learning effects can be observed. Merely in BASE subjects earn more than they would have earned by always stating uniform beliefs. (ii) In BASE and PI subjects reveal the same pattern of learning. The column players best respond more often in later periods while the row players do not. (iii) On average beliefs are significantly less accurate in NF than in BASE since there is no learning at all in NF. (iv) While for the row player the accuracy of beliefs is the same in RM as in BASE on average and over time, the column players exhibit less accuracy on average which is due to less improvements over time.*

4 Summary

This paper reports on an experiment aimed at a better understanding of the development of strategic reasoning over a limited number of rounds. To classify the strategies of the 3x3 normal-form game employed in our study, we used the level-of-reasoning model of Stahl and Wilson. This classification allows us to track strategic play over time. In order to understand the determinants of strategic play, we varied the information available to the players and elicited their beliefs about opponents' play.

First consider behavior aggregated over all periods. We find that both, feedback and information about the payoffs of the opponent have an impact on the average choice and belief structure. Both types of information lead to an increase of non-strategic (L1) and a decrease of efficient (or Rawls) play on average. Moreover, the lack of feedback reduces the average accuracy of beliefs which is not the case for missing information about the opponent's payoff structure. We also find that stated beliefs are a better proxy for the underlying true beliefs than theoretical belief models such as weighted fictitious play or Cournot best response (as in Nyarko and Schotter, 2002).

In addition, we investigate choices and beliefs over time. There is considerable learning in a relatively small number of rounds. In the baseline treatment, subjects exhibit less non-strategic and more Nash play over time. Furthermore in later periods their actions become more consistent with their stated beliefs (significantly so for row players). The accuracy of subjects' beliefs with respect to the opponent's choices is also increasing over time (this is significant for column players). In the no-feedback treatment, neither the strategic sophistication of the choices nor the accuracy and consistency of stated beliefs show any significant time trend. These results support the experimental approach of studying behavior in different one-shot games by using repeated choices of the same subject without giving any feedback. On the other hand missing information about the opponent's payoff has almost no impact on the learning path that we have observed in the baseline treatment. Both, the strategic sophistication of choices and the accuracy of beliefs are not altered in their development over time. Only the increase in best-response rates observed in the baseline treatment is diminished without information about the opponent's payoffs.

This study should be seen as a first step in understanding the development of strategic thinking with the help of stated beliefs in a game. Many issues remain to be investigated, e.g. other games should be used in order to abstract from the specifics of our game. Also, the accuracy and consistency of beliefs over time is by now very little understood and requires thorough empirical scrutiny.

References

[1] Camerer, C., Ho, T., and Chong, J.-K. (2004). A Cognitive Hierarchy Model of Games, *Quarterly Journal of Economics*, 119, 861-898.

[2] Clarke, K. A. (2003). Nonparametric Model Discrimination in International Relations, *Journal of Conflict Resolution*, 47, 72-93.

[3] Clarke, K. A. (2007). A Simple Distribution-Free Test for Nonnested Model Selection, *Political Analysis*, 15, 347-363.

[4] Costa-Gomes, M.A., Crawford, V. and Broseta, B. (2001). Cognition and Behavior in Normal-form Games: An Experimental Study, *Econometrica*, 69, 1193-1235.

[5] Costa-Gomes, M.A. and Weizsäcker, G. (forthcoming). Stated Beliefs and Play in Normal-form Games, *Review of Economic Studies*.

[6] Crawford, V. and Iriberri, N. (2007a). Level-k Auctions: Can a Non-Equilibrium Model of Strategic Thinking Explain the Winner's Curse and Overbidding in Private Value Auctions, *Econometrica,* 75, 1721-1770.

[7] Crawford, V. and Iriberri, N. (2007b). Fatal Attraction: Salience, Naivete, and Sophistication in Experimental Hide-and-Seek Games, *American Economic Review*, 97, 1731-1750.

[8] Ehrblatt, W. Z., Hyndman, K., Özbay, E.Y. and Schotter, A. (2006). Convergence: An Experimental Study of Teaching and Learning in Repeated Games, Working Paper.

[9] Ellison, G. (1994). Cooperation in the Prisoner's Dilemma with random matching, *Review of Economic Studies*, 61, 567-588.

[10] Fischbacher, Urs (2007). z-Tree: Zurich toolbox for ready-made economic experiments, *Experimental Economics*, 10, 171-178.

[11] Gerber, A. (2006). Learning In and About Games, Working paper.

[12] Gneezy, U. (2005). Step-Level Reasoning and Bidding in Auctions, *Management Science*, 51, 1633-1642.

[13] Ivanov, A. (2006). Strategic Play and Risk Aversion in One-Shot Normal-Form Games: An Experimental Study, Working Paper.

[14] Kandori, M. (1992). Social Norms and Community Enforcement, *Review of Economic Studies*, 59, 63-80.

[15] McKelvey R. and Palfrey T. (1995). Quantal Response Equilibria for Normal Form Games, *Games and Economic Behavior*, 10, 6-38.

[16] Kalai, E. and Lehrer, E. (1993). Subjective Equilibrium in Repeated Games, *Econometrica*, 61, 1231-1240.

[17] Nyarko, Y. and Schotter, A. (2002). An Experimental Study of Belief Learning Using Elicited Beliefs, *Econometrica*, 70, 971-1005.

[18] Oechssler, J. and Schipper, B. (2003). Can You Guess the Game You Are Playing?, *Games and Economic Behavior*, 43, 137-152.

[19] Rey Biel, Pedro (forthcoming). Equilibrium Play and Best Response to (Stated) Beliefs in Constant Sum Games, *Games and Economic Behavior*.

[20] Schwarz, G. (1978). Estimating the Dimension of a Model, *The Annals of Statistics*, 6, 461-464.

[21] Vuong, Q. H. (1989). Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses, *Econometrica*, 57, 307-333.

[22] Weber, R. (2003). 'Learning' With No Feedback in a Competitive Guessing Game, *Games and Economic Behavior*, 44, 134-144.

[23] Weber, R. and Rick, S. (2008). Meaningful learning and transfer of learning in games played repeatedly without feedback, Working Paper.

SFB 649 Discussion Paper Series 2008

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

001 "Testing Monotonicity of Pricing Kernels" by Yuri Golubev, Wolfgang Härdle and Roman Timonfeev, January 2008.
002 "Adaptive pointwise estimation in time-inhomogeneous time-series models" by Pavel Cizek, Wolfgang Härdle and Vladimir Spokoiny, January 2008.
003 "The Bayesian Additive Classification Tree Applied to Credit Risk Modelling" by Junni L. Zhang and Wolfgang Härdle, January 2008.
004 "Independent Component Analysis Via Copula Techniques" by Ray-Bing Chen, Meihui Guo, Wolfgang Härdle and Shih-Feng Huang, January 2008.
005 "The Default Risk of Firms Examined with Smooth Support Vector Machines" by Wolfgang Härdle, Yuh-Jye Lee, Dorothea Schäfer and Yi-Ren Yeh, January 2008.
006 "Value-at-Risk and Expected Shortfall when there is long range dependence" by Wolfgang Härdle and Julius Mungo, Januray 2008.
007 "A Consistent Nonparametric Test for Causality in Quantile" by Kiho Jeong and Wolfgang Härdle, January 2008.
008 "Do Legal Standards Affect Ethical Concerns of Consumers?" by Dirk Engelmann and Dorothea Kübler, January 2008.
009 "Recursive Portfolio Selection with Decision Trees" by Anton Andriyashin, Wolfgang Härdle and Roman Timofeev, January 2008.
010 "Do Public Banks have a Competitive Advantage?" by Astrid Matthey, January 2008.
011 "Don't aim too high: the potential costs of high aspirations" by Astrid Matthey and Nadja Dwenger, January 2008.
012 "Visualizing exploratory factor analysis models" by Sigbert Klinke and Cornelia Wagner, January 2008.
013 "House Prices and Replacement Cost: A Micro-Level Analysis" by Rainer Schulz and Axel Werwatz, January 2008.
014 "Support Vector Regression Based GARCH Model with Application to Forecasting Volatility of Financial Returns" by Shiyi Chen, Kiho Jeong and Wolfgang Härdle, January 2008.
015 "Structural Constant Conditional Correlation" by Enzo Weber, January 2008.
016 "Estimating Investment Equations in Imperfect Capital Markets" by Silke Hüttel, Oliver Mußhoff, Martin Odening and Nataliya Zinych, January 2008.
017 "Adaptive Forecasting of the EURIBOR Swap Term Structure" by Oliver Blaskowitz and Helmut Herwatz, January 2008.
018 "Solving, Estimating and Selecting Nonlinear Dynamic Models without the Curse of Dimensionality" by Viktor Winschel and Markus Krätzig, February 2008.
019 "The Accuracy of Long-term Real Estate Valuations" by Rainer Schulz, Markus Staiber, Martin Wersing and Axel Werwatz, February 2008.
020 "The Impact of International Outsourcing on Labour Market Dynamics in Germany" by Ronald Bachmann and Sebastian Braun, February 2008.
021 "Preferences for Collective versus Individualised Wage Setting" by Tito Boeri and Michael C. Burda, February 2008.

SFB 649, Spandauer Straße 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

022 "Lumpy Labor Adjustment as a Propagation Mechanism of Business Cycles" by Fang Yao, February 2008.
023 "Family Management, Family Ownership and Downsizing: Evidence from S&P 500 Firms" by Jörn Hendrich Block, February 2008.
024 "Skill Specific Unemployment with Imperfect Substitution of Skills" by Runli Xie, March 2008.
025 "Price Adjustment to News with Uncertain Precision" by Nikolaus Hautsch, Dieter Hess and Christoph Müller, March 2008.
026 "Information and Beliefs in a Repeated Normal-form Game" by Dietmar Fehr, Dorothea Kübler and David Danz, March 2008.

SFB 649, Spandauer Straße 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".