Artificial Agents as an Application to Policy Design: The Market Entry Game

Hernan Bejarano
Penn State University
Agricultural Economics & Rural Sociology
University Park, PA 16802
Email: hdb117@psu.edu

Maciej M. Łatek
George Mason University
Department of Computational Social Science
Krasnow Institute for Advanced Study
4400 University Drive, Fairfax, VA 22030
Email: mlatek@gmu.edu

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association's 2011 AAEA & NAREA Joint Annual Meeting, Pittsburgh, Pennsylvania, July 24-26, 2011

Preliminary version. Please do not cite.

1 Introduction

Experimental economics (ExE) and agent based economics (ACE) offer the opportunity of observing and explaining aggregated behavior that was difficult to predict a priori by the researcher. This difficulty of prediction can be either a result of complex emergent dynamics or scale effects as it is common in ACE simulation or due to non-classical preferences, cognitive biases or framing effects as often observed in ExE experiments. The purpose of this paper is to present a case study of the concurrent application of ExE and ACE and a research methodology behind it.

Complementarity between ExE and ACE approaches have already been discussed. Roth (2002) discussed how simulation can help experimental economics with the problem of generalizing results beyond the limited scale of what can be implemented in laboratory. Kurzban and Houser (2005) discussed the role of controlled participatory experiments in validating agent-based models. This angle has been extended by Duffy (2006) who focused on achieving the external validity of computational models of learning and decision making, including both calibration of algorithms using data from human subject experiments and using simulation to eliminate competing hypothesis about unobservable beliefs, cognition or reasoning processes.

In our work we argue for an additional type of complementarity obtained by an iterative coupling experiments and simulations the researcher can improve both the experimental design and the learning rules applied to the artificial agents. The research cycle is laid out as follows. First, simulations with generic agents are used to identify the best experimental design necessary to answer a set of hypothesis. Afterwards, the same computational environment created for agents is used to conduct experiments with human subjects. Next, acquired data is used to calibrate the behavioral parameters of agents. The cycle concludes with the next round of simulations used to redesign new treatments for human experiments. By following such a process, one minimizes the cost of conducting the human experiments and obtain good estimates for the artificial agents parameters. Additionally, one gathers evidence for robustness of the tested hypothesis. and reinforces the validity of the simulation by a sequence of out-of-sample prediction and recalibration.

We implement such a research process using a derivative of the Market Entry Game (MEG) (Rapoport, 1998). MEG is binary decision game where agents (humans) face a environment with strategic substitutes incentives and multiple equilibria. In this game an additional entrant generates a negative externality to all those taking the entry decision. The MEG data from previous studies defies notions of Nash Equilibrium being played by the individuals, leading to problems with identifying preferences and indicators of strategic or heuristic reasoning that motivated the observed behavior. Given prevalence of real life environments with MEG setup, examples of these are not only entrance into new markets but also traffic congestion (Selten et al., 2007), it is important to find ways in which the market designs and different policies can alleviate the generally observed levels of over entry and persistent stochastic behavior:

This paper first outlines the game and applicable approaches to computational modeling of bounded rationality. Next we present initial results from our previous experiments and compare them to the behavior of a crudely calibrated artificial market. We conclude the paper with a presentation of our predictions of an alternative market design, one planned to be tested in the next round of experiments.

2 The Market Entry Game

In the MEG, individuals' payoffs depend on the market capacity, individual cost, number of other entrants and an entry fee. If they choose to stay out then their payoff is certain, if they choose to enter then their payoffs depend on the number of agents choosing this option. We analyze a version of the MEG payoff function where payoffs for entering are linearly decreasing with the increasing number of entrants. The two possible actions are either enter or stay out (denoted s^1 and s^0, respectively):

- Each market has and intrinsic value I, a capacity K and a fixed entry fee f is imposed;
- Each agent i has cost of entry c_i and reservation price R.
- Ex post payoffs for agent i depend on number of entrants $\|s^{-i}\|$ and his decision in a following way:
  - **Enter** $\pi\left(s_i^1, s_{-i}\right) = I + v * \left(K - s^{-i}\right) - c_i - f$
  - **Stay Out** $\pi\left(s_i^0, *\right) = R$.

The game has multiple Nash equilibria. Denote e^* as the equilibrium number of subjects entering the market when the remaining $N - e^*$ stay out. Given that payoff functions are homogeneous then any combination of agents behavior that satisfies these conditions is an equilibrium. In the canonical MEG, subjects have homogeneous payoff functions with no cost for entry $\forall_i\ c_i = c = 0$ and no entry fees $f = 0$ as well as I=R. In previous experimental work (Rapoport, 1998; Amnon Rapoport and Sundali, 1998; Erev and Roth, 1998) the payoff function and parameters values were such that e^* was usually e^*=K or e^*=K-1. This feature of the canonical MEG led Daniel Kahneman to address the observed aggregated behavior as Magic, given that the observed number of entrants for these values is between K and K-1. Once an entry fee or subsidy has been introduced in the payoff function then the equilibrium e^* may usually be different than K or K-1.

3 Experimental Design and Data

We use the data of new version of the MEG for our analysis. This one is modified from the canonical one in two ways. First, we added a new treatment called Expected Market Entry Game (EMEG) in this treatment the capacity of the market is uncertain. This uncertainty is represented by a random variable that has equal probability for three possible values [1]. Second, different that the canonical MEG this experiment introduced different costly entry fees. The main effect of this introduction is that separates the number of entrants in equilibrium e^* from the capacity of the market K by numbers that are greater than one.

[1]In the MEG treatments the traditional strategic risk presented to players in a simultaneous move game. On the other hand, the EMEG two types of risk are present. The uncertainty regarding the capacity of the market adds to the strategic risk.

We conducted 7 experimental sessions where 84 undergraduate and graduate students participated as subjects at the Interdisciplinary Center for Economic Sciences at George Mason University. Upon arrival to the laboratory the subjects were seated at one of the 12-computer terminal separated by partitions. Communication between subjects was prohibited. Written instructions and a series of record sheet were given to each participant. The participants first read the instructions in private. After they read the instructions the lab coordinator come back and read the instructions out loud. A short hypothetical example and a following quiz was performed. After checking that the participants answered the quiz correctly the experiment started. The experimental sessions lasted approximately two hours. Subjects earned an average of 16.5 dollars in addition to a 5 dollars show up fee.

Each session consisted on 150 rounds of a binary decision environment. The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007). The rounds session is divided in six blocks of 25 rounds. For each block the parameters remain constant and are those displayed in Table 1, but the order of block may be permutated to control for possible artifacts.

The payoff function we implement in our experiment is a representation of the payoff function described in the former section. In each sub-treatment of our experiment, the subjects will face the following payoff function.

- **Enter** $\pi\left(s_i^1, s_{-i}\right) = 15 + 2 * \left(K - \|s^{-i}\|\right) - 2 - f$;
- **Stay Out** $\pi\left(s_i^0, *\right) = 17$.

Contrary to the classical approach, entering the market in our experiments carries an implicit cost. The experimental design checks different values K and f. K is deterministic in MEG treatments or is a random variable in the EMEG treatments. The expected market capacity can either be of one third of the number of participants or two thirds. If the expected value of K is 4 and EMEG condition is active, then the realizations of this random variable can take the values 2, 4 and 6 all with equal probability. For those sub-treatments where expected capacity equals 8, then the momentary market capacity is 6, 8 or 10.

Treatment	R	I	c	K	f	lower RNNE	upper RNNE
1	17	17	2	4	-1	2.5	3.5
2	17	17	2	4	0	2	3
3	17	17	2	4	2	1	2
4	17	17	2	8	0	6	7
5	17	17	2	8	2	5	6
6	17	17	2	8	4	4	5

Table 1: Values of default parameters for all treatments. Last column, RNNE, contains calculations of the Risk Neutral Nash Equilibrium. For each treatment, exactly two RNNE exists.

4 Applicable Behavioral and Learning Models

Traditionally, modeling of procedural rationality has aimed to represent learning processes. T. Brenner in (Tesfatsion and Judd, 2006, Chapter 18) categorizies learning models into non-conscious learning, routine-based learning and belief learning. In recent years, experimental economics and psychology have made increasingly larger contributions to this knowledge, but no model has been found to be uniformly superior to others ((Tesfatsion and Judd, 2006, Chapters 18 and 19)). Learning-based approaches reviewed by Brenner suffer from two common problems. First, they are geared towards situations where each of the agents does not distinguish between individual opponents and treats them as a part of a noisy environment. Second, the applications are most often limited to iterated single-stage (stateless) games.

The field of artificial intelligence offers solutions that do not suffer from these weaknesses, but with limited behavioral support. The first interesting approach is variable learning rate reinforcement learning such as AWESOME Sandholm and Conitzer (2003) and GIGA-WoLF Bowling (2005). Those approaches account for the fact that the optimal policy at any moment depends on the policies of the other agents. A variable learning rate is only a partial solution to the issue of learning a moving target. Another venue is joint action reinforcement learning. Suematsu and Hayashi (2002) and Hu and Wellman (2003) offer frameworks EXORL and Nash-Q for learning in stochastic games[2]. Given a stochastic game, the proposed algorithms converge to a Nash equilibrium when other agents are adaptable, otherwise an optimal response will be played. All those approaches have been demonstrated to lead to better performance than the traditional single agent learning methods in case of strategic interactions.

All multi-agent reinforcement learning algorithms require agents to build internal representations of their environment and opponents, progressively updated as experience is accrued. Those internal models are based on principles of stochastic optimization and do not incorporate any domain specific knowledge. This induces a high sample complexity (measuring duration of "burn-in" phase necessary for agents to become model-consistent), possibly too high for realistic economic modeling applications[3]. This relates to an observation by Costa-Gomes in Costa-Gomes et al. (2009), where he argues that adaptive methods produce good fits to the long-run behavior, but short term dynamics often seem to underestimate rates of human learning, because adaptive methods forego a great deal of background information available to humans and rarely assume that agents perform introspection or counterfactual reasoning.

A way of fixing this issue is belief or fictitious learning. A fictitious player observes (a) the strategies of all players and (b) can calculate the payoffs he would have obtained had he and the other players played any other possible combination of strategies. This requires such a player to have access to an accurate model of the environment. Fictitious players have no manner of anticipating the behavior of opponents and each derives his behavior by myopically best responding to the history. The best understood and tested fictitious learners include models of Experience Weighted Attraction Camerer and Ho (1999); Ho et al. (2004),

[2]A stochastic game is a set of n-agent normal-form games augmented with rules for transitions depended on actions of agents. See (Shoham et al., 2004, Section 2) for the definition.

[3]Y. Shoham raises this point with respect to the Trading Agent Competition (Shoham et al. (2004)).

which also generalize reinforcement learning models.

An interesting extension of the fictitious learning that goes beyond removing myopia of players is presented in Camerer et al. (2002). That model assumes a mixture of adaptive learners and sophisticated players. An adaptive learner adjusts his behavior the EWA way. A sophisticated player rationally best-responds to her forecasts of all other behaviors and can be either myopic or farsighted. A farsighted player develops multiple-period rather than single-period forecasts of others' behaviors and how those would change in response to her initial actions and engages in strategic teaching of others by choosing a COA that gives her the highest discounted net present value.

Last remaining is so called n-th order rationality. A formal definition is presented in Michihiro (1997). An agent is first-order rational if it calculates the best response to his beliefs about strategies of zero-order agents[4] and the state of the world. An agent is n-th order rational if it determines its best response assuming that the other agents are $(n-1)$-th order rational.

The most advanced example of this class of reasoning mechanisms is constituted by the Cognitive Hierarchies (later CH) model, presented in Camerer et al. (2004)[5]. The CH model consists of iterative decision rules for players doing n steps of thinking, and the frequency distribution $f(k)$ of order k players. The iterative process begins with step 0 types who don't assume anything about their opponents and merely choose according to some probability distribution. Step k thinkers assume their opponents are distributed, according to a normalized Poisson distribution (with mean τ), from step 0 to step $k-1$. CH has been validated with human-based experiments where it has been found that $\tau = 1.5$ fits data from many canonical games much better than extant learning-based approaches.

Several papers have already computationally analyzed behavior in the MEG, for example in Amnon Rapoport and Sundali (1998); Erev and Roth (1998). Erev and Rapoport have also documented results from previous experimental studies. It has been found that variants of the basic reinforcement learning model of reinforcement learning accounts for the quick convergence to equilibrium play in market entry games with a large number of agents has been found and explain variations in the population strategies for different treatments. In order to dock our results with those of the previous studies, we will use reinforcement learning approaches called Q-learning Sutton and Barto (1998), but also augment it with two other approaches: CH and EWA. Default parameters for each of those algorithms are listed in Table 2.

5 Experimental and Computational Dynamics

In Figure 1 and Figure 2 we observe how closely the presence and the dynamic of over entry in the experimental data predicted is replicated by the agents with each of the three learn-

[4]As discussed later, the exact choice of behavioral rule for zero-order agent's is usually case-specific and could include random behavior, continuation of historical behavior (our default) or any non-strategic learning rule.

[5]The first review articles with empirical support for CH and related concepts of n-th order rationality as plausible cognitive architecture for individuals appeared in early 2000s, see Crawford and Iriberri (2007) and Costa-Gomes et al. (2001).

Algorithm	Parameter	Default range	Meaning
RL	α	0.25	Learing step size
	τ	1	Payoff sensitivity
	γ	0.75	Discounting factor
CH	τ	1.5	Average depth of recursion
EWA	δ	0.1	Weight of counterfactual observations
	ϕ	0.75	Discounting factor
	ρ	0.8	Observation-experiences discount
	λ	1	Payoff sensitivity

Table 2: Simulation parameters used to calibrate behavioral algorithms.

ing rules. Each learning rule has shown the potential to predict the direction of population-level adaptation in response to treatment switches. Nevertheless, CH approach overestimates the initial speed of adaptation, whereas RL and EWA approaches underestimate the speed and magnitude of movements for an experienced population.

For treatments with non-stochastic capacity of the market the predictions are more accurate. In these treatments the agents are able to closely replicate the variance of the number of total entrants, although the experimental data shows a faster convergence to a small over entry of one subject. The analysis is slightly different for those treatments with stochastic capacity. The introduction of variable capacity K but keeping the mean constant does not changes the expected profit of entrance but it does change the observed profits at each round of the experiment. This feature seems to have a large effect on the human participants and CH agents, making their behavior more erratic than the one observed by the RL and EWA agents, see Figure 4.

6 Conclusions and Future Outlook

We found that the ACE with different learning rules constituted a useful tool for assessing human behavior in an experimental design with varied treatment effects. Introduction of capacity risk seems to alter the human behavior increasing the observed individual heterogeneity, but had little effect on the average number of entrants. Learning algorithms consistently mis-estimated the range of human heterogeneity and had problems with capturing how the speed of population adaptation changes in time.

The fact that the human behavior observed during the experiments presents much more heterogeneity than the one generated by the learning rules calls for an approach that allow for distribution of different agents to be implemented instead of just representative agents. This implementation might not be essential for environments like the one analyzed here where subjects where presented with homogeneous payoff functions. Then on aggregate level, the heterogeneity of human behavior does not have an important effect on the social welfare and humans achieve similar performance as homogeneous computational agents. Nevertheless, we plan to adapt approaches which include heterogeneous populations of computational agents, in particular through the estimation of CRRA utility functions Geweke (2001). We also want to implement and test a different learning rule, one which actively tracts the environment and conditions the learning rates on the it's dynamics.

Lastly, we plan to investigate treatments where individual entry costs are asymmetric. We performed a number of computational experiments to investigate how the amount of cost asymmetry influences the population outcomes, see Figure 5. In line with results of Amnon Rapoport and Gisches (2008); Rapoport et al. (2002), it seems that a careful application of assymetry may help overcome coordination failures. After the computational framework is improved, we plan to use it to design the next round of experiments.

References

Amnon Rapoport, Darryl A. Seale, I. E. and Sundali, J. A. (1998). Equilibrium play in large group market entry games. *Management Science*, 44:119–141.

Amnon Rapoport, Tamar Kugler, S. D. and Gisches, E. J. (2008). Braess Paradox in the Laboratory: Experimental Study of Route Choice in Traffic Networks with Asymmetric Costs. *Decision Modeling And Behavior In Complex And Uncertain Environments*, 21:309–337.

Bowling, M. (2005). Convergence and No-Regret in Multiagent Learning. In *Proceedings of the Annual Conference on Neural Information Processing Systems 2005*, page 209216.

Camerer, C. and Ho, T. (1999). Experience-weighted Attraction Learning in Normal Form Games. *Econometrica*, 67(4):827–974.

Camerer, C., Ho, T., and Chong, J. (2004). A Cognitive Hierarchy Model of Games. *Quarterly Journal of Economics*, 119(3):861–898.

Camerer, C. F., Ho, T. H., and Chong, J. K. (2002). Sophisticated Experience-Weighted Attraction Learning and Strategic Teaching in Repeated Games. *Journal of Economic Theory*, 104(1):137–188.

Costa-Gomes, M., Crawford, V. P., and Broseta, B. (2001). Cognition and Behavior in Normal-Form Games: An Experimental Study. *Econometrica*, 69(5):11931235.

Costa-Gomes, M. A., Crawford, V. P., and Iriberri, N. (2009). Comparing Models of Strategic Thinking in Van Huyck, Battalio, and Beila's Coordination Games. *Journal of the European Economic Association*, **7**:365–376.

Crawford, V. P. and Iriberri, N. (2007). Level-k Auctions: Can a Nonequilibrium Model of Strategic Thinking Explain the Winner's Curse and Overbidding in Private-Value Auctions? *Econometrica*, 70(5):17211770.

Duffy, J. (2006). *Handbook of Computational Economics, vol 2*, chapter Agent-Based Models and Human Subject Experiments, pages 949–1011. Elsevier B.V.

Erev, I. and Roth, A. E. (1998). Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria. *American Economic Review*, 88(4):848–81.

Fischbacher, U. (2007). Z-Tree: Zurich Toolbox for Ready-made Economic Experiments. *Experimental Economics*, 10:171–178.

Geweke, J. (2001). A Note on Some Limitations of CRRA Utility. *Economics Letters*, 71:341–345.

Ho, T., Camerer, C., and Chong, J. (2004). Self-tuning Experience Weighted Attraction Learning in Games. *Journal of Economic Theory*, 133(1):177–198.

Hu, J. and Wellman, M. P. (2003). Nash Q-Learning for General-Sum Stochastic Games. *Journal of Machine Learning Research*, 4(6):1039–1069.

Kurzban, R. and Houser, D. (2005). Experiments Investigating Cooperative Types in Humans: A Complement to Evolutionary Theory and Simulations. *PNAS*, 102:1803–1807.

Michihiro, K. (1997). *Advances in Economics and Econometrics: Theory and Applications, Seventh World Congress*, chapter Evolutionary Game Theory in Economics, pages 244–279. Cambridge University Press.

Rapoport, A. (1998). Individual strategies in a market-entry game. *Group Decision and Negotiation*, 4:117–133.

Rapoport, A., Seale, D. A., and Winter, E. (2002). Coordination and Learning Behavior in Large Groups with Asymmetric Players. *Games and Economic Behavior*, 39(1):111–136.

Roth, A. E. (2002). The Economist As Engineer: Game Theory, Experimental Economics and Computation As Tools of Design Economics. *Econometrica*, 70:1341–1378.

Sandholm, T. and Conitzer, V. (2003). AWESOME: A General Multiagent Learning Algorithm That Converges in Self-Play and Learns Best Response Against Stationary Opponents. In *Proceedings of the Twentieth International Conference on Machine Learning*, pages 506–513.

Selten, R., Schreckenberg, T., Pitz, T., Chmura, and Kube, S. (2007). Commuters' route choice behavior. *Games and Economic Behavior*, 58:394–406.

Shoham, Y., Powers, R., and Grenager, T. (2004). Multi-Agent Reinforcement Learning: A Survey.

Suematsu, N. and Hayashi, A. (2002). A Multiagent Reinforcement Learning Algorithm Using Extended Optimal Response. *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems AAMAS '02*, page 370.

Sutton, R. and Barto, A. (1998). *Reinforcement Learning: An Introduction.* MIT Press.

Tesfatsion, L. and Judd, K. L. (2006). *Handbook of Computational Economics: Agent-based Computational Economics.* North-Holland.

Average of profit		Nash level 1.5	2.5	3.0	4.5	5.5	6.5	Grand Total
Approach	Cog. Hier.	15.3	15.7	15.6	15.6	16.2	16.3	15.8
	EWA	14.5	14.7	15.1	15.7	16.3	17.0	15.5
	Experiment	16.3	15.7	16.0	16.3	16.6	16.6	16.2
	Q-learning	14.4	14.9	15.2	16.2	16.9	17.6	15.8

(a) Average profits

Std. deviation of profit		Nash level 1.5	2.5	3.0	4.5	5.5	6.5	Grand Total
Approach	Cog. Hier.	2.7	2.6	2.7	3.3	3.4	3.4	3.1
	EWA	3.3	3.1	2.9	3.0	3.1	3.3	3.2
	Experiment	2.2	3.1	2.8	2.5	3.0	3.0	2.8
	Q-learning	3.3	3.0	2.8	2.6	3.1	3.3	3.2

(b) Standard deviation of profits

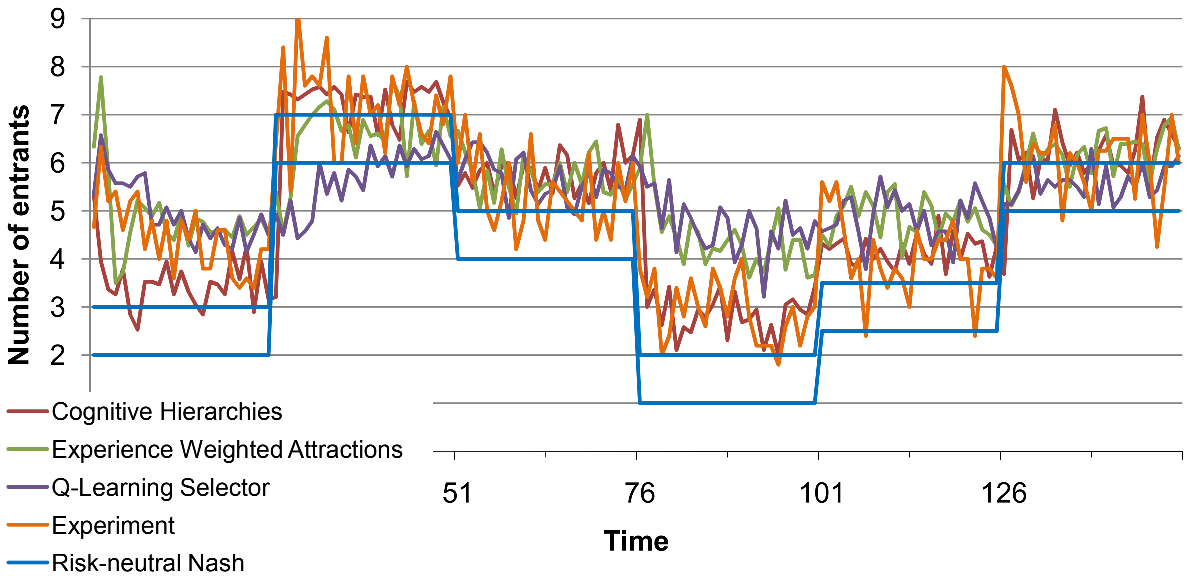
Average number of entrants		Nash level 1.5	2.5	3.0	4.5	5.5	6.5	Grand Total
Approach	Cog. Hier.	3.1	3.5	4.2	5.5	6.1	7.1	4.9
	EWA	4.5	4.9	5.0	5.6	6.2	6.5	5.4
	Experiment	2.9	4.3	4.2	5.3	6.0	7.0	4.9
	Q-learning	4.7	5.0	4.9	5.4	5.7	5.9	5.3

(c) Average number of entrants

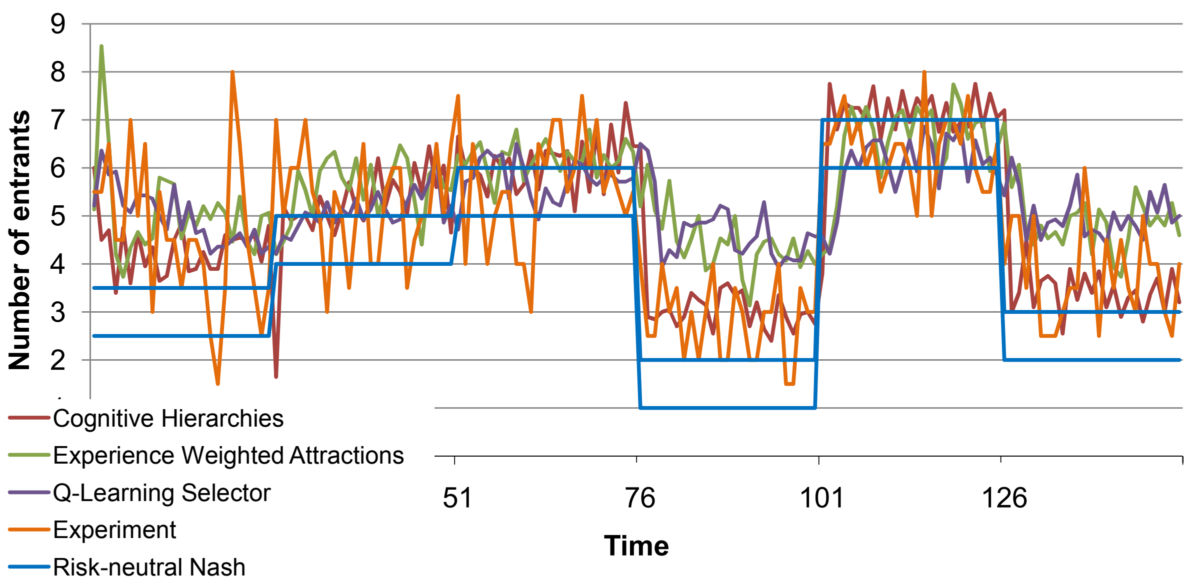
Std. deviation of number of entrants		Nash level 1.5	2.5	3.0	4.5	5.5	6.5	Grand Total
Approach	Cog. Hier.	2.1	1.9	1.8	2.1	2.1	2.0	2.5
	EWA	1.8	1.9	1.8	1.9	1.8	1.9	2.0
	Experiment	1.3	1.7	1.7	1.5	1.6	1.6	2.0
	Q-learning	1.6	1.6	1.6	1.5	1.6	1.6	1.6

(d) Standard deviation of number of entrants

Figure 1: Summary statistics of profits per participant per round and number of entrants.



(a) Treatment order 1, 2, 3, 4, 5, 6.



(b) Treatment order 5, 3, 6, 4, 2, 1.

Figure 2: Predicted and observed average number of entrants for each of the tested treatmet permutations. All time series are simple averages across multiple independent runs.

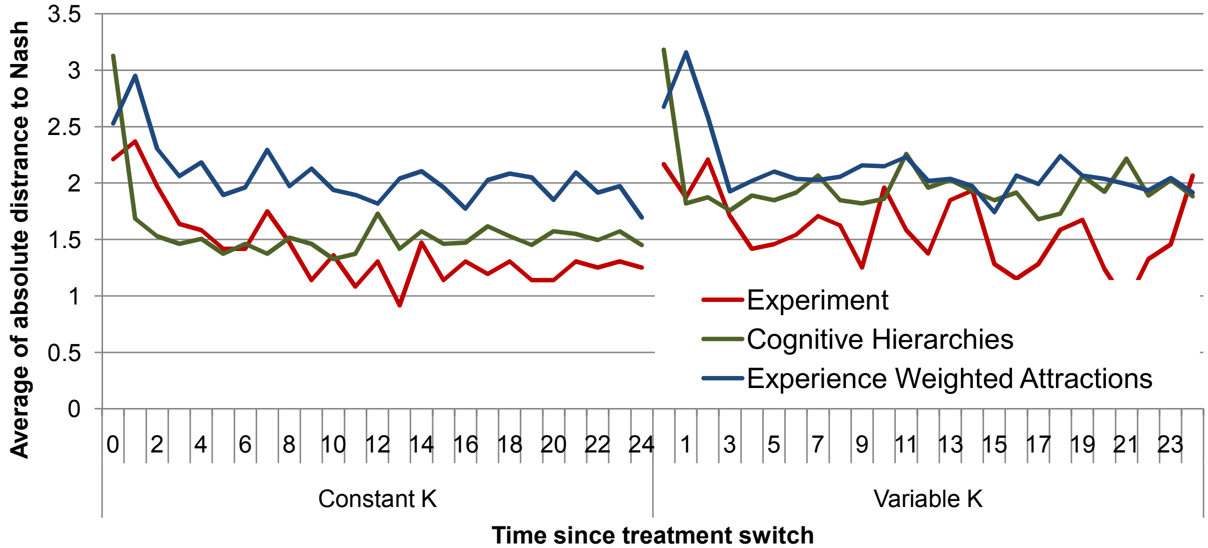


Figure 3: Predictions and observations of distance to the average RNNE as a function of time elapsed since the treatment change.

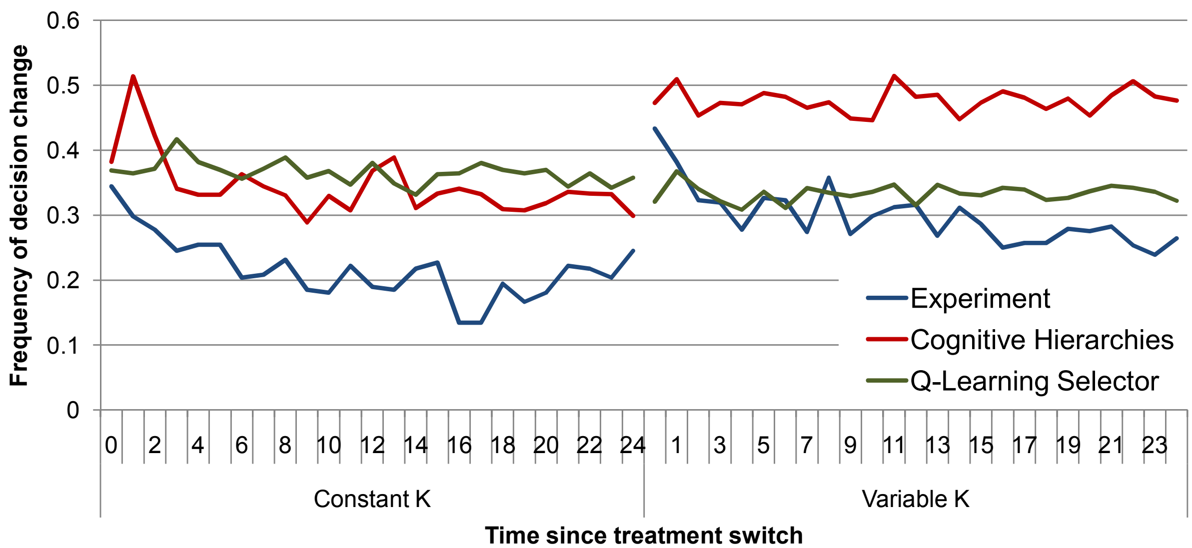


Figure 4: Predictions and observed frequencies of changes of individual decisions.

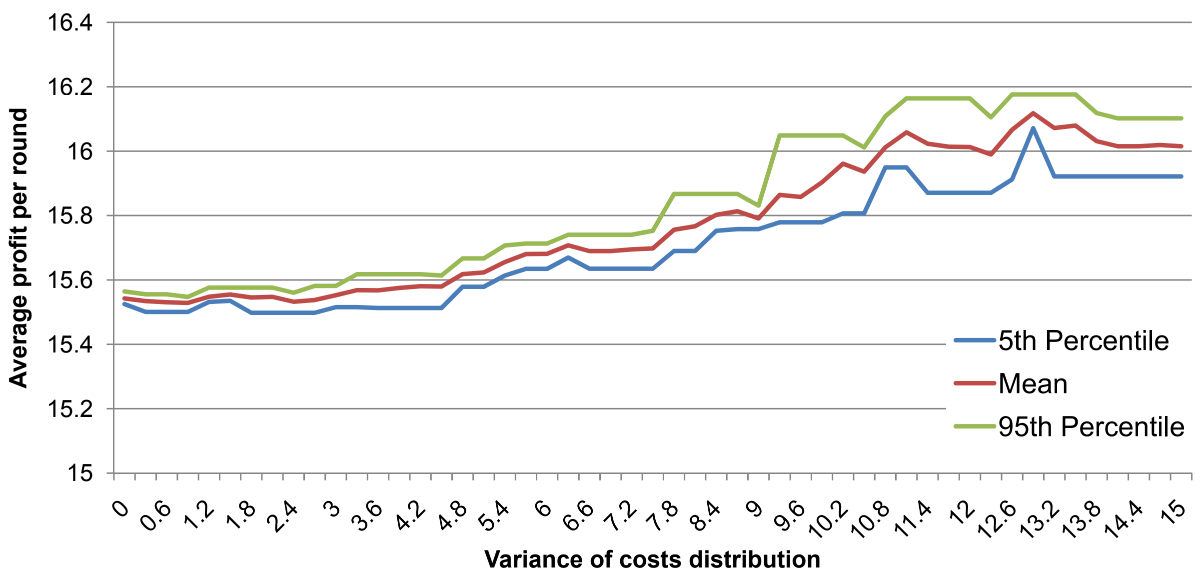


Figure 5: Prediction of influence degree of entry cost assymetry on the performance of the market. Individual costs were drawn from a lognormal distribution with constant mean 1 and different variances. In order to preserve the expected Nash structure, we subtracted 1 uniformly from all fees f for all treatments.