On the measurement of the predictive success of learning theories in repeated games

by Atanasios Mitropoulos*

Faculty of Economics and Management,
Otto-von-Guericke-Universität Magdeburg, Germany

version 1.0
July 2001

Abstract:
The growing literature on learning in games has produced various results on the predictive success of learning theories. These results, however, were based on various methods of comparison. The present paper uses experimental data on a set of four games in order to check on the robustness of rankings among learning rules across measures. We characterise measures along three dimensions: (i) the scoring rule, (ii) the method of comparison, and (iii) the definition of observations and apply all thus defined measures to 12 learning rules. The results show that rankings are indeed sensitive to the measure used. Furthermore, we point at deficiencies of certain measures that have been applied in the past and suggest the use of simulated data when learning rules are supposed to predict realisations of random variables.

Keywords: learning, experimental games, predictive success, forecasts

JEL classification: C53, C72, C91, D83

* The author would like to thank seminar participants at the Otto-von-Guericke-Universität Magdeburg and the following persons for valuable discussions: Bertrand Koebel, Axel Ockenfels, Frank Silber. Support by the Deutsche Forschungsgemeinschaft is gratefully acknowledged.

1. Introduction

In the past literature comparisons of learning rules have often been made without clear statement about what purpose the investigation serves. Many different researchers have applied many different measures but rarely cared about the impact of their choice of measure on the result. As far as we are aware of, Feltovich (2000) and Erev and Haruvy (2000) are the only works that use more than two different measures. This practice neglects the abundance of conceivable methods of evaluation. The present work shows that the choice of the method of evaluation does very well have an impact on the result of the comparison. Evaluations consist of different components. We specify 3 components, namely

(i) the *measure* that implements the comparison between observations and predictions (it is equivalent to the *scoring rule*, if that exists),

(ii) the *method* of comparing observations with predictions and

(iii) the way observations are defined.

For each of these components there are several ways to fill them. The combinations of the elements of these components result in a large number of alternative methods of evaluation of the predictive power of learning rules. We apply all these methods of evaluation to 12 learning rules and analyse the impact of each component on the ranking of learning rules according to their ability to predict experimental data from various games.

The MSD

In the past, the measure that has most often been used to compare the success of learning rules in predicting experimental data is the *mean squared deviation* (*MSD*). This measure sums the squared deviations of prediction probabilities from occurred events and normalizes this value by finally taking the square root. Selten (1998) has argued in favour of this measure since it is the only one that can be generated via a scoring rule that simultaneously fulfils symmetry, elongation invariance, incentive compatibility, and neutrality. As noted earlier by Friedman (1983) the *MSD* has also the virtue of being “effective” with respect to the Euclidean metric, i.e. the measure produces larger (i.e. less favourable) values as the Euclidean distance between prediction and observation increases. The *MSD* has often been applied to experimental data by way of computing the sum of squared deviations between probability vectors and the according unit vectors describing the observations made, thereby ignoring the fact that the favourable properties apply only to comparisons between probability distributions and probability distributions. As we will argue in section 6.1 the application of the *MSD* to dependent realisations of random variables causes several problems, the most important of which is that among two learning rules that produce the same expected hit rates, it selects the one that makes more probabilistic predictions close to the uniform distribution over states.

There are several ways out of this problem. Purists would still argue in favour of the MSD, with the difference that it should be applied to independent histories of play, i.e. the state to be predicted would be the evolution of decisions for the whole of a *T*-round repeated game. The problem with this approach is that even for a small number of players, a small number of alternative strategies, little informational feedback and a moderate number of repetitions of the game the number of alternative states (i.e. histories) quickly becomes computationally intractable.

Component (i)

The second solution is to use different measures (component i). The *mean absolute deviation* (*MAD*), for example, does not have the above property, since deviations are treated proportionally. It may, thus, be an attractive alternative. Even more so, since it does fulfil the properties of the loss functions as derived from Selten's four axioms for scoring rules. Another way of dealing with predictions is to transform the predicted probabilities into point predictions. This has been done, for example, by Erev and Roth (1998) by using the *proportion of inaccuracy* (*POI*). The authors, thereby, deliberately ignored that by reducing probabilities to point predictions the probabilistic rules were deprived of their very nature. Furthermore, the transformation cuts off valuable information provided by the learning rules. Another alternative measure suffering from the same deficiency is the *Kuipers Score* (*KS*). This measure goes back to Pierce (1884) and is nowadays widely used in meteorological literature. However, it has the advantage of normalising the value of rules that either predict with uniform distribution or constantly predict the same action to zero and the value of perfect forecasts to 1 (see e.g. Gandin and Murphy 1992). In our analysis we will take all of these measures into account. We will have a closer look at the comparison between the *MSD* and the *KS* in section 6.2.

Component (ii)

A third solution involves a different way of transforming the probabilistic statements into point predictions. This is done by looking at component (ii), i.e. by looking at the way predictions are formed and compared to aggregates of the observations. In particular, one may expand the prediction probabilities to a large set of simulated events whose relative frequencies correspond to the predicted probabilities. The comparison would then involve realisations of the predicted probabilities and the actual realisation of the true random variable. Predictions and observations would, thus, be put on equal footing without throwing away valuable information. This is also addressed in section 6.1, where it is shown that for dichotomous choice variables this method of comparison produces the same rankings for the *MSD*, the *MAD*, and the *POI*.

The inconsistency problem of comparing probabilities with realisations of random variables can also be overcome by using aggregates of observations over individuals which are compared with aggregates of probabilistic predictions. The resulting method treats learning rules as predictors of average play. Whether it is useful to use disaggregated or aggregated data as the base is a matter of purpose. Those researchers who are more interested in individual decision making and the investigation of microeconomic dynamics should prefer using the disaggregated data of observations (e.g. Feltovich 2000). Those who are – within a given environment – interested in the general tendency of decisions over time do better taking averages over individuals (e.g. Erev and Roth 1998). Section 6.3 deals with this issue not only at the level of probabilities but also by considering the effect of aggregation on the results derived from simulated predictions.

Component (iii)

A further topic concerns the way observations are recorded and, thus, relates to the way the term "observation" is interpreted. We already pointed at the possibility to take aggregates over individuals as observations. A different problem arises when dealing with repeated games, i.e. with games that are repeatedly played by the same individuals. Strictly speaking the choices of individuals of the same group are not independent, since they receive feedback that is correlated with the actions taken by the other players. One way of addressing this (without falling back into the purist's view outlined above) is to take the group outcome of one period as one observation (see component iii). Experimenters of public good games and coordination games played by cohorts of players started to incorporated this issue long ago. However, since the usual learning environment involves a random matching protocol that complicates matters a lot, this topic has been neglected in the past literature on learning. We will see in section 6.5 that, for our data, this distinction does not have a large impact on rankings between learning rules. However, it is likely that the importance of this aspect rises as the number of players that constitute a group increases.

One Further Issues

Apart from the question of how to overcome disadvantageous properties of some methods of evaluation we may also address another topic that partly has already been recognised in earlier literature.

This topic involves the question of whether learning rules are supposed to use all information that has been gathered during play until the period for which the action has to be predicted, or whether to do predictions for a number of periods beforehand. The two extremes of usage of information illustrate that this question, again, is a matter of research purpose. One may use all information an agent has collected until period t in order to make a prediction for the choice in period $t+1$. This is called the *one-period ahead* prediction. These predictions are more valuable for short-run investigations of adaptive play and are better suited to elicit the cognitive processes of players within a given environment (e.g. Camerer and Ho 1999). The other extreme of making predictions is to completely simulate the course of play from the first period to the last (e.g. Erev et al. 1999). We call this the *complete simulation*. The advantage of this approach is that it better captures whether a learning rule is able to replicate long-run dynamic trends. We incorporated this distinction between types of predictions into our component (ii). The difference between these two approaches is being dealt with in section 6.4.

Related work

Some brief discussions of the impact of the choice of method of evaluation on the comparative performance of adaptive theories of behaviour can be found in the experimental investigations by Erev and Roth (1998), Camerer and Ho (1999), Feltovich (2000), and Chen and Khoroshilov (2000). However, the most closely related work to ours is that of Erev and Haruvy (2000). They show on a dataset of single-person decision tasks that depending on the way the comparison between rules is performed either of three rules may perform best. In particular, they replicate three differing rankings between three rules by using three methods of comparison used in previous literature.

Our study differs from theirs in three respects. First, Erev and Haruvy also consider the way parameter estimations of the respective rules were determined. To the contrary, we simplify the comparison by taking estimations as given. Second, the data on which our comparison is based were taken from an experiment that’s informational conditions do not allow for the use of the experienced weighted attraction learning (EWA) first discussed in Camerer and Ho (1999). Third, the scope of our paper is not to replicate a set of diverse former results on learning rules within a single dataset, but to discuss problems and pitfalls associated with the choice of method of comparison between dynamic rules.

Outline

The remainder of the paper is organised as follows. Sections 2 to 5 are devoted to thoroughly introduce all definitions. Section 2 deals with observations and predictions. Section 3 introduces all elements of the three components, i.e. it presents the notation for aggregates of data, the various methods of comparison and the measures. Section 4 describes the experimental data set. Section 5 defines all learning rules and reports the parameter estimations. Section 6 presents the data analysis while section 7 concludes.

Table 1 gives an overview of all dimensions and levels of comparison.

Component (i): Measure

- mean squared deviation — *MSD*
- mean absolute deviation — *MAD*
- proportion of inaccuracy — *POI*
- Kuipers score — *KS*

Component (ii): Method of comparison

- measure of actual observations and probabilistic one-period ahead predictions $M(Y,P)$
- average measure of actual observations and simulated one-period ahead predictions $M(Y,X)$
- average measure of actual observations and completely simulated actions $M(Y,\overline{Z})$
- measure of observations aggregated over individuals or pairs and average of completely simulated probabilistic predictions $M(\overline{Y},\overline{P})$
- measure of observations aggregated over individuals and average probabilistic one-period ahead predictions $M(\overline{Y},\overline{Q})$

Component (iii): Aggregation level of observations

- individual actions — Y
- group outcomes — Y^O

Game

- mutual fate control — MFC
- fate-control behaviour-control — FCBC
- simple coordination — CO
- matching pennies — MP
- data on all games combined — all

Rule

- Bush-Mosteller — BM
- Mookherjee-Sopher — MS
- Cross — CR
- Börgers-Sarin — BS
- Roth-Erev — RE
- Variant on RE — REL
- Karandikar et al. — KA
- Experimentation — EX
- Sarin-Vahid — SV
- Win-Stay Lose-Change — WSLC
- Win-Stay Lose-Randomise — WSLR
- Randomisation — RAND

Table 1: Overview of Dimensions and Levels of Comparison

2. Notation for observations and predictions

Individual observations

We denote the set of players as J and the number of repetitions of the game as T. We assume that each player can choose among a set of actions A which is finite, is the same for all players, and stays fixed over time[1]. Elements of all these sets are denoted by the corresponding lower-case letters. In each period t of the game player j plays action $a_j(t)$ and receives payoff $p_j(t)$. For later reference we identify the action chosen with the mixed-strategy vector $Y(j,t)$ that assigns probability 1 to the observed action, i.e.

$$Y(j,t,a) = 1\lfloor a_j(t) = a \rfloor$$

whereby 1[.] denotes the indicator function.

Depending on the kind of feedback given to the players, the individual subjective history is given by the vector that subsumes all previous feedback. Since for our exemplary data set it will be the case that subjects knew only about which action they themselves had chosen and which payoff resulted for themselves, the individual history $h_j(t)$ is given by $(a_j(1), p_j(1), \ldots, a_j(t\text{-}1), p_j(t\text{-}1))$[2]. Correspondingly, the set of possible histories in period t is denoted by $H(t)$. Theories of decision making have produced a multitude of rules that do not directly select actions but assign probabilities to actions. For this reason we also need to denote the set of mixed-strategies on the set of actions, $\Delta(A)$. For each period, a decision rule L maps the individual history into the set of mixed-strategies, i.e. $L(t): H(t) \rightarrow \Delta(A)$. Because of their adaptive nature such decision rules are usually called *adaptive rules* or *learning rules*. Note that rules that make point predictions are simply mapping into unit-vectors.

Observations on group outcomes

In experiments on repeated games the J players are grouped into K groups of I players that interact T times with each other. Group $k \subset J$ is identified by its members, and the members are each assigned a position $i \in I$ by the function $pos(k,j)$. Similarly to individual actions we refer to the outcome of a group k within period t by the vector $Y^O(k,t)$ that assigns probability 1 to the outcome actually observed and probability zero to all other A^I-1 possible outcomes, i.e. for any $o = (a_1,...,a_I) \in A^I$

[1] We may easily generalise to player-specific and time-varying action spaces.
[2] Again we may easily generalise to accustom different informational settings.

$$Y^O(k,t,o)=\prod_{j\in k}1\left[a_j(t)=a_{pos(k,j)}\right]$$

Comparisons in previous studies have not always measured the predictive success on single observations, but sometimes the predictive success on aggregates, particularly on the average choice probability, whereby the average was taken over all individuals. In our notation this aggregate observation over all individuals will be denoted as

$$\bar{Y}(t,a)=\frac{1}{J}\sum_{j=1}^{J}Y(j,t,a).$$

Correspondingly, we will aggregate outcomes over groups and denote the result by

$$\bar{Y}^O(k,t,o)=\frac{1}{K}\sum_{k=1}^{K}Y^O(k,t,o).$$

Probabilistic one-period ahead predictions

There are several ways to make predictions. The first and widely applied rule is to use the actual observations to form the individual history $h_j(t)$ for a player j at time t and to make a probabilistic prediction $P(j,t)$ for this period via the applied learning rule, i.e. $P(j,t)=L(h_j(t))$ is the prediction vector that's elements sum to one. Just as before we may also aggregate these one-period ahead predictions over players and get $\bar{P}$. Correspondingly, we may make one-period probabilistic predictions for outcomes of group interactions P^O whereby the histories of all players determine the outcome probabilities of that period. The aggregate over groups is then denoted by $\bar{P}^O$.

Simulated one-period ahead predictions

An alternative way of using one-period ahead prediction probabilities is to directly carry out the so-defined random variables and to use the realisations of the random prediction instead of the probabilities. Since some measures will prove not to be linear, measuring the distance between actual observations and prediction probabilities will make a difference to averaging over the identically produced distances between actual observations and simulated prediction realisations. We denote the s^{th} simulated one-period ahead prediction realisation of actions by X_s and the according one-period ahead prediction of group outcomes by X_s^O.

Completely simulated predictions

Predictions may be made for more than only one period ahead. The learning rules may be applied to forecast two, three or maybe more periods beforehand. We chose to use only the two extremes of this predictive ability of the learning rules. Apart from the one-period ahead prediction mentioned above we also examine the predictive power of complete *T*-round predictions of subjects' play. We, hence, simulated predictions which are based on the history of simulated choices of an individual and not on the actually observed choices. For this purpose we need to simulate the complete *T*-period repeated interaction of all the *I* group members. From a single simulation of choices among group members we may either record the individual choice probabilities or the actual choices. In case of probabilities the s^{th} such simulation is denoted by $\overline{Q}_s$, and in case of the simulated choices we get $\overline{Z}_s$. The simulations may also be used to produce predictions for group outcomes. And again we may record either outcome probabilities $\overline{Q}_s^O$ or the simulated outcomes $\overline{Z}_s^O$. Note that when simulating a complete history of play, characteristics of the history of subjects are ignored. In order to bear this in mind the "bar" in this notation signifies that the corresponding predictions are not indexed by individuals or groups, just as the average observations $\overline{Y}$ and $\overline{Y}^O$ are not indexed by individuals or groups.

3. Measures and Aggregates

The MSD and aggregates of data and predictions

In the past literature only few measures for the predictive success of probabilistic learning rules have been used. The most prominent is probably the *mean squared deviation* (MSD; also called the *quadratic scoring rule*) which has first been described by Brier (1950) and has extensively been discussed by Selten (1998). In broad terms the MSD is being defined as the mean of the squared difference between prediction and observation. However, since for experimental observations we may use different aggregates of the data and for the learning rules we may use different aggregates of predictions, the MSD measure may be implemented in various ways. The most common way of using it (see e.g. Tang 1998, Feltovich 2000, Chen and Khoroshilov 2000) is to calculate the mean of the mean squared difference between actually observed actions and probabilistic one-period ahead predictions. For reasons of normalisation one should finally take the square root of the result. With our notation, using individual observations *Y* and individual one-period ahead predictions *P*, we may write this measure as:

$$MSD(Y,P) = \sqrt{\frac{1}{JTA}\sum_{j=1}^{J}\sum_{t=1}^{T}\sum_{a=1}^{A}\left(Y(j,t,a) - P(j,t,a)\right)^2}$$

Note that within this definition the actual observations *Y* are represented as probability distributions which place all probability on the actually chosen action.

Instead of using the probabilistic predictions *P* we may use the corresponding realisation of the implicitly defined random variable *X*. The complete measure is then defined as the average of the *MSD*s between actual observations and the *S* simulated realisations of predictions[3].

$$MSD(Y,X) = \frac{1}{S}\sum_{s=1}^{S}\sqrt{\frac{1}{JTA}\sum_{j=1}^{J}\sum_{t=1}^{T}\sum_{a=1}^{A}\left(Y(j,t,a) - X_s(j,t,a)\right)^2}$$

We may not only compare actual observations with one-period ahead predictions but also with predictions $\bar{Z}$ which simulate the realised history of a pair in which both players use the same learning rule. The resulting measure is

$$MSD(Y,\bar{Z}) = \frac{1}{S}\sum_{s=1}^{S}\sqrt{\frac{1}{JTA}\sum_{j=1}^{J}\sum_{t=1}^{T}\sum_{a=1}^{A}\left(Y(j,t,a) - \bar{Z}_s(t,a)\right)^2}$$

Note that a complete simulation of the history of a pair $\bar{Z}_s$ is independent of the characteristics of the actually observed pairs *j*.

There have been some arguments as to whether predictions are meant to predict individual behaviour or aggregate behaviour. We may apply predictions to different aggregates of data, for example aggregates over pairs or aggregates over periods. Similar to using individual actions we may compare aggregates over individuals $\bar{Y}$ with aggregates over completely simulated paths of actions. Erev et al. (1999) have already pursued this approach. However, they calculated the average of simulated actions while we use the computationally more efficient way of calculating the average of simulated choice probabilities $\bar{Q}_s$, i.e.

$$MSD(\bar{Y},\bar{Q}) = \sqrt{\frac{1}{TA}\sum_{t=1}^{T}\sum_{a=1}^{A}\left(\frac{1}{J}\sum_{j=1}^{J}Y(j,t,a) - \frac{1}{S}\sum_{s=1}^{S}\bar{Q}_s(t,a)\right)^2}\ .$$

Finally, one might compare aggregates over individuals with aggregates over one-period ahead probabilistic predictions $\bar{P}$:

$$MSD(\bar{Y},\bar{P}) = \sqrt{\frac{1}{TA}\sum_{t=1}^{T}\sum_{a=1}^{A}\left(\frac{1}{J}\sum_{j=1}^{J}Y(j,t,a) - \frac{1}{J}\sum_{j=1}^{J}P(j,t,a)\right)^2}\ .$$

We may calculate the same five measures using outcomes of groups instead of individual actions. The according measures $MSD^O(Y^O,P^O)$, $MSD^O(Y^O,X^O)$, $MSD^O(Y^O,\bar{Z}^O)$, $MSD^O(\bar{Y}^O,\bar{Q}^O)$, $MSD^O(\bar{Y}^O,\bar{P}^O)$ would differ in two ways. First, one would take the average

[3] We used *S* = 10,000.

over outcomes $o \in O$ instead of actions $a \in A$, and second, one would average over groups $k \in K$ instead of individuals $j \in J$.

MAD, POI* and the *Kuipers Score

We further calculate all these measures not only using the *MSD* but also using the mean absolute deviation, *MAD*, i.e. instead of using the squared difference between observation and prediction, we take the absolute difference[4], and the proportion of inaccuracy, *POI*, which has already been applied to learning theories by Erev and Roth (1998) and Feltovich (2000). The *POI* measure transforms probabilistic predictions into point predictions by way of treating the most probable event as *the* predicted event. For our dataset this means that if one of the two possible actions is predicted with probability higher than 0.5, then this action is assigned probability 1, while the other action is assigned probability 0. In case both actions are predicted with equal probability (which is typically the case in period 1) then each action is assigned probability 0.5. When examining outcomes the *POI* measure transforms the probability distributions over the four possible outcomes into unit vectors that assign all probability to the most probable outcome. Ties are again broken by assigning equal probability to all most probable outcomes. The measure then reports the mean number of wrong predictions, whereby the mean is taken over all observations. Note that the *POI* measure requires point predictions as well as pure strategy observations. Hence, for comparisons involving aggregates of observations, i.e. $\bar{Y}$ or $\bar{Y}^O$, we transform the relative frequencies into pure strategy observations by using the threshold of 0.5, just as probabilistic predictions are rendered deterministically. The same applies to the next measure.

Meteorologists usually use a different measure for assessing the predictive power of a dynamic theory. The Kuipers score, *KS*, for two events with base probability of ½ each is defined as the difference between the proportion of correct predictions of an event and the proportion of false predictions when the alternative event occurred. Even though this measure has some desirable properties, it lacks applicability to more than two states of the world. For our dataset this means that we can apply the *KS* to the prediction of individual actions, but we are unable to specify an appropriate generalisation that allows us to use this measure for the four possible group outcomes[5]. We define the Kuipers score for observed actions and probabilistic predictions as

[4] Of course, the normalisation by way of taking the square root has to be dropped.

[5] Gandin and Murphy (1992) show that the *KS* is the only equitable rule for two-state predictions. They show, that equitability imposes necessary constraints on the scoring rule that for predictions on more than two states do not suffice to characterize the rule. Thus, we could have generated a scoring rule that shares the equitability

$$KS(Y,P) = \frac{1}{J}\sum_{j=1}^{J}\left(\frac{T_{BB}(j)}{T_{BB}(j)+T_{BA}(j)} - \frac{T_{AB}(j)}{T_{AB}(j)+T_{AA}(j)}\right)$$

where $T_{ab}(j) = \#\{t \in T \mid (a_j(t) = a) \wedge (P(j,t,b) > 0.5)\}$[6].

Note that, as are all the other measures, the *KS*, too, is invariant to a relabeling of actions. For probabilistic observations it also fulfils equitability as defined by Gandin and Murphy (1992).

In the following, when referring to a method of comparison, say the fit of *P* to *Y*, without specifying the measure we will denote this by the letter *M*, i.e. *M*(*Y*, *P*) in the above case.

4. The Data Set and the Games

The data set is taken from an experiment on behaviour under little information[7]. In each of the 12 sessions 10 subjects were randomly paired to play T = 40 repetitions of a randomly assigned 2x2-game, i.e. A = 2 and I = 2. After the first 40-period repetition, subjects were randomly re-matched another three times to pairs to each play another 40 repetitions of another randomly assigned game. Each session, thus, provided data on 20 pairs, which totals to k = 240 pairs over all sessions. Subjects were not informed of what game was actually being played. Instead, they were provided a probability distribution over four games from which the actual game was randomly drawn. Without further knowledge of the payoff matrix subjects repeatedly made decisions on the two available actions (labelled A and B). They knew, however, that repetitions of the game would end after 40 periods. After each period subjects were only told their own payoff and were neither informed of the action chosen by their opponent nor of the payoff received by their opponent. The design was chosen such that information was minimised to the amount necessary to perform simple adaptation. The four games involved were *mutual fate control* (*MFC*), *fate-control behaviour-control* (*FCBC*), the simple *coordination game* (*CO*), and the *matching pennies game* (*MP*). Table 2 displays the according payoff matrices of the games:

At the beginning of each session subjects were told that each of the four games could occur with equal probability and that their position as player 1 or player 2 as well as the actual labelling of the actions was determined randomly with equal probabilities.

properties and can be applied to predictions of four states. Since this iterature is not well developed yet, we refrain from determining our own generalisation of the *KS*.

[6] Within the computations, ties in prediction probabilities, i.e. p=0.5 for either strategy, are broken in favour of action 1.

[7] More details on the experimental design can be found in Mitropoulos (2001).

MFC

Player 1 \ Player 2	A	B
A	0, 0	1, 0
B	0, 1	1, 1

FCBC

Player 1 \ Player 2	A	B
A	1, 0	0, 0
B	0, 1	1, 1

CO

Player 1 \ Player 2	A	B
A	1, 1	0, 0
B	0, 0	1, 1

MP

Player 1 \ Player 2	A	B
A	1, 0	0, 1
B	0, 1	1, 0

Table 2: The Four Games: Mutual Fate Control (MFC), Fate-Control Behaviour-Control (FCBC), Coordination (CO), and Matching Pennies (MP)

Each game has certain characteristics that may be crucial to the evaluation of the predictive power of a learning theory. *MFC*, for example, is a game in which all mixed-strategy profiles are Nash-equilibria. Nevertheless, there is only one strategy profile (B,B) that can be assumed to be stable. *FCBC* still has an infinite number of Nash-equilibria and only one efficient cell, but contrary to *MFC* there is one player (player 1 in table 2) having direct influence on that player's own payoff. The *CO* game has two pure strategy Nash-equilibria. And *MP* has only one Nash-equilibrium in mixed strategies. As Mitropoulos (2001) reports, in *MFC* and *FCBC* about 50 percent of the pairs eventually ended up coordinating on the efficient cell, while for the *CO* game coordination on efficiency occurred for about 90 percent of the pairs. In *MP*, as expected, none of the pairs coordinated on any particular cell. Except for the strategic similarity between *MFC* and *FCBC* all games show diametrically different characteristics. They led to more or less volatility in subject behaviour. For this reason, we may expect certain adaptive rules to perform better in some games while others perform best in other games. However, we do not refrain from checking on the predictive power of rules for the entire set of data.

5. The Learning Rules

We consider twelve different rules that allow for adaptive behaviour based solely on feedback about the own payoff. For 9 of the twelve rules we use a parameterised version of the rule and estimate two parameters using all observations from players labelled 1. We deliberately chose to use two parameters to give each of the learning rules the opportunity to be calibrated to the specifics of the data set. The number of parameters is a compromise between the need for degrees of freedom which in the past have proven to be useful for a reasonable fit of certain learning rules to the data and the lack of such a need for certain other learning rules. We fixed the number of parameters to two in order to base the comparison on equal footing and to avoid another dimension of comparison on the level of information criteria. The last three decision rules are free of parameters and thus provide benchmarks for comparison. Since the experiment did not allow to infer any advantage of one action over the other before play began, we initialised all learning rules to choose both strategies with equal probability.

In the following, for the sake of exposition we use $p_i^a(t) = P(i,t,a)$ denotes the probability of action *a* being chosen by player *i* at period *t*, while a and b denote the parameters to be estimated.

Bush-Mosteller (BM)

The most prominent early formulation of an adaptive rule that does not need information on payoffs or actions of opponents is that by Bush and Mosteller (1955). The basic idea of this rule is that an action is played with higher probability in the next period, if the action was chosen and resulted in success or if the alternative action was chosen and resulted in failure.

$$\begin{aligned} p_i^B(t+1) = p_i^B(t) + a^{BM} \cdot \left(1 - p_i^B(t)\right) \cdot 1\begin{bmatrix} a_i(t) = B \\ \wedge p_i(t) = 1 \end{bmatrix} \cdot \left(- p_i^B(t)\right) \cdot 1\begin{bmatrix} a_i(t) = A \\ \wedge p_i(t) = 1 \end{bmatrix} \\ + b^{BM} \cdot \left(1 - p_i^B(t)\right) \cdot 1\begin{bmatrix} a_i(t) = A \\ \wedge p_i(t) = 0 \end{bmatrix} \cdot \left(- p_i^B(t)\right) \cdot 1\begin{bmatrix} a_i(t) = B \\ \wedge p_i(t) = 0 \end{bmatrix} \end{aligned}$$

Mookherjee-Sopher (MS)

Mookherjee and Sopher (1994) use the same underlying idea but a different parameterisation. Instead of using separate parameters for the case of success or failure, the MS scheme parameterises the extend of reinforcement separated according to whether a particular action had been vindicated or refuted.

$$\begin{aligned} p_i^B(t+1) = p_i^B(t) + a^{MS}\left(1 - p_i^B(t)\right) \cdot 1\left[\begin{bmatrix} a_i(t) = B \\ \wedge p_i(t) = 1 \end{bmatrix} \vee \begin{bmatrix} a_i(t) = A \\ \wedge p_i(t) = 0 \end{bmatrix}\right] \\ + b^{MS} p_i^B(t) \cdot 1\left[\begin{bmatrix} a_i(t) = B \\ \wedge p_i(t) = 0 \end{bmatrix} \vee \begin{bmatrix} a_i(t) = A \\ \wedge p_i(t) = 1 \end{bmatrix}\right] \end{aligned}$$

Cross (CR)

The Cross (1973) dynamic is also based on probabilistic reinforcement but uses adjustments based on the amount of payoff received.

$$\begin{aligned} p_i^B(t+1) = p_i^B(t) + \left(a^{CR} \cdot p_i(t) + b^{CR}\right)\left(1 - p_i^B(t)\right) \cdot 1[a_i(t) = B] \\ - \left(a^{CR} \cdot p_i(t) + b^{CR}\right) p_i^B(t) \cdot 1[a_i(t) = A] \end{aligned}$$

Börgers-Sarin (BS)

The model by Börgers and Sarin (1997) adds the consideration of an aspiration level *b* to the determination of choice probabilities.

$$p_i^B(t+1) = p_i^B(t) - |\pi_i(t) - b_i(t)| p_i^B(t) + |\pi_i(t) - b_i(t)| \cdot 1\left[\begin{bmatrix} a_i(t) = B \\ \wedge \pi_i(t) > b_i(t) \end{bmatrix} \vee \begin{bmatrix} a_i(t) = A \\ \wedge \pi_i(t) \le b_i(t) \end{bmatrix}\right]$$

with the aspiration level evolving according to

$$b_i(t+1) = \beta^{BS} b_i(t) + (1 - \beta^{BS}) \pi_i(t)$$

The second estimated parameter is the initial aspiration level, i.e. $\alpha^{BS} = b_i(1)$.

Roth-Erev (RE)

Instead of using probabilities directly, Roth and Erev (1995) proposed a reinforcement learning rule based on accumulated propensities:

$$u_i^a(t+1) = \beta^{RE} \cdot u_i^a(t) + \pi_i(t) \cdot 1[a_i(t) = a]$$

The propensities then determine the choice probabilities:

$$p_i^a(t) = \frac{u_i^a(t)}{\sum_{k \in \{A,B\}} u_i^k(t)}$$

The second parameter assesses the initial sum of propensities, i.e. $\alpha^{RE} = u_i^A(1) + u_i^B(1)$.

Variant to Roth-Erev (REL)

Erev et al. (1999) propose a variant to the RE rule that uses adjustments via approximates to average payoffs and payoff variance and uses a logit form for the assessment of choice probabilities. The propensities, thus, evolve according to

$$u_i^a(t+1) = \begin{cases} \dfrac{u_i^a(t) \cdot \left(c_i^a(t) + 0.5 \cdot \alpha^{REL}\right) + \pi_i(t)}{c_i^a(t) + 0.5 \cdot \alpha^{REL} + 1} & if\ a_i(t) = a \\ u_i^a(t) & otherwise \end{cases}$$

whereby $c_i^a(t)$ denotes the number of times action *a* had been played by player *i* until period *t*. The assessment of probabilities then follows

$$\sigma_i^a(t) = \frac{\exp\left(\beta^{REL} \cdot p_i^a(t) / PV(t)\right)}{\sum_{k \in \{A,B\}} \exp\left(\beta^{REL} \cdot p_i^k(t) / PV(t)\right)}$$

whereby the payoff variability *PV* is determined via

$$PV(t+1) = \frac{PV(t) \cdot \left(t + \alpha^{REL}\right) + |\pi_i(t) - PA(t)|}{t + \alpha^{REL} + 1}$$

and the approximate for the average payoff *PA* is given by

$$PA(t+1) = \frac{PA(t) \cdot \left(t + a^{REL}\right) + p_i(t)}{t + a^{REL} + 1}$$

and the initial *PA*(1) is the expected payoff from random choice and *PV*(1) is the expected absolute difference between the obtained payoff from random choice and the average payoff from random choice.

Karandikar et al. (KA)

The model proposed by Karandikar et al. (1998) is a halfway conditionally deterministic process that bears similarity to the win-stay lose-randomise scheme we present below. The KA rule, however, involves an evolving aspiration level. Formally, a simplified version of the original model states:

$$\begin{aligned} p_i^a(t+1) = 1\begin{bmatrix} a_i(t) = a \\ \wedge p_i(t) \geq b_i(t) \end{bmatrix} + h(b_i(t) - p_i(t)) \cdot 1\begin{bmatrix} a_i(t) = a \\ \wedge p_i(t) < b_i(t) \end{bmatrix} \\ + (1 - h(b_i(t) - p_i(t))) \cdot 1\begin{bmatrix} a_i(t) \neq a \\ \wedge p_i(t) < b_i(t) \end{bmatrix} \end{aligned}$$

whereby the aspiration level is given in the same way as in the BS scheme,

$$b_i(t+1) = \left(1 - b^{KA}\right) b_i(t) + b^{KA} p_i(t)$$

and the randomisation function *h*(.) is given by

$$h(x) = \arctan\left(\frac{g}{x^2} + \tan\left(\frac{p}{2} \cdot d\right)\right) \cdot \frac{2}{p}$$

with *g* fixed at 0.2 and ***d*** fixed at 0.1.

The second parameter to be estimated is the initial aspiration level, i.e. $a^{KA} = b_i(1)$.

Experimentation Learning (EX)

In a slightly different version the experimentation learning scheme has been proposed by Mitropoulos (forthcoming). Again, the idea is to basically play the win-stay lose-randomise scheme, but contrary to KA the EX scheme postulates that players record the results of a certain number of trials for each action. The discrete value of an action *a* is given by

$$\begin{aligned} v_i^a(t+1) = v_i^a(t) \cdot \left(1[a_i(t) \neq a] + 1[a_i(t) = a] \cdot 1[p_i(t) = 1] \cdot 1\left[v_i^a(t) = b^{EX}\right]\right) \\ + \left(v_i^a(t) + 1\right) \cdot 1[a_i(t) = a] \cdot 1[p_i(t) = 1] \cdot 1\left[v_i^a(t) < b^{EX}\right] \end{aligned}$$

which is initialised with zero, and the choice probabilities are then determined via

$$p_i^a(t) = 0.5 + \mathrm{sign}\left(v_i^a(t) - v_i^b(t)\right) \cdot \max\left(v_i^A(t), v_i^B(t)\right) \cdot \alpha^{EX}$$

Sarin-Vahid (SV)

The learning rule proposed by Sarin and Vahid (1997) is based on approximates for average payoffs of each action

$$u_i^a(t+1) = \left(\left(1-\alpha^{SV}\right)u_i^a(t) + \alpha^{SV}\pi_i(t)\right) \cdot 1[a_i(t) = a] + u_i^a(t) \cdot 1[a_i(t) \neq a]$$

which are deranged by a normally distributed variable with zero mean *Z*, i.e.

$$\tilde{u}_i^a(t) = u_i^a(t) + Z$$

and then determine the action chosen:

$$a_i(t) = \arg\max_{a \in \{A,B\}} \tilde{u}_i^a(t)$$

The second estimated parameter is the standard deviation of the random variable *Z*, i.e. $\beta^{SV} = \sqrt{\mathrm{var}(Z)}$.

Win-Stay Lose-Change (WSLC)

The win-stay lose-change strategy has become famous because in dilemma situations it turned out to be an evolutionarily very successful simple adaptation rule (Nowak and Sigmund 1993). In our setting it also coincides with the more general concept of learning direction theory (Selten and Stoecker 1986, Selten and Buchta 2000). The rule simply postulates

$$p_i^a(t+1) = 1[a_i(t) = a] \cdot 1[\pi_i(t) = 1] + 1[a_i(t) \neq a] \cdot 1[\pi_i(t) = 0]$$

Win-Stay Lose-Randomise (WSLR)

As argued by Mitropoulos (2001), for the games considered, a randomisation after receiving zero payoff is much more effective in pursuing the efficient cell than is the strict change of action. Therefore, we also investigate the rule win-stay lose-randomise, whereby the randomisation is assumed to occur with equal probability over both actions.

$$p_i^a(t+1) = 1[a_i(t) = a] \cdot 1[\pi_i(t) = 1] + 0.5 \cdot 1[\pi_i(t) = 0]$$

As can easily be verified, this rule is a special case of the KA scheme.

Randomisation (RAND)

As a representative of complete lack of adaptive ability, the rule that assigns equal probability over actions regardless of what happened before will serve as a benchmark.

$$p_i^a(t) = 0.5$$

Estimation Results

It is beyond the scope of this paper to study the method of calibration of learning rules. In the following the estimation results will be taken as given, without questioning the validity or the robustness. The results on the comparison of different learning rules will, however, give rise to discussions about the impact on methods of estimation.

The calibration of the learning rules was done via a maximum-likelihood estimation on the probabilistic one-period ahead predictions for half of the data set. The other half of the data set will serve as the base for predictions. Table 3 shows the estimation results for the 9 rules that involve parameters. All parameters, except one, turned out to be highly significant ($p<0.001$) according to a Wald-Test. The only parameter not being significantly different from 0 (not even to the 10% level) is b^{KA}, the adjustment parameter for the aspiration level. Since the initial aspiration level a^{KA} is very close to 0.5, the resulting KA scheme is statistically indistinguishable from WSLR. This similarity is useful for detecting whether measures sufficiently discriminate between similar and dissimilar rules.

Rule	Alpha	Beta	Log-Likelihood	Rank
BM	0.27	0.12	-4245.53	1
MS	0.22	0.23	-4338.23	2
CR	0.18	0.03	-4406.63	3
BS	0.40	0.06	-4852.49	8
RE	3.00	0.91	-4446.94	4
REL	13.76	11.22	-4618.87	7
KA	0.49	0.00	-8894.75	9
EX	8.00	0.99	-4489.30	5
SV	0.13	0.27	-4503.46	6

Table 3: Estimation Results on the 9 Rules Involving Parameters

The log-likelihood after calibration already gives a measure of the goodness of fit and allows for a comparison between rules. However, as can easily be seen, rules that make almost point predictions do have a considerable disadvantage as compared to those rules that make predictions with moderate probabilities, since few predictions close to unit vectors that turn out to be false dramatically reduce the likelihood of the observations. Selten (1998) has pointed out this over-sensitivity of the log-likelihood measure. We follow him in rejecting this measure as a means to compare different rules with each other and, thus, will not include it in the following analysis.

6. Hypotheses and Results

6.1 Characteristics of the classical *MSD*

Selten (1998) points out the desirable properties of the *MSD* measure when comparing probabilistic observations with probabilistic predictions. The method *MSD*(*Y*,*P*) is the corresponding measure for the closeness of probabilistic predictions to probabilistic disaggregated events. Contrary to this assumption, experimental data usually consist of a number of correlated observations on events that appear to be realisations of dynamically determined random variables. As a consequence, one must admit the structural difference between predictions and observations[8].

One way of dealing with this difference is to compare observations with realisations of probabilistic predictions. The corresponding measure is *MSD*(*Y*,*X*). Representing *Y*, *P*, and *X* as column vectors one easily calculates the difference between these two measures as

$$\begin{aligned}\Delta_{MSD}^{X,P} &= E[MSD(Y,X)] - MSD(Y,P) \\ &= \sqrt{v + 1'P} - \sqrt{v + P'P} \\ &> 0\end{aligned}$$

where $v = Y'Y - 2 \cdot Y'P$ and 1 is the vector consisting of ones. More important than the sign of Δ is its dependence on the structure of the probabilities *P*. The difference is substantially higher if the predictions *P* mainly contain intermediate values (around 0.5 for individual actions, or around 0.25 for group outcomes) than if these predictions frequently contain almost sure predictions, i.e. many entries close to 0 or close to 1. As a consequence, cautious probabilistic learning rules that mainly predict probabilities around the equal distribution may perform well under *MSD*(*Y*,*P*) but badly under *MSD*(*Y*,*X*). Table 4 shows the comparison of ranks for each rule over all measures and for the methods *MSD*(*Y*,*P*) and *MSD*(*Y*,*X*). We restricted the table to show only the results for the aggregation level of individual actions *Y* and the data set containing observations from all games.

Table 4 shows that, indeed, the performance of certain rules varies between *MSD*(*Y*,*P*) and *MSD*(*Y*,*X*). Most intriguing is the jump of the rule WSLC within the measure *MSD* when turning from probabilistic to simulated predictions. On the other hand, most purely probabilistic

[8] The problem that predictions and observations are of different type has been discussed by Bossuyt and Roskam (1987) for static probabilistic choice models.

rules rank worse when using simulated predictions than when using probabilistic ones. A notable exception is the Börgers-Sarin scheme BS.

Actions Y	All games	Rule											
Measure	Method	BM	MS	CR	BS	RE	REL	KA	EX	SV	RAND	WSLC	WSLR
MSD	$M(Y,P)$	2	1	6	3	5	7	8	12	4	11	10	9
	$M(Y,X)$	5	6	9	2	8	10	4	12	7	11	1	3
MAD	$M(Y,P)$	5	6	9	2	8	10	4	12	7	11	1	3
	$M(Y,X)$	5	6	9	2	8	10	4	12	7	11	1	3
POI	$M(Y,P)$	5	3	10	1	9	8	7	12	4	11	2	6
	$M(Y,X)$	5	6	9	2	8	10	4	12	7	11	1	3
KS	$M(Y,P)$	7	3	10	1	9	8	6	11,5	4	11,5	2	5
	$M(Y,X)$	6	5	10	2	9	8	4	11	7	12	1	3

Table 4: Comparison of ranks for all rules over all measures and over methods *M*(*Y*,*P*) and *M*(*Y*,*X*), restricted to levels of actions *Y* and values on all games

As long as we are dealing with bivariate observational data the *MAD* measure does not suffer from this deficiency. This measure produces the same results for both methods of comparison, i.e. $\Delta_{MAD}^{X,P} = 0$. The measures *POI* and *KS* are not quite as stable, but also show less variance between methods than the *MSD*. As illustrated in subsection 6.5 *MAD*'s independence of the method of comparison is no longer true if we switch to group outcomes as observations; that is, to data with more than two possible states.

Furthermore, if the set of events to be predicted is a finite unordered space, we have *MSD*(*Y*,*X*) = *MAD*(*Y*,*X*) = *POI*(*Y*,*X*)[9]. This means that the method based on simulations of one-period ahead predictions has the advantage of being invariant to whether the *MSD*, the *MAD*, or the *POI* measure is being used. And, trivially, this holds even when turning to outcomes as observational data. But the prime argument in favour of using simulations is that the observations are taken as realisations of random variables, which is exactly the way they are treated by the probabilistic theories. To the contrary, the *M*(*Y*,*P*) method, implicitly treats observations as probabilistic events.

6.2 The *MSD* versus the *KS*

In earlier literature it has been noted that the *MSD* and the *KS* differ in the way they treat the prediction that consists of a randomisation with uniform distribution over all actions available, which in our case is the RAND rule (see e.g. Gandin and Murphy, 1992). While the *MSD* gives some positive value the *KS* always normalizes the value to 0. For practitioners in applied fields of research this characteristic of the *KS* may be a valuable feature. For the testing and comparison of theories, however, what value is given to the RAND rule is not as important an issue as its rank compared to other rules.

Actions *Y*	Rule RAND	Game				
Method	Measure	MFC	FCBC	CO	MP	all
$M(Y,P)$	*MSD*	10	10	11	7	11
	KS	11.5	11.5	11.5	11.5	11.5
$M(Y,X)$	*MSD*	11	11	12	11	11
	KS	12	11	12	12	12
$M(Y,\bar{Z})$	*MSD*	3	6	11	1	2
	KS	1	6	10	1	2
$M(\bar{Y},\bar{Q})$	*MSD*	10	5	10	1	9
	KS	12	4		2	1
$M(\bar{Y},\bar{P})$	*MSD*	11	11	11	7	11
	KS	6.5	11.5		7	11.5

Table 5: Ranks of the rule RAND for the aggregation level of actions *Y* and for measures *MSD* and *KS*

Table 5 shows all ranks of the RAND rule for the measures *MSD* and *KS* over all methods of comparison and over all games. We restricted the table to only show ranks for the aggregation level of actions *Y*, because for outcomes we do not have a proper definition of the *KS*. Note that, for the CO game, after aggregating observations over individuals, i.e. $\bar{Y}$, and after transforming the relative frequencies into a point decision we were left with the same observation over all 40 rounds, so the *KS* is not well defined. The table shows that there are only few instances of a significant difference in ranks between the *MSD* and the *KS*. Whether RAND performs rather well or rather badly depends on the method of comparison as well as the characteristics of the data.

The three notable exceptions are implicitly or explicitly discussed below when the performance of rules is investigated for each game separately.

6.3 Disaggregated data *Y* versus aggregated data $\bar{Y}$

Research on adaptive behaviour may serve different purposes. While some researchers may be interested in explaining and predicting individual play others may be purely interested in aggregate numbers. Within our set of methods of comparison there are two pairs of methods that basically do the same, the only difference being the level of aggregation. In particular, $M(\bar{Y},\bar{P})$ is the same as $M(Y,P)$ on aggregated data, and $M(\bar{Y},\bar{Q})$ is the same as $M(Y,\bar{Z})$ on aggregated data. Tables 6 and 7 show these two comparisons on the ranks generated from the data of all games. The usage of data on particular games exhibits a similarly large discrepancy in ranks between levels of aggregation of observations.

[9] Except for the normalisation, which however does not affect ranks.

Actions Y	All games	Rule											
Measure	Method	BM	MS	CR	BS	RE	REL	KA	EX	SV	RAND	WSLC	WSLR
MSD	$M(Y,P)$	2	1	6	3	5	7	8	12	4	11	10	9
	$M(\bar{Y},\bar{P})$	2	1	8	5	9	4	6	12	3	11	10	7
MAD	$M(Y,P)$	5	6	9	2	8	10	4	12	7	11	1	3
	$M(\bar{Y},\bar{P})$	5	1	8	2	10	4	7	12	3	11	9	6
POI	$M(Y,P)$	5	3	10	1	9	8	7	12	4	11	2	6
	$M(\bar{Y},\bar{P})$	3.5	3.5	10.5	7.5	10.5	3.5	3.5	12	3.5	7.5	9	3.5
KS	$M(Y,P)$	7	3	10	1	9	8	6	11.5	4	11.5	2	5
	$M(\bar{Y},\bar{P})$	3.5	3.5	9.5	7	9.5	3.5	3.5	11.5	3.5	11.5	8	3.5

Table 6: Comparison of ranks for all rules over all measures and over methods $M(Y,P)$ and $M(\bar{Y},\bar{P})$, restricted to levels of actions *Y* and values on all game

Actions Y	All Games	Rule											
Measure	Method	BM	MS	CR	BS	RE	REL	KA	EX	SV	RAND	WSLC	WSLR
MSD	$M(Y,\bar{Z})$	8	9	2	11	3	6	4	7	12	5	10	1
	$M(\bar{Y},\bar{Q})$	4	5	8	1	7	11	3	10	6	9	12	2
MAD	$M(Y,\bar{Z})$	8	9	2	11	3	6	4	7	12	5	10	1
	$M(\bar{Y},\bar{Q})$	4	5	9	1	7	11	3	10	6	8	12	2
POI	$M(Y,\bar{Z})$	8	9	2	11	3	6	4	7	12	5	10	1
	$M(\bar{Y},\bar{Q})$	6	9.5	11	6	8	6	2.5	9.5	2.5	12	2.5	2.5
KS	$M(Y,\bar{Z})$	10	4	7	9	5	3	8	1	6	2	12	11
	$M(\bar{Y},\bar{Q})$	8	6	12	8	10	8	3.5	11	3.5	1	3.5	3.5

Table 7: Comparison of ranks for all rules over all measures and over methods $M(Y,\bar{Z})$ and $M(\bar{Y},\bar{Q})$, restricted to levels of actions *Y* and values on all games

The tables show considerable differences in rankings dependent on whether before calculation data were aggregated over players or not. This is true for both comparisons of methods and is independent of the measure used.

We first discuss the comparison between $M(Y,P)$ and $M(\bar{Y},\bar{P})$ from table 6. As noted in section 6.1 the $M(Y,P)$ method treats observations on pure strategy choices the same as probabilistic predictions. Taking the mean of all observations for each period over all players, thus, transforms data into observations of approximates on choice probabilities for a representative agent. As a result, the method $M(\bar{Y},\bar{P})$ can be viewed as the comparison between observed and predicted choice probabilities of a representative agent who is composed from agents with diverse histories. By aggregation, hence, the method regains internal consistency. The consequence for the rankings depends on the measure used. Since *MSD* punishes large deviations stronger than small deviations, rules making point predictions suffer a disadvantage under disaggregated data as compared to the application of the *MSD* after aggregation (REL, KA, WSLR). Conversely, rules that regularly predict probability distributions rather close to uniformity tend to perform worse after aggregation (CR, BS, RE). This observation is not true for the measure *MAD*, since *MAD* treats large deviations proportionally to small deviations. Still, there are differences in ranks between disaggregated data and aggregated data. Correlation coefficients for each measure on the ranks from the two methods even reveal that the *MSD* produces fewer differences (r = 0.85) than *MAD*, *POI*, or the *KS* (0.43, 0.45, 0.57, respectively).

Discrepancies in rankings between aggregation over players are more pronounced for the comparison between methods $M(Y,\bar{Z})$ and $M(\bar{Y},\bar{Q})$ which are based on complete simulations of 40-round play. Correlation coefficients on the differences in rankings between methods reveal that after aggregation the order of rules is significantly affected (r_{MSD} = 0.03, r_{MAD} = 0.01, r_{POI} = -0.23, r_{KS} = -0.22). This finding is confirmed when looking at the corresponding figures for specific games. As an example, table 8 shows this comparison of methods for the *MSD* measure and for each game separately. Note that for $M(Y,\bar{Z})$ and $M(\bar{Y},\bar{Q})$ the measures *MSD* and *MAD* produce the same rankings. The according correlation coefficients between methods vary a lot (r_{MFC} = -0.41, r_{FCBC} = 0.81, r_{CO} = 0.46, r_{MP} = 0.08) with the worst correlation between ranks for the game MFC.

Actions Y	*MSD*	Rule											
Game	Method	BM	MS	CR	BS	RE	REL	KA	EX	SV	RAND	WSLC	WSLR
MFC	$M(Y,\bar{Z})$	9	11	3	10	5	6	4	8	12	7	1	2
	$M(\bar{Y},\bar{Q})$	3	2	7	11	6	9	4	8	1	10	12	5
FCBC	$M(Y,\bar{Z})$	8	9	2	10	3	4	7	6	12	5	11	1
	$M(\bar{Y},\bar{Q})$	9	10	4	12	3	7	2	6	11	5	8	1
CO	$M(Y,\bar{Z})$	8	9	2	11	3	6	4	7	12	5	10	1
	$M(\bar{Y},\bar{Q})$	5	9	1	11	2	3	6	8	4	10	12	7
MP	$M(Y,\bar{Z})$	7	10	6	2	8	5	11	9	3	4	12	1
	$M(\bar{Y},\bar{Q})$	7	9	5	12	4	6	3	8	11	1	10	2

Table 8: Comparison of ranks over all rules, over methods $M(Y,\bar{Z})$ and $M(\bar{Y},\bar{Q})$, and over all games restricted to levels of actions Y and the measure *MSD*

Note that for the game CO the method $M(Y,\bar{Z})$ is not really meaningful, since observations are distributed in a bimodal way. Most often pairs eventually coordinated on either (A,A) or (B,B) with both outcomes being almost equally likely. Simulated players, hence, cannot capture the trend towards a particular action of an individual over time. This deficiency is eliminated when using the according method on aggregated data, i.e. $M(\bar{Y},\bar{Q})$, since both observations as well as predictions are formulated independently of individuals.

Returning to tables 6 and 7 we also note that *POI* and *KS* are better not used on aggregated data because they do not sufficiently discriminate among rules. In our case, after aggregation there are only 40 data points left for comparison. Since aggregated data as well as aggregated predictions are internally highly correlated, these measures by transforming into point predictions and pure strategy observations eliminate too much information.

6.4 One-period ahead predictions *P* versus complete simulations $\overline{Z}$, $\overline{Q}$

When forecasting economic indicators or indices from stock markets it naturally is an important issue for which time horizon predictions are formed. Apart from seasonal influences that have to be taken into account, there is a trade-off between the size of the prediction period and accuracy. When experimentalists use models of adaptive decision making then the same decision upon the size of the prediction interval involves more methodological considerations. While forecasting stocks and bonds or economic indicators serve the sole purpose of making the optimal decision on investment or policy, experimental economics has an additional interest in finding a good model of the underlying cognitive processes. As a consequence, different methods of evaluation may serve different aspects of research in learning. If models are supposed to fit as closely as possible to the actual process of decision making of subjects, then it is best to compare the performance of a decision rule by measuring its efficacy in predicting the immediately following decision, given the history of decisions before. This is represented by our set of one-period ahead predictions *P*. The design of market institutions, however, requires the anticipation of general dynamic trends. This is better captured by predictions of long sequences of dynamic play. For this purpose we chose to also simulate complete 40-round sequences of repeated play with artificial agents. We denote one such sequence as $\overline{Z}$ and the average of a number of such sequences as $\overline{Q}$.

We, first, stay at the disaggregated level of data and look at the difference in ranks among rules between using realisations of the probabilistic predictions $M(Y, X)$ and the completely simulated predictions $M(Y, \overline{Z})$. Table 9 shows the corresponding figures.

Note, first, that for simulated disaggregated data the *MSD* produces the same ranks as *MAD* and the *POI*. Note further that for the reasons mentioned in the preceding subsection we do not present figures for the game CO.

Actions Y	*MSD*	Rule											
Game	Method	BM	MS	CR	BS	RE	REL	KA	EX	SV	RAND	WSLC	WSLR
MFC	$M(Y,X)$	5	6	9	1	7	10	4	12	8	11	3	2
	$M(Y,\bar{Z})$	9	11	3	10	5	6	4	8	12	7	1	2
FCBC	$M(Y,X)$	5	6	7	4	8	10	2	12	9	11	1	3
	$M(Y,\bar{Z})$	8	9	2	10	3	4	7	6	12	5	11	1
MP	$M(Y,X)$	7	5	10	2	9	8	4	12	6	11	1	3
	$M(Y,\bar{Z})$	7	10	6	2	8	5	11	9	3	4	12	1

Table 9: Comparison of ranks over all rules, over methods $M(Y,X)$ and $M(Y,\bar{Z})$, and for each game separately restricted to levels of actions Y and the measure MSD

It is fairly obvious that the method employed has a larger impact on the rank of a rule than the game. Most striking is the good performance of the simple model WSLC over all games, if one-period ahead predictions are considered. For the game MFC performance stays good even when turning to simulated predictions. But for the other games WSLC rather badly predicts the long-run trend of pairs. To the contrary, the more appropriate rule WSLR performs well independent of the game and independent of the way predictions are being generated. A further striking observation concerns the rule RAND. The bad performance of the purely randomising decision rule when predicting individual decisions had to be expected. Also, no surprise causes the good performance of RAND for the game MP, since, by the dynamics of the game, subjects can be expected to be driven to randomising behaviour. However, the intermediate ranks for the two other games show that more sophisticated rules have difficulties in capturing coordination by pairs.

We now have a look at the same figures for the comparison of rules based on aggregated data. Table 10 shows the ranks of rules once for the comparison between aggregate one-period ahead predictions and aggregated data, i.e. $M(\bar{Y},\bar{P})$, and once for the comparison between completely simulated predictions and aggregated data, $M(\bar{Y},\bar{Q})$.

Actions Y	*MSD*	Rule											
Game	Method	BM	MS	CR	BS	RE	REL	KA	EX	SV	RAND	WSLC	WSLR
MFC	$M(\bar{Y},\bar{P})$	5	3	7	1	10	8	2	12	6	11	9	4
	$M(\bar{Y},\bar{Q})$	3	2	7	11	6	9	4	8	1	10	12	5
FCBC	$M(\bar{Y},\bar{P})$	5	2	9	1	10	8	3	12	6	11	7	4
	$M(\bar{Y},\bar{Q})$	9	10	4	12	3	7	2	6	11	5	8	1
MP	$M(\bar{Y},\bar{P})$	2	1	6	4	5	10	8	12	3	7	11	9
	$M(\bar{Y},\bar{Q})$	7	9	5	12	4	6	3	8	11	1	10	2

Table 10: Comparison of ranks over all rules, over methods $M(\bar{Y},\bar{P})$ and $M(\bar{Y},\bar{Q})$, and for each game separately restricted to levels of actions Y and the measure MSD

Comparing tables 9 and 10 with each other we see that after turning to aggregated data the rules that require little sophistication do not perform as extreme as for the disaggregated data. For one-period ahead predictions of aggregated data the BS rule performs particularly well. However, the same rule performs worst when turning to complete simulations. It appears that data on the MFC game produce less differences in ranks (r_{MFC} = 0.38) than the data for the other two games for which correlations of ranks even tend to be negative (r_{FCBC} = -0.34, r_{MFC} = -0.27). We, hence, find that when looking at aggregated data, the incentives of the games are as important a factor for the ranks as is the way predictions are generated.

6.5 Individual actions Y versus group outcomes Y^O

Up to now, in order to assess the predictive power of a learning rule we always used individual actions as the underlying observations. This is in line with general practice in experimental learning literature. However, without loss of information we may also use the group outcome as the observed variable. Similar to the comparison of methods $M(Y,P)$ and $M(Y,X)$ from section 6.1 we get that the $MSD^O(Y^O,P^O)$ as compared to $MSD^O(Y^O,X^O)$ favours predictions close to uniform distributions over predictions close to unit vectors. We have seen that the *MAD* measure did not discriminate between probabilistic rules and learning rules that make point predictions as long as the expected number of hits is the same. However, *MAD* suffers from the same deficiency as the *MSD* measure when turning to outcomes as observational data. That is, $MAD^O(Y^O,P^O)$ will favour probabilistic over point predicting rules as compared to the usage of simulations, i.e. $MAD^O(Y^O,X^O)$. A set of stylised data illustrates this point. We use the simplest possible observation, namely that both players use the same action (action 1) over all four periods. Observations and predictions are displayed in Table 11.

			Prediction			
	Observation		Point predicting rule		Probabilistic rule	
Period	Player 1	Player 2	Player 1	Player 2	Player 1	Player 2
1	1	1	0	1	0.5	0.5
2	1	1	1	0	0.5	0.5
3	1	1	0	1	0.5	0.5
4	1	1	1	0	0.5	0.5

Table 11: Exemplary data for observations, point predictions and probabilistic predictions

The resulting values for the two rules over aggregation levels and over measures for the method *M*(*Y*,*P*) are depicted in table 12.

Measure	Aggregation	Rule: Point predicting rule	Rule: Probabilistic rule
MSD	Actions	0.707	0.500
	Outcomes	0.707	0.433
MAD	Actions	0.500	0.500
	Outcomes	0.500	0.375

Table 12: Values of predictive power for the exemplary data

The result is that among those learning rules that produce the same number of correct predictions on group outcomes the $MSD(Y,P)$ favours those that make predictions close to the uniform distribution. It does even more so when using the observational level of group outcomes rather than individual actions. Differently, $MAD(Y,P)$ does not discriminate between types of rules as long as we deal with individual actions as observations. However, probabilistic rules perform better when applied to the group outcomes. The intuition behind it is that on the level of outcomes the probabilistic prediction scatters probability mass over several possible outcomes and thus reduces the expected penalty of a wrong prediction.

Table 13 shows the corresponding values for the actual data. It is plain to see that the above result for extreme learning rules and extreme data does not carry over to the analysis of our data. The method and the measure do have a larger impact on rankings than whether we use individual actions or group outcomes.

All games			Rule											
Measure	Method	Aggregation	BM	MS	CR	BS	RE	REL	KA	EX	SV	RAND	WSLC	WSLR
MSD	$M(Y,P)$	Actions	2	1	6	3	5	7	8	12	4	11	10	9
		Outcomes	1	2	4	5	6	7	8	12	3	10	11	9
	$M(Y,X)$	Actions	5	6	9	2	8	10	4	12	7	11	1	3
		Outcomes	5	6	7	4	8	10	3	11	9	12	1	2
MAD	$M(Y,P)$	Actions	5	6	9	2	8	10	4	12	7	11	1	3
		Outcomes	5	6	7	4	8	10	3	12	9	11	1	2
	$M(Y,X)$	Actions	5	6	9	2	8	10	4	12	7	11	1	3
		Outcomes	5	6	7	4	8	10	3	11	9	12	1	2
POI	$M(Y,P)$	Actions	5	3	10	1	9	8	7	12	4	11	2	6
		Outcomes	5	3	10	1	9	7	8	11	4	12	2	6
	$M(Y,X)$	Actions	5	6	9	2	8	10	4	12	7	11	1	3
		Outcomes	5	6	7	4	8	10	3	11	9	12	1	2

Table 13: Comparison of ranks over all rules, over methods $M(Y,P)$ and $M(Y,X)$, over measures *MSD, MAD* and *POI* and over aggregation levels, restricted to data from all games

Still, the problems described above again lead us to support the usage of simulated predictions, which do not suffer from similar problems of inconsistency. We suspect, that the impact of the representation of the data as actions or outcomes rises as the number of group members rises and as the number of alternative actions rises, since this would quickly lead to outcome spaces that are much larger and allow probabilistic rules to scatter probability mass even more.

7. Discussion

This work was devoted to eliciting the impact of the choice of the method of comparison on the resulting ranking between learning rules. We found that the choice of method has a large impact on the ranking. The paper helps to clarify the characteristics of the various methods of evaluation and offers ways how to produce robust results. Where methods differ from each other in fundamental ways we argue that each method may be conceived of serving different purposes. Furthermore, we find some applications in earlier literature to suffer from lack of consistency. As a result, we suggest that future research on learning should be more sensitive to methodological issues. Either researchers support their choice of method of comparison by arguments related to the purpose of their study, or they support the robustness of their results by checking on a wide range of alternative methods.

In particular we find that the popular *MSD* measure suffers from some deficiencies and should better not be used in conjunction with a comparison between observational data on individual actions and probabilistic predictions. The classical way around this problem is to use measures that transform probabilistic predictions into point predictions by using a cut-off prediction probability (in our case 0.5). This is done in the *POI* and *KS* measures. This procedure, however, is accompanied by a severe loss of information. Instead, we suggest to use simulations of probabilistic predictions. So, a mean of measures of predictive power between actual observations and simulated predictions can be calculated.

We further find that the method of comparison is crucial for the resulting ranking between rules. It does matter whether one uses disaggregated data or mean values averaged over individuals. It also does matter whether one chooses one-period ahead predictions or complete simulations of play. Experimentalists analysing their data may justify their decision on these two dimensions by pointing at their research purpose. Disaggregated data and one-period ahead predictions are better suited for studying individual cognitive processes, while aggregated data and long-run predictions better elicit the ability of forecasting general dynamic trends. In any case the purpose of the study should be made clear.

For our data, which are generated by repeated interaction of groups consisting of only two players, the question of whether predictions are made for individual actions or for group outcomes seems not to matter much. However, the analysis on exemplary data shows that its importance may be expected to rise when groups consist of more players. Another crucial factor for the comparison of learning rules is the question whether to predict aggregated data or individual data. Some measures perform better in the former task while others are better in the latter.

The scope of this work is not limited to repeated games. The insights gained into the use of measures for the predictive power of dynamic theories carry over to the literature on econometrics and meteorological forecasting. Considering that measures may as well be used for estimation purposes our conclusion is that the estimation of learning theories of the kind used in this study involves a lot of sensitivity towards the employed method of estimation. For example, the maximum-likelihood estimation procedure has long been criticised (see e.g. Friedman 1983). Modern econometric approaches further investigate the variance of predictions based on the variance of parameter estimations given the probabilistic nature of the observations. The usual approach is to use bootstrapping techniques, such as for example in Pascual et al. (2001). At this stage, due to the large number of candidates as methods of evaluation and the large number of learning rules we deem it not promising to follow this approach in the present study.

References

Brier, Glenn W. (1950) „Verification of forecasts expressed in terms of probability“ *Monthly Weather Review*, 78, 1 – 3.

Börgers, Tilman and Rajiv Sarin (1997) "Learning Through Reinforcement and Replicator Dynamics" *Journal of Economic Theory*, 77, 1 – 14.

Bossuyt, Patrick M., and Edward E. Roskam (1987) "Testing Probabilistic Choice Models" *Communication and Cognition* 20, 5 – 16.

Bush, Robert R. and Frederick Mosteller (1955) "Stochastic Models for Learning" New York, John Wiley and Sons

Camerer, Colin and Teck-Hua Ho (1999) "Experience-weighted attraction learning in normal form games" *Econometrica*, 67, 827 – 874.

Chen, Yan and Yuri Khoroshilov (2000) "Learning under limited information" mimeo. December 2000.

Cross, John G. (1973) "A Stochastic Learning Model of Economic Behavior" *Quarterly Journal of Economics*, 87, 239 – 266.

Erev, Ido, Yoella Bereby-Meyer and Alvin E. Roth (1999) "The effect of adding a constant to all payoffs: experimental invetigation, and implications for reinforcement learning models" *Journal of Economic Behavior and Organization*, 39, 111 – 128.

Erev, Ido and Ernan Haruvy (2000) "On the Potential Value and Current Limitations of Data Driven Learning Models" mimeo. April 2000.

Erev, Ido and Alvin E. Roth (1998) “Predicting how people play games: reinforcement learning in experimental games with unique, mixed strategy equilibria” *American Economic Review,* 88, 848 – 881.

Feltovich, Nick (2000) “Reinforcement-based vs. belief-based learning models in experimental asymmetric-information games” *Econometrica*, 68, 605 – 641.

Friedman, Daniel (1983) “Effective Scoring Rules for Probabilistic Forecasts” *Management Science*, 29, 447 – 454.

Gandin, Lev S. and Allan H. Murphy (1992) “Equitable skill scores for categorical forecasts” *Monthly Weather Review*, 120, 361 – 370.

Karandikar, Rajeeva, Dilip Mookherjee, Debraj Ray, and Fernando Vega-Redondo (1998) “Evolvong Aspirations and Cooperation” *Journal of Economic Theory*, 80, 292 – 331.

Mitropoulos, Atanasios (2001) “ Little information, efficiency and learning – an experimental study” mimeo. February 2001.

Mitropoulos, Atanasios (forthcoming) “Learning under minimal information – an experiment on mutual fate control” *Journal of Economic Psychology*.

Mookherjee, Dilip and Barry Sopher (1994) “Learning Behavior in an Experimental Matching Pennies Game” *Games and Economic Behavior*, 7, 62 – 91.

Nowak, Martin and Karl Sigmund (1993) “A Strategy of Win-Stay, Lose-Shift that Outperforms Tit-for-Tat in the Prisoner’s Dilemma Game” *Nature*, 364, 56 – 58.

Pascual, Lorenzo, Juan Romo, and Esther Ruiz (2001) “Effects of Parameter Estimation on Prediction Densities: A Bootstrap Approach” *International Journal of Forecasting* 17, 83 – 103.

Peirce, C.S. (1884) “The numerical measure of the success of predictions” *Science*, 4, 453 – 454.

Roth, Alvin E. and Ido Erev (1995) “Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term” *Games and Economic Behavior* 8, 164 – 212.

Sarin, Rajiv and Farshid Vahid (1999) “Payoff Assessments without Probabilities: A Simple Dynamic Model of Choice” *Games and Economic Behavior*, 28, 294 – 309.

Selten, Reinhard (1998) “Axiomatic characterization of the quadratic scoring rule” *Experimental Economics*, 1, 43 – 62.

Selten, Reinhard and Joachim Buchta (2000) “Experimental Sealed Bid First Price Auctions With Directly Observed Bid Functions” in *Games and Human Behavior: Essays in Honour of Amnon Rapoport*, D. Budescu, I. Erev and R. Zwick (eds.) Lawrenz Erlaaum Ass., Mahwak, NJ.

Selten, Reinhard and Rolf Stoecker (1986) “End Behavior in Sequences of Finite Prisoner’s Dilemma Supergames” *Journal of Economic Behavior and Organization*, 7, 47 – 70.

Tang, Fang-Fang (1998) “A comparative study on the Learning and stability in normal form games: some experimental and simulational results” mimeo. September 1998.

Technical note

All computations have been done using Stata 6.0 and 7.0. Stata is a trade mark of Stata Corp.