National Research University Higher School of Economics

as a manuscript

Chernov V. Grigory

MODELS OF LEARNING IN ECONOMIC EXPERIMENTS

PhD Dissertation Summary

for the purpose of obtaining academic degree

Doctor of Philosophy in Economics

Academic supervisors:

PhD in Economics Alexis Belianin

JEL: C52, C73, C92, D91

Moscow – 2022

Problem description Modern economics increasingly becomes an experimental science. Laboratory and field experiments, natural and quasi-experiments are everywhere in the economics literature. Experimental results have shed light on a wide range of questions, ranging from individual rationality, heuristics and biases, to efficiency and implementation of public policy programs. One of the main goals of experimental research is to improve our understanding of the mechanisms and reasons for making particular economic decisions. Yet progress in this direction so far has been limited. On the one hand, human decisions result from complicated, multi-dimensional processes, most of which remain unobservable or unidentifiable to the observer. On the other, human decision-makers are known to be boundedly rational and have limited abilities to figure out their best strategies even in relatively simple strategic environments, let alone to decode strategic intentions of the other player(s). In the heart of all these problems lies the problem of gradual accumulation of knowledge in the process of interactions, i.e. strategic learning.

The study of strategic learning is the main topic of this thesis. The work consists of three interrelated chapters. In the first, we describe modeling approaches common to theories of learning. In the second, we motivate the problem of experimental comparisons between the theories. Finally, in the thrid chapter we prepare the ground to and introduce a new type of strategy-based learning model in repeated games, and then we evaluate it in comparison to the traditional ones in an example of an experimental game of Rock-Paper-Scissors.

As we will learn from the first chapter, there is a huge pool of prospective models that fall into several large groups. These groups possess particular features: some models care about realized wins, some look only at the behavior of the opponent while ignoring that the opponent may do the same. There is also a hybridization between these models, resulting in “a model zoo”. Further, many of these models are observationally equivalent, and can be substituted for one another without loss of descriptive power. Which models are more useful or more theoretically sound than others, which are – these are open scientific questions. We propose to start answering those questions by finding testable assumptions between classes of models and their specifications. We show that it can be done by simulations because the resulting empirical distributions reflect both finite-sample and asymptotic properties of experimental samples.

We then discuss the common empirical problem for model comparisons – we have no analytical and general way to compare models and to establish whether one model explains behavior better than another or even that it is better in clearly defined conditions (specific experimental game). Moreover, our econometric tools of comparison are lacking power in this context – even when we can say that one model fits the data “better”, it is hard to specify “by how much” and “how robust” is this comparison even for a different realization of the same game, without speaking of generalizability of this conclusion to arbitrary games.

The second chapter covers Formal Theory Approach (FTA) to the experimental problem: how can we identify and estimate the model within the laboratory bounds.

Constraints on possible designs of laboratory experiments are underappreciated by econometric theorists (Basse and Bojinov, 2020). Experiments are naturally constrained by the length of the experimental session and, to a lesser extent, by a number of observations of the specific player after a specific history. We cannot expect our subject to play for an infinite time. We can, however, add more subjects or select a longer or shorter experimental session. Further, we usually can change the game itself, or add new features (e.g. information, or random moves) to the design. Finally, in order to study the motives of real players, we can put our experimental subjects to play against a robot with known Data Generating Process (DGP). Robot player enables us to check more specific hypotheses about our subjects by fine-tuning robot's (DGP), an option unavailable with a human player — and in our third chapter we do exactly this.

Thus the chapters of the dissertation cover all experimental stages, yet it is the second chapter which does the heavy-lifting here, and constitutes the main conceptual contribution of the thesis. Specifically, it sets the criterion for whether we can distinguish different learning models, and proposes an empirical strategy to do so in the context of a particular (Rock-Paper-Scissors) experimental game.

Finally, we run the experiment itself, and analyse its results using our criterion. We conclude that our data provide clear evidence that our specific extension of the classical learning models to strategic learning tracks human behavior demonstrably better than the baseline.

Objectives of the research. Learning the opponent's strategy in repeated games and optimally reacting to it requires time and more complex strategies require more time to learn. Thus, proper understanding and modeling of this process (both theoretical and empirical) are of utmost importance for game theory and economics in general. This learning may take too long while the player is losing most of the time. To insure learning models from such a potential loss, many of them only learn "the latest empirical frequencies of actions", and quickly adapt to them. This has two consequences: a) they can prevent rough manipulation by an opponent and win at least as much as if they know opponent action frequencies in advance (see (Hannan et al., 1957) for details) b) they can't learn complex strategies and therefore can't react optimally to them. It allows the models to be flexible and to play "without losing too much" against any type of opponent. At the same time, it remains unclear whether they can learn the optimal response to even a simple pattern that human subjects easily detect.

As a consequence, we formulate our main research question as follows: "In a laboratory experiment with a repeated game how to check that the subjects can recognize patterns?" Can we distinguish between strategy learning, when a human player tries to recognize the contingent action plan of her opponent, and action-based learning that may produce complex behavior out of just simple actions history?

Thus the main aim of the research is to conduct an experimental and structural econometric assessment of the participants' adaptive response to a strategy with a fixed complexity in a repeated game. This problem was divided into the following tasks.

- Analyze the existing theoretical and empirical approaches to the classification of learning models

- Construct game-theoretic learning models that are able to process existing simple regularities in opponent played sequence of action and react to them considering consequences of current play for opponent's future action.

- Analyze whether popular learning models can be distinguished in an experiment. For that purpose:

- Create a synthetic dataset to test the performance of the maximum likelihood estimator under various conditions
- Formulate the criteria and the procedure for testing models through simulations
- Test the procedure on popular learning models

- Based on the conducted analysis, formulate the criteria applicable to the experimental design, allowing to correctly identify the models on the data
- Develop a laboratory experiment design that meets the above criteria, and that allows to identify the subjects' type of learning
- Conduct simulations according to the previously developed procedure and check that the criteria are met
- Run the developed experiment and obtain structural econometric estimates of the developed learning models
- Determine with the structural assessments which particular class of learners the participants belong to
- Find which model type predicts the subjects' behavior better

Methodology The subject of this dissertation is learning in experimental games from both theoretical and empirical perspectives. We argue that the existing approaches to model learning are restrictive, and in particular, they fail to recognize different degrees of the players' congition abilities in games with more than two strategies. To capture these differences, we develop a generalized concept of strategy-based learning and to test its application, including empirical identification, in a class of simple experimental games. We set up and run an experiment whose results supports the viability of this methodological approach: learning models based on our concept shows superior explanatory power in comparison with the classical models.

The first chapter analyzes the central ideas and the current state of the economic theory of learning in games. Within the framework of game theory, learning can be seen as both an alternative to equilibrium analysis, and as a way to investigate the nature of equilibrium concept(s). Outside of this framework, learning in games (starting from the classical Cournot dynamics) sheds new light on economic interactions, sets interesting theoretical and non-trivial econometric problems, and can be studied experimentally. Learning in games connects economics with other (sometimes unexpected) scientific disciplines: biology, philosophy of rationality, and computer science. The first chapter examines in detail why there are so many learning models, what properties in a dynamic context are crucial, and what are the criteria for the "goodness" of these models. At the end of the chapter, a classification of models of learners based on their crucial properties is presented.

The second chapter is devoted to the question of why it is so hard to study learning even in the laboratory setting, outlining several theoretical and practical concerns (like the limited length of an experimental session). In particular, simulations by (Salmon,

2001) show, in a cross-model (or "blind") testing of several models, the data generated by those models does not correspond to the estimated parameters. Thus, even when the real data generation process is known we cannot distinguish correct models from incorrect ones by looking at the estimates. However, we demonstrate that part of these problems could be resolved through simulations and experimental design. We also present the simulation-based toolbox for testing weak identification for any particular experimental sample.

The third chapter studies learning in a strategic environment using experimental data from the Rock-Paper-Scissors game. In a repeated game framework, we explore the response of human subjects to the behavior of a strategically sophisticated opponent. We model this opponent as a robot that plays a stationary strategy with superimposed noise varying across four experimental treatments. Using experimental data from 85 subjects playing against such a stationary robot for 100 periods, we show that humans can decode its strategies, on average outperforming the random response to such a robot by 17%. Further, we show that the human ability to recognize such strategies decreases with exogenous noise in the behavior of the robot. Further, we fit learning data to classical Reinforcement Learning (RL) and Fictitious Play (FP) models and show that the classic action-based approach to learning is inferior to the strategy-based one. We adapt the criteria from the second chapter and provide specific algorithms for the strategy-based class of learning from the first chapter into a 3x3 game. We also show, using a combination of experimental and post-experimental survey data, that human participants are better at learning separate components of the opponent's strategy than in recognizing this strategy as a whole. This decomposition offers a shorter and more intuitive way to figure out their own best response. We build a strategic extension of the classical learning models accounting for this behavioural fact and calibrate its practical application to our experimental data.

Brief literature review The theory of learning in games originated in the Cournot model and nowadays is a well-developed theory ((Young, 2004); 585. However, its development is hindered by a lack of development of methods at the intersection of experimental inference methods and microeconometrics. While separately they are quite developed and sophisticated, their intersection requires special conditions: advanced experimental designs and taking into account finite samples.

The complexity of this problem is illustrated by several relatively recent works on model selection and testing in learning. First popular learning models on 2x2 games were tested in (McKelvey and Palfrey, 2001) who found that the models fit experimental data extremely poorly when played on some types of games, such as coordination games. A series of tournaments (starting with (Arifovic et al., 2006)) tested the potential difference between data generated by the model and human subjects. Time after time the models did not follow the dynamics similar to humans. Later, literature turned to rethink simple goodness of fit measure as a criterion and researchers began to experiment not only with the composition of models pool but also with the criteria. In ((Erev et al., 2007); (Erev et al., 2010)) tournament, authors started experimenting with out-of-sample predictions and comparing different samples by using aggregated choices in one sample as a predictor to another. A bit different approach was demonstrated by (Mathevet and Romero, 2012), namely the theory of predictive metrics in a game based on average payoffs (started by

(Selten, 1998) but not developed until (Mathevet and Romero, 2012)). All these papers test a pool of models on multiple datasets, but instead of balancing between context and accuracy they prioritize only one of these. Simple model can be generalized to most, but not all contexts and in the remaining contexts they perform abysmally bad. Complex models may fit well in all contexts separately, one by one, but do not generalise across them. The consensus we know today is to move towards the accumulation of large datasets and the development of specific criteria ((Fudenberg et al., 2020); (Fudenberg et al., 2019)).

In our view, however, the accumulation of data may not be sufficient. For example, (Salmon, 2001) shows on 500 synthetic datasets that the common methods do not provide a correct statistical inference. This problem has long been discussed in the econometric literature as "weak identification" (Lewbel, 2019) and is aptly described by ((Morton and Williams, 2010) p. 202) as: "Inspired-By Evaluations of Formal Theory Predictions: When a researcher evaluates a formal theory prediction using a Rubin Causality Model-based approach and assumes consistency with all model imposed assumptions but does not explicitly investigate whether it holds or not." We are aware of only one recent work that tries to find an analytical solution for this problem in the case of a linear dynamic model (Bojinov et al., 2020). We take the-similar approach, namely to find a simulation-based solution to provide an experiment planning tool.

Main findings We present a class of learning models that avoids losing too much against an arbitrary opponent, and at the same time can learn simple conditional strategies of the "win-stay-lose-shift" type (i.e. the strategy that prescribes keeping the same action once it has been successful in the past, and shifting to another action once the current one has been unsuccessful). We have developed such models in the context of Rock-Papers-Scissors game, but it can be without loss of generality extended to any repeated game with interval action space (e.g. setting a price in an oligopoly).

Based on (Salmon, 2001), direct identification testing of the learning model was carried out on the example of the most general class of learning models, known as hybrid Experience Weighted Attraction (EWA) (Camerer and Ho, 1999). This is a generic multi-parameter model which embeds two of the most popular approaches, belief-based (Fictitions Play, FP) and action-based (Reinforcement Learning, RL) as particular cases. Salmon's previous results are reproduced and extended. We confirm and elaborate Salmon's result that point identification in a realistic experimental setting is problematic for EWA, and construct indices to assess identifiability of the learning model. It is shown that if we consider only the basic representatives included in the model (individual points of the hybrid model), then we can point identify them. However, we also show that this approach is not sufficiently rich to distinguish different kinds of learning sophistication even in relatively simple strategic environment of games of complete information with more than two actions. To capture this generalised learning, we draw a distinction between action-based and strategy-based learners, and develop empirically-based criteria to test whether participants can be classified as action-based or strategy-based learners. While the concept of strategy-based learning is not new itself ((Hanaki, 2004); (Ioannou and Romero, 2014)), we are apparently the first to provide formal criteria for identification of such models, and extend it to the space of the games with more than 2 actions. We establish the experimental conditions that satisfy such criteria. To test our approach, we propose the specific experiment design involving a controlled opponent (robot) who is pre-programmed to play a particular strategy unbeknown to the opponent human player. Specific algorithms for a strategy-based class are proposed and formalized as well. A pool of models has been selected for comparison, including representatives of belief-based learning, reinforcement learning, action-based learning, and strategy-based learning. We set up an conduct an experiment satisfying these properties, and run it in several treatments characterised by different levels of noise (probability of random move rather than pre-programmed strategy on the part of the robot). Our experiment involved 85 subjects who made individual decisions: they all played agains a robot over 100 rounds, with the goal to recognize it strategy and 'beat' it under various levels of noise. oTree programme software has been used (Chen e.a., 2015).

Experimental evidence confirms that: (a) many people are capable to defeat our preprogrammed artificial opponent; (b) usually it happens in the span of 30-60 rounds, depending on the noise level; (c) often when subject's behavior shows learning, they can explain what they have learned, typically in belief-based terms; (d) among the three subparts of the best response strategy of the win-stay-lose-shift type, the lose-shift part is more easily recognized.

These observations are not very surprising, but more importantly, we show that the empirical learning patterns in this context are at odds with action-based learning. Strategy-based learning model, by contrast, explains it much better, as confirmed by subject reported strategies, and simulations that allow us to compare how well different models would perform in this experimental design against our robot. In our simulations, we have compared the standard action-based models (namely, Fictitious Play and Reinforcement learning) with strategy-based approach.

Simulation results show that: (a) pre-experimental check strongly suggests that free-parameters models are distinguishable in simulation against robot, (b) action-based models are unable to adapt to simple patterns in actions of the opponent, (c) strategy-based models do adapt to exploit preprogrammed simple pattern strategy, (d) strategic belief-based model recognizes such strategy faster than the strategic reinforcement one.

One more interesting result requires a synthesis of or experimental and simulation results: the speed of human learning is the closest to the strategic reinforcement-based model, action-based models don't learn at all and our strategic belief-based model learns too fast.

As a by-product of this analysis, we conclude that pre-experimental simulations are a sound addition to asymptotic criteria, a sanity check for their applicability. Although they require computational power, they allow any particular design to be verified before the actual costs of the experiment are incurred, which will certainly be a useful tool for any experimental economist and might be indispensable for experiments on learning in games.

Contribution The contribution of this dissertation begins with a survey chapter that reviews and reevaluates the existing learning model classifications and properties of models that follow from them. In addition to writing down approaches specific to this literature or the convergence analysis traditional for the reviews[1], this section deals with the

[1] We should mention the works (Marimon, 1996, Fudenberg and Levine, 1998, 2009, 2016). Erev and Haruvy (2016) outlines a view of this theory that is close to experimental and behavioral economics. Not

cognitive aspects of learning and its representation in models, as well as the connection between the theoretical conceptualization of different properties of learning models and the issues of their empirical testing and comparison of models. Latest is discussed in more detail at the beginning of the second chapter.

Further in chapter 2, we present common approaches to evaluate those models, discuss the weaknesses of those approaches and propose a new way to avoid the largest pitfalls of the existing studies (namely the horse-race approach). In particular, we scrupulously discuss the issue of"weak identification" (Lewbel, 2019) for learning models in conditions having the same sample constraints as (Bojinov et al., 2020). From the general criteria provided by (Matzkin, 2005, 2007) we induce the numerical criterion called "simulation ratio" (SM) specifically for simulations with learning models and test its work in the (Salmon, 2001) setting. Salmon (2001) starts from the fact that the EWA learning model is not correctly identified through the statistical criteria and routines used in the original papers. We go further and show that for a given sample size the EWA learning model cannot be identified at all by any test through information derived from the likelihood function. The proposed procedure and numerical SM indicator are thus available to evaluate any arbitrary set of learning models (both nested models like EWA family and different models) and are used by us to evaluate the identifiability of strategic learning models in Chapter 3.

The work also develops the approach of repeated strategies in learning. Although the approach proposed by (Hanaki, 2004) has already been tested by (Ioannou and Romero, 2014) in 2-by-2 games, its implementation has been limited for using in games with larger action space both in terms of necessity for "model training" [2], as in terms of computational limitations of the subjects' resources. In previous studies (e.g. in Axelrod's tournament) implementation of repeated strategies has been looked at through the prism of evolutionary theory and such strategies are understood as an exhaustive indivisible prescription in different situations. For example, the Tit-for-tat strategy for the prisoner's dilemma can be taken as an instruction how to act in 2 different situations: if the player-opponent cooperates or not. This is why (Ioannou and Romero, 2014) learning models use repeated strategies as the basis and work with strategies in their entirety. We show however that dividing repeated strategies into small component parts (we called it "elementary strategies") has a number of comparative advantages in terms of modeling.

Firstly, conceptually, it allows the learner to construct a complex pattern "on the fly" during the learning process, which sufficiently simplifies the computation and even, as we will show at the end of the first chapter, can be implemented in games with continuous action space.

Secondly, the model comparison may not be trivial, and the simpler conceptualization of repeated strategies is easier to test empirically. For example, let us present two specially simplified models for the Battle of the sexes game: strategic and action-based models. The strategic model knows how to choose between alternating on even periods and alternating on odd periods. A model on action repeats an opponent's action if it was successful (imitation) and if it wasn't, it randomizes. If the actual subjects after some time since the beginning of the repeated Battle of the sexes game cooperate in practice, then for us

to mention the article (Nachbar, 2020), which the most succinctly describes the main results in the field.

[2] i.e. there is no element of online learning, the model has to play with itself before it can predict people

these two models as an explanation would be observationally equivalent.

Exploring this problem in general, in chapter 2 we show how empirically the conceptualization of repeated strategies through “elementary strategies” can be separated from the action-based class of algorithms through experimental intervention. It’s possible if experimenter “freezes” the behavior of one of the players in the pair and equips it with a strategy with a fixed complexity. In the conditions of the laboratory experiment in Chapter 3, we achieve this through the use of a robot opponent (which, however, does not imply that the robot is needed in practice, because through the experiment we investigate the properties of the Lerner-human, and they are identical in the laboratory and outside). We use noise in the actions of the robot in order to break the cycle of winnings in which the player could get “occasionally” in some rounds. We observe an increase in the frequency of successful “elementary strategies” at different periods of the game (from early to late), as well as the verbalization of these strategies in the post-experimental questionnaire. The behavior of the participants demonstrates a correct adaptive response, however, the results of the experiment can be interpreted wider, that “elementary strategies ” are kept by the memory of the participants as separate irreducible elements.

Finally, registered data on learning dynamics which were obtained in this work, methodological analysis of the human reaction to the robot opponent, and a conceptual revision of how players behave in the Rock-Paper-Scissors game (Wang et al., 2014) are valuable in their own right.

List of author’s original articles

- Chernov G. V. How to Learn to Defeat Noisy Robot in Rock-Paper-Scissors Game: An Exploratory Study // HSE Economic Journal. 2020. Vol. 24. No. 4. P. 503-538. doi: 10.17323/1813-8691-2020-24-3-503-538

- Chernov G. V., Susin I. S. Models of learning in games: the review // Journal of the New Economic Association. - 2019. - P. 77. doi: 10.31737/2221-2264-2019-44-4-3

- Chernov G. V., Susin I. S. Heuristics Recognition and Learning in Rock-paper-scissors Game: Experimental Study // Russian Journal of Economic Theory. - 2018. - T. 15. - №. 3. - C. 408-419. doi: 10.31063/2073-6517/2018.15-3.6

- Chernov G. Cheparuhin S. Susin I., , Evaluation of Econometric Models of Adaptive Learning by Predictive Measures / SSRN. Series ”Working Papers”. 2020. doi: 10.2139/ssrn.3658087

Also, the candidate participated in the following international conference with presentations on the topic of the thesis:

- XXI April International Scientific Conference on problems of development of economy and society (Moscow). Presentation: Identification and predictive power of learning models in economic experiments. - 2020.

- The workshop ”Causality in the Social Sciences II” , Germany. Presentation: ”Conditional Learning in Non-Transitive Game: An Exploratory Study”. - 2020.

- iCare 6th International Conference on Applied Research in Economics . Presentation: Heuristics recognition and learning In rock-paper-scissors game: experimental study. -2018.

References

(1) J. Arifovic, R. McKelvey, and S. Pevnitskaya. An initial implementation of the turing tournament to learning in repeated two-person games. *Games and Economic Behavior*, 57(1):93–122, 2006.

(2) G. Basse and I. Bojinov. A general theory of identification. preprint, arXiv, 2020.

(3) I. Bojinov, A. Rambachan, and N. Shephard. Panel experiments and dynamic causal effects: A finite population perspective. preprint, arXiv, 2020.

(4) C. Camerer and H. Ho. Experience-weighted attraction learning in normal form games. *Econometrica*, 67(4):827–874, 1999.

(5) I. Erev and E. Haruvy. Learning and the economics of small decisions. In *The handbook of experimental economics Vol. 2*, pages 638–716. Princeton University Press, 2016.

(6) I. Erev, A. Roth, R. L. Slonim, and G. Barron. Learning and equilibrium as useful approximations: Accuracy of prediction on randomly selected constant sum games. *Economic Theory*, 33(1):29–51, 2007.

(7) I. Erev, E. Ert, A. E. Roth, E. Haruvy, S. M. Herzog, R. Hau, R. Hertwig, T. Stewart, R. West, and C. Lebiere. A choice prediction competition: Choices from experience and from description. *Journal of Behavioral Decision Making*, 23(1):15–47, 2010.

(8) D. Fudenberg and D. Levine. *The theory of learning in games*, volume 2. MIT press, 1998.

(9) D. Fudenberg and D. Levine. Learning and equilibrium. *Annual Review of Economics*, 1(1):385–420, 2009.

(10) D. Fudenberg and D. Levine. Whither game theory? towards a theory of learning in games. *Journal of Economic Perspectives*, 30(4):151–170, 2016.

(11) D. Fudenberg, J. Kleinberg, A. Liang, and S. Mullainathan. *Measuring the Completeness of Theories.* arXiv. org, 2019.

(12) D. Fudenberg, W. Gao, and A. Liang. How flexible is that functional form? Quantifying the Restrictiveness of Theories. arXiv preprint, 2020.

(13) N. Hanaki. Action learning versus strategy learning. *Complexity*, 9(5):41–50, 2004.

(14) J. Hannan, M. Dresher, A. Tucker, and P. Wolfe. Approximation to bayes risk in repeated play. In *Contributions to the Theory of Games*, pages 97–139. Princeton Univ. Press, Princeton, 1957.

(15) C. A. Ioannou and J. Romero. A generalized approach to belief learning in repeated games. *Games and Economic Behavior*, 87:178–203, 2014.

(16) A. Lewbel. The identification zoo: Meanings of identification in econometrics. *Journal of Economic Literature*, 57(4):835–903, 2019.

(17) R. Marimon. Learning from learning in economics. In *Advances in economic theory:the 7th Word Congress. CUP*. European University Institute, 1996.

(18) L. Mathevet and J. Romero. Predictive repeated game theory: Measures and experiments. 2012.

(19) R. L. Matzkin. Identification of consumers: preferences when individuals choices are unobservable. *Economic Theory*, 26:423–443, 2005.

(20) R. L. Matzkin. Nonparametric identification. In J. Heckman and E. Leamer, editors, *Handbook of Econometrics*, pages 5307–5368. Amsterdam, 2007.

(21) R. D. McKelvey and T. R. Palfrey. *Playing in the dark: Information, learning, and coordination in repeated games*. California Institute of Technology, California, 2001.

(22) R. Morton and K. C. Williams. *Experimental political science and the study of causality: From nature to the lab*. Cambridge University Press, Cambridge, 2010.

(23) J. Nachbar. Learning in games. In Springer, editor, *Complex Social and Behavioral Systems: Game Theory and Agent-Based Models*, pages 485–498. Springer, New York, 2020.

(24) T. C. Salmon. An evaluation of econometric models of adaptive learning. *Econometrica*, 69(6):1597–1628, 2001.

(25) R. Selten. Axiomatic characterization of the quadratic scoring rule. *Experimental Economics*, 1(1):43–61, 1998.

(26) Z. Wang, B. Xu, and H. Zhou. Social cycling and conditional responses in the rock-paper-scissors game. *Scientific reports*, 4(1):1–7, 2014.

(27) H. P. Young. *Strategic learning and its limits*. Oxford University Press, Oxford, 2004.