

Estimating choice parameters under coarse feedback

Juni Singh and Philippe Jehiel*

March 2017

Abstract

This paper tries to focus on the choices made by people under uncertainty with feedback via an experiment. The experiment consists of a one player decision task across two different situations. It tries to mimic a real life setting where agents need to make decisions, often not knowing the payoff from a particular option completely. There are some options which have been experimented enough by people and are fairly unambiguous. Often it is too demanding to consider each option individually, so we tend to simplify our choice set by grouping similar options together. The paper uses this heuristic to present information to players across situations. We find that people do not stay away from the ambiguous option but instead learn about it via the repetitive feedback. A reinforcement learning rule for updating and a logit rule for choice helps us establish that in the long run people do not discount ambiguity. Further the keenness of choosing the ambiguous urn depends on the payoff of the outside option.

Key words: Ambiguity, Bounded Rationality, Experiment, Learning, Updating

JEL Classification: D81, D83, C12, C91

*We would like to thank PSE Research grants and LABEX OSE for having provided us with the funds to conduct the laboratory experiments. We would also like to thank Prof Nicolas Jacquemet and Philipp Ketz for helping with statistical estimation. For the execution of the experiment and logistics, we would like to thank Maxim Frolov. I have benefited from the initial comments of Prof Itzhak Giboa, Julien Combe and Elias Bouacida.

1 Introduction

Often in life, we have too many choices to make. In order to simplify our decisions, we use different heuristics. Grouping similar options together is one of the probable ways of simplifying decision making. It could be a financial decision (investment, stocks), personal decision (jobs, friendship) and so on. This is often because the information we have is not clear but coarse. We get feedback about our options in an aggregate form which further makes it convenient for us to group options. Given the coarse nature of the feedback, the option becomes ambiguous. In this paper, we try to study how people learn and make choices given a coarse feedback. It is going to be an experimental approach where we try to enrich the literature on reinforcement learning by introducing ambiguity and aspects of bounded rationality.

As humans, we are likely to repeat actions with a higher positive payoff. This tendency of people is the simple implication of reinforcement learning literature. However often it is the case that we have little or no information regarding the outcome of our actions/options. Even if feedback exists, it is in a very coarse form. Using this as a motivation we have designed an individual decision problem that aims to understand how people make choices under such circumstance. The coarse nature of the feedback allows us to introduce an ambiguity parameter in our experiment which is borrowed from the ambiguity literature. It would be one of the first attempts to reconcile the reinforcement learning and ambiguity literature.

The current literature of individual decision making under ambiguity is limited to mostly one period plays. Individuals are either ambiguity-loving or ambiguity averse based on their decision. The evolution of this ambiguity parameter given feedback has not been explored. Jehiel and Samet (2007) have proposed a model of learning through valuation approach when the feedback is coarse. The basic idea is that people group similar moves together and assign a common value as they get feedback. The coarseness of feedback is introduced in our experiment via aggregated feedback across states for similar options. (Similarity is induced by colors in our experiment and can be interpreted as categorization). The learning model that we will be employing is in line with the reinforcement learning model proposed by Erev and Roth (1998).

The aggregation of feedback in order to evaluate options is in line with bounded rationality

models that give importance to different heuristics like labeling to simplify everyday decisions. (eg: think of bio label on a variety of food products) .This is also in line with the prospect theory literature by Kahneman and Tvresky (1974).This way of providing aggregate feedback establishes an extra layer of ambiguity in our experiment which cannot be reduced over time. Some of the options have information that is not coarse while some others have coarse feedback. This is motivated by the fact that some actions might be better documented or more familiar than the others in real life. We plan to do additional treatments to understand the current data and to make robust conclusions.

The central idea is to understand the decisions people make by introducing new aspects to the learning literature. By allowing ambiguity via a coarse feedback mechanism, the experiment will allow us to draw conclusions on the reinforcement learning and the variation of the ambiguity through time. The next section provides a comprehensive review of existing literature which is most closely related to this paper. This is followed by the theory and a couple of examples that would fit in this framework.We move on to explain the experimental setting followed by the initial results. Section 6 focuses on the estimation strategy for the analysis which is then followed by the results and conclusions.

2 Literature Review

This paper draws inspiration from various strands of the literature. The experiment uses an Ellsberg(1960)kind of framework to study the evolution of ambiguity via learning across periods. Players play a repeated Ellsberg decision problem across two states and 70 periods. The paper then tries to use reinforcement learning under bounded rationality to explain choices with ambiguous options. The literature on ambiguity is relevant to understanding how ambiguity has been dealt with in the past. The kind of ambiguity we are interested in is due to the absence of information about certain options. There are a bunch of papers which study how people choose when faced with such ambiguous option.

Choquet expected utility introduced by Schmeidler (1989) uses the concept of consideration of the lowest outcome by an agent and addition of successive increments weighted by his personal

estimation capacity to represent a belief. Gilboa and Schmeidler(1989) introduced the concept of maxmin expected utility where players have multiple priors and minimal expected utility is used for updating. Klibanoff, Marinacci, and Mukherjee(2005) propose another decision model under ambiguity that allows a set of priors without the maxmin criteria drawing a link between risk models and ambiguity.

In our paper, we allow players to update their ambiguous belief with repeated feedback across periods. Players have a single prior and they learn about their options by simple reinforcement learning, that is to choose the option that gave them a higher payoff more frequently. In the ambiguity literature players often use either Bayesian updating or maximum likelihood updating. The maximum likelihood updating is studied by Epstein and Schneider(2007).They claim that Ellsberg-type behavior exists in the short run but disappears with learning in the long run. We arrive at a similar conclusion using the reinforcement learning framework. The seminal paper on reinforcement learning is the one by Erev and Roth (1998).

We use this way of updating since we believe it is the most intuitive, given the limited computing capacity of agents. In line with using the reinforcement learning model,the Experienced Weight Attraction (EWA) learning concept introduced by Camerer and Ho (1999) introduces a new learning concept by combining reinforcement learning with response learning. Learning by valuation (Jehiel 2005) also studies reinforcement versus response effect. There is also a strand of psychology literature that studies how people assimilate a feedback (Gluck et al. 2002) by analyzing the heuristics behind weather prediction tasks in experiments.

Drawing inspiration from Jehiel and Samet(2007), we imagine that players simplify the decision making and updating process by assigning values to each of the options. The authors introduce similarity classes(a concept of grouping similar options together) which we induce in our experiment via similar colors and combined feedback. Valuation equilibrium and how people change the assigned values across moves is also included in our experiment. The way feedback is provided to the agents automatically inclines us in using this equilibrium concept. Given the learning rule,we assume that people chose between options by using a simple logit rule (probabilistic choices). This has widely been used in the literature of individual decision making which claims that given bounded rationality; people do not necessarily play the best response.

For the statistical estimation of the experimental data, the paper uses structural parametric estimation. Similar to the concept of Quantal Response Functions (McKelvey 1995) the idea was to determine a λ parameter that would capture the noise in the decision making of the players in addition to other parameters. Players try to respond optimally but with noise in the decision captured by the logit scale parameter. The authors examine reinforcement versus belief learning based on a choice probability model using logit.

Looking at experiments related to updating and learning under ambiguity, Vanet and Zeelenberg (2003) consider discounting of ambiguous information by means of three experimental designs. The authors introduce components of ambiguity related to price and other factors to study the difference in decision making with varied information. Charness and Levine (2003) study the alignment of reinforcement learning and Bayesian updating by means of a simple experiment. Our experimental setup of uncertain payoffs has been inspired by their design. Experiments studying the evolution of ambiguity and learning have been fairly limited in the literature. In addition to bringing the various strands of literature together, the paper also contributes to the existing experimental literature.

3 Theory

Consider a decision making problem under ambiguity which is repeated for n periods. The decision choice for each player is then a finitely repeated game. The decision problem is modeled as a dynamic situation where ambiguity is reducible over time. There are two states of the world $s1, s2$ and each player is randomly assigned to either of the states. Each player knows his state but does not know the state of the other players. In each state, players need to choose between two options, the payoff of which is unknown to the players. Assume the options are B,R in $s1$ and R,G in $s2$. The action R in the two states have different payoff but share a common name.

Players have a prior belief about the options which they update based on existing belief and current values of the feedback. After every round, individuals are provided with the feedback f_B, f_{R1}, f_{R2}, f_G for control group and for treatment f_B, f_R, f_G . The fact that the feedback for the treatment group is combined across states for Red is the idea of grouping similar options together as motivated from valuation reasoning. Here the common feedback forces them to consider both

the outcomes Red in state1 and Red in state2 as similar. The similarity group is given exogenously in this case. It would be ideal to assume that groupings are endogenously formed by agents but that makes the model intractable.

Each option has a payoff associated with it say, x_B, x_{R1}, x_{R2}, x_G . Players have greater information of x_B and x_G but no information about x_{R1} and x_{R2} . An ambiguity averse agent would probably stay away from the Red urns and pick the blue and the green option more frequently. Going by the valuation equilibrium notion (Jehiel and Samet 2007), people would tend to decide by assigning valuations to each option and in particular, a common valuation for both the Reds. Valuation equilibrium is the idea of assigning a single value to a group of moves. Given that the feedback on the Reds is combined, agents might account for the ambiguity. However, with more information about the urns, we would argue that the behavior converges towards certain options.

The idea is to then study what valuation equilibrium would have to say about the frequency of choices. With respect to the learning rules, to keep it as simple as possible we use reinforcement learning. If the player knew for certain that $x_B > x_R$ then Blue would be picked. Given the uncertainty, choice is guided by a probability function (logit) in our case. The value assigned to the urns or belief about the x changes with the feedback. In the end, choice should stabilize to the moves/options with highest valuation. In explaining the stable choice for each situation in the experiment that follows, this theoretical concept has been utilized.

In order to clarify the concept, consider an example in real life where we often need to decide between options. Take for instance, a farmer who is trying to decide whether to adopt a new technology. He has more information about the efficiency of the methods he is using now rather than the new technology (existing tech would represent options B and G). The information about the new technology often is not very clear and is generated from farmers who both used it correctly and did not (Think of the two types of farmers as the two different state across which information is assimilated.) Again think of someone deciding whether they need a complementary health insurance. The information is often not clear and is combined across different kinds of people (say old and young) for instance. The decision then depends on how promising the outside option is. This could also be applied to explain homophily and discrimination where people often assign higher values in connecting with similar people.

4 Experiment

The experimental design consists of a simple one player decision game. The central idea was that people tend to bundle their moves according to certain similar features. Here, the same color of the urn is to induce such behavior in the treatment group as represented in Figure 1. There are two states of the world and people make a choice (which urn to pick) in each state. The state is known to the agent but the composition of the urns is unknown. In State 1 players choose between the urns Blue and Red (Red 1) whereas in State 2 players choose between Green and Red (Red 2).

IV.1. Design

Each urn contains 10 balls which are either black (B) or White (W). W gives zero payoff whereas B gives a fixed amount. After each round, feedback regarding the choices of previous round is given. To be more precise, information about absolute numbers of B and W from each urn is presented. The difference between Control and Treatment group stems from the way feedback is reported. As seen in Figure 2, for the control group information about the urns are given separately. For the treatment group as seen in Figure 3 the information on the Red urn is given collectively across states. This implies in addition to no initial information on the Red urn, another level of ambiguity is introduced. In order to reduce the levels of ambiguity the subjects were provided with initial information regarding the Blue and the Green urn as seen in Table 1

Blue	30 Black	70 White
Green	68 Black	32 White

Table 1: Initial information given to the players. The absolute no. of black balls from the two urns after 100 random draws are reported

The experiment consisted of 70 rounds. Four sessions each of the control and treatment group were carried out in order to take into consideration the group effect. Given that the quality of feedback depended on the choices of the participants, four sessions were necessary. The experiments were conducted in the Laboratory at Maison de Sciences Economiques (MSE). The participants, 20 in each session were obtained from a common pool of participants at the Parisian Experimental Economics Laboratory (LEEP). The sample of population was fairly representative. It consisted of a mixture of mostly students and others.

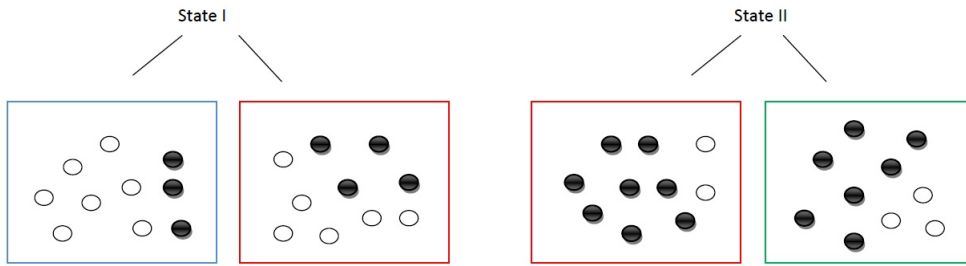


Figure 1 *The experimental setup*

IV.2 Procedure

Every player was presented with an instruction sheet (Appendix) describing the experiment and the rules. At the start of the experiment the instructions were read out in addition to a printed version. It was a simple decision task. There were two states the players may face referred to as states 1 and 2. In each state, they had to choose one of the two urns. Each urn was composed of ten balls either black or white in color. When an urn was chosen, one of the balls in the urn was drawn at random (by the computer) and it was immediately replaced by the same ball after the computer noted the color of the ball. It is important to note here that there was no private information that is the players did not know the color of the ball they obtained from the choice of the urn.

The players were clearly informed that while the compositions of the various urns remain the same throughout the experiment, the compositions of the Red urn in state 1 need not be the same as the composition of the Red urn in state 2. The sequence of choices from states 1 or 2 was decided randomly by the computer. The only task of the player was to choose one urn out of the two in each state. After one round of choices, the players were presented with feedbacks which can be seen in the screen shots. For determining the payoff, two of the rounds out of 70 were randomly chosen at the end of the experiment. If one of the balls in these two rounds was Black, players earned additional 5euros. If both of the balls in these two rounds were Black, players earned an additional 5euros.

A cette période, vous faites un choix dans l'état 2.

Balles tirées dans l'urne de chaque couleur lors de la période précédente :

Urne		 (état 1)	 (état 2)	
Nombre de balles de chaque couleur (● est une balle gagnante)	$\frac{1}{0}$  	$\frac{1}{0}$  	$\frac{0}{1}$  	$\frac{0}{1}$  

Quelle urne choisissez-vous ?



Figure 2 The feedback for control group

A cette période, vous faites un choix dans l'état 2.

Balles tirées dans l'urne de chaque couleur lors de la période précédente :

Urne			
Nombre de balles de chaque couleur (● est une balle gagnante)	$\frac{1}{0}$  	$\frac{1}{1}$  	$\frac{0}{1}$  

Quelle urne choisissez-vous ?



Figure 3 The feedback for treatment group

In addition we conducted extra treatments by varying the payoff of the Red urn. This was to ensure robustness in our results and learning model via variation in the feedback. An alternative explanation for the data could be aspiration based or reference point heuristics. Through our additional treatments, we wanted to be sure that we can establish that players pay attention to the feedback and learn. The additional treatment has been described in detail in the next section.

5 Initial Choice behavior

The figures below show how players played across time and across states. They primarily give an answer to our question if there was convergence in choice behaviors. It seems that choice stabilizes moreover around time $t=30$. After $t=30$, there is less ambiguity in the situation of choice and people tend to follow a stable strategy. It is interesting to see how people learn with the feedback and behave differently across different treatments.

In Figure 4, we see that the choice of players converges to Red in State 1 and Red in State 2. The average payoffs of each urn in the experiment was as follows, Blue=0.3, Red=1 in State 1 and Red=0.6 and Green=0.7 in State 2. According to Nash predictions, the player should choose Red in State 1 and Green in State 2. Given that the feedback for the treatment for Red was given jointly across states, the idea of valuation equilibrium would suggest that players converge to Red in both states. If players were completely ambiguity averse, then the choice should have stabilized to Blue in State 1 and Green in State 2 given the initial information. The behavior of players away from Nash equilibrium and ambiguity aversion is of interest in this case.

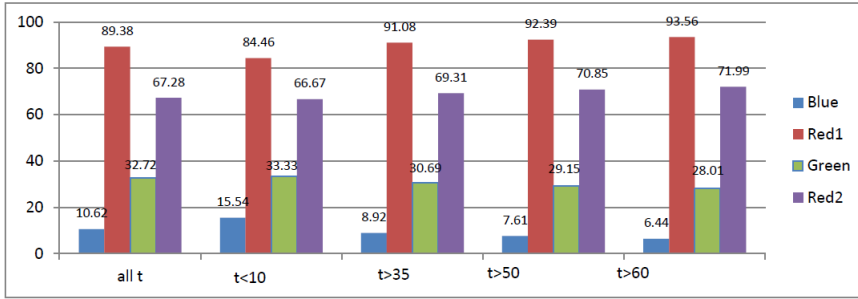


Figure 4 *Convergence to Red (State1) and Red (State2) in Treatment1*

In Figure 5, we see that the choice of players converges to Red in State 1 and Green in State2. The average payoffs of each urn in the experiment was as follows, Blue=0.3, Red=0.4 in State 1 and Red=0.8 and Green=0.7 in State 2. According to Nash predictions, the player should choose Red in State1 and Red in State2. Similar to treatment 1 above, given that the feedback for the treatment for Red was given jointly across states, the idea of valuation equilibrium would suggest that players converge to Red in State1 and Green in State2.

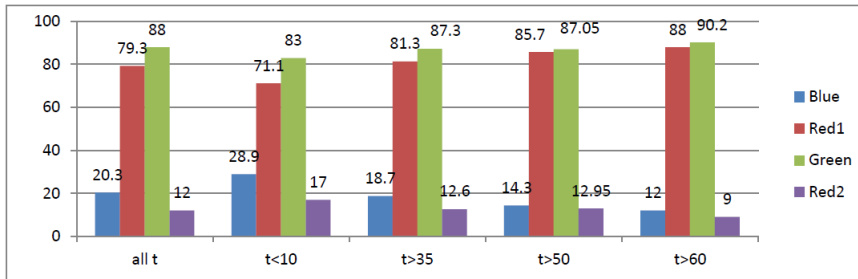


Figure 5 *Convergence to Red (State1) and Green (State2) in Treatment2*

Similarly in Figure 6, the choice of players converges to Red in State 1 and Red—Green in State2. The average payoffs of each urn in the experiment was as follows, Blue=0.3, Red=0.4 in State 1 and Red=0.8 and Green=0.7 in State 2. According to Nash predictions, the player should

choose Red in State1 and Red in State2. Since this was the control group, the feedback for the for Red urn was given separately across states. Since there is no grouping of similar options here, the idea of valuation equilibrium would not apply.

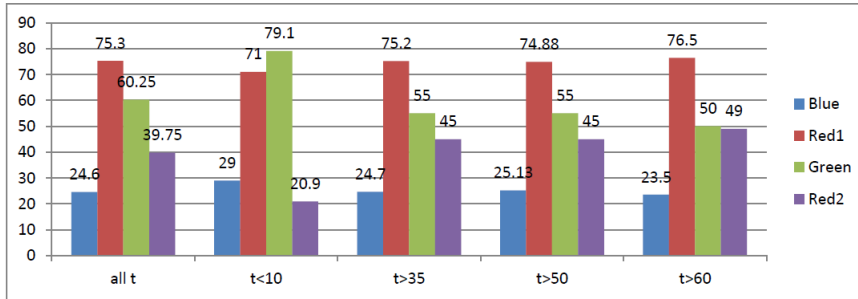


Figure 6 *Convergence to Red (State1) and Red—Green(State2)in Treatment1*

Given the descriptive statistics of the choice behavior, we propose a learning model to understand the evolution of ambiguity and choice behavior. This has been presented in the next section. The main idea is to use the preliminary findings of the initial choice description to fit parameters that would help generalize the choice behavior. This sort of exercise could be useful in applied works whenever an agent has to decide between different options with more information for some and less/no information for the others.

6 Estimation Strategy

Given the initial choice behavior, the data looked promising with respect to studying the evolution of ambiguity and the learning process. In order to formalize the learning story idea, we used maximum likelihood estimations and Monte Carlo simulations to arrive at the correct model specification. For the structural estimation of the data, there are two important components to be emphasized on. In line with the logit choice models, players follow probabilistic decision rule. With the provision of feedback every period, players are assumed to use a simple reinforcement learning update rule where players learn from their past beliefs and with new information.

1.Update rule:

$$BR_t = \rho \times BR_{t-1} + (1 - \rho) \times UR_t \tag{1}$$

where BR :Belief about the payoff from the urn, ρ : weightage on previous period belief, UR :Current period feedback

2. Choice Rule:

$$Prob(y = R) = \frac{exp^{\lambda(BR-\delta)}}{exp^{\lambda(BR-\delta)} + exp^{\lambda BB}}$$

where δ includes ambiguity in the Red urn and λ takes into consideration the randomness of play

This logit formulation is similar to Quantal Response Equilibrium (Palfray 1997). The idea is thus to estimate the fur parameters ρ , δ and λ in addition to BR_1 for the Red urn. For the other urns, players are assumed to have priors $BB_1 = 0.3$ and $BG_1 = 0.7$ respectively. We pool the data across treatments together in order to estimate the learning model. We have tried a couple of variations around the original model and also reinforce our results with likelihood ratio tests. The results have been laid out in the next section.

7 Results

With the structural estimation we have estimated two possible models to fit the data generating process. The first model propose that given the combined feedback for Red across states, players discount the information with a parameter δ . As mentioned in the estimation strategy above, the idea is to then see if people remain ambiguity averse or does the behavior towards ambiguity evolve over time. For all period t

$$\rho \text{ Red} = 0.89, \quad \rho \text{ Blue} = 0.96, \quad \delta = 0.007, \quad \lambda = 5.862, \quad BR_1 = 0.6$$

This shows that the ambiguity averse agents almost reduces to ambiguity neutral over a period of time. The repeated feedback after each play drives this result. Given this intuition, we checked the model with the hypothesis that $\delta = 0$. It turns out that we cannot reject the hypothesis after a likelihood ratio test. Thus the initial aversion towards ambiguity reduces and even disappears when agents receive enough information even across states. The weightage of beliefs for the urns

seems to be a little higher but that is explained by the fact that the new feedback reinforces previous period beliefs. This leads people to assign maximum weight to their own beliefs after certain period. The λ accounts for the noise/randomness in people's behavior. If people played randomly λ would equal 0 and if people played Nash it would be infinity.

Another way to assimilate the behavior of people could be that people behave differently depending on what the outer option looks like. If the outer option seems less rewarding, based on an aspiration argument players might be attracted towards the ambiguous option (Red in our experiment) This reasoning in fact holds true,

$$\rho_{\text{Red}} = 0.79, \quad \rho_{\text{Blue}} = 0.95, \quad \delta_1 = -0.07, \quad \delta_2 = 0.04, \quad \lambda = 4.862 \quad BR_1 = 0.6$$

Therefore, given the poor payoff of Blue in State1, people tend to choose more of the Red whereas given the high payoff of Green in State2, people choose red less frequently. This could be seen through the behavior of players in period $t=1$ as well. In order to establish the robustness of the results, we conducted addition hypothesis with alternative models. We could successfully reject the hypothesis that players chose randomly with probability 0.5 for each urn.

For the control group, we observe similar patterns that is in State1 people seem to be ambiguity loving and in State2 players tend to be ambiguity averse. Again the level of ambiguity averse/love declines with the availability with new information. Similar to what is established in the ambiguity literature, with our learning rule, Ellsberg-type behavior tends to vanish in the long run. With sufficient information and repetition, people tend to solve the ambiguity problem. In addition as argued by valuation equilibrium, people reason by assigning a common valuation to a group of moves.

8 Conclusion

The one player decision task was the simplest framework to study the sort of questions we were interested in. Mimicking a real life decision problem, a player was aware of the different states/situations he was faced with. (For instance happy mood/sad mood or good investment environment/ bad investment environment). Moving away from how ambiguity is generally portrayed in the literature,

ambiguity in our decision problem was due to the uncertainty of payoff from options (Consider not knowing in advance if you would like your new job, or not knowing if it is profitable to buy a particular stock). There are some options in life which we are more familiar with and were depicted in our experiment by providing initial information to the players. This can be thought of options that are frequently chosen by people and thus enough information has been established about (for instance think of use of a technology that has been widely used before).

The idea was to then study how the ambiguity present in the unknown option changed over time. As seen through the results, in the long run ambiguity decreases almost making people ambiguity neutral. Further we exploited the bounded rationality of the players by assuming that they group similar moves together and value them jointly. In the experiment, combining information across states led to exogenously formed similarity classes. Players did not seem to stay away from the grouped moves but instead learnt from the combined feedback, as predicted by the valuation equilibrium literature.

This was seen through the initial behavior of players where choices were stabilized with different options across treatments. People did use the feedback that was provided. It seemed through the data that most of the learning occurred in the first 30 periods, after which the behavior fairly stabilized. The simple reinforcement learning rule seemed to be much more intuitive than computationally heavy updating rules.

Hence faced with uncertain options, people learned enough with repeated feedback. Further it seemed to matter what the alternative option was. If the less ambiguous option offered not very satisfactory payoff, people tended to experiment more. This seems intuitive with real life situations (Consider people leaving their own country for better prospects). Through this paper we are successfully able to connect the dots between learning with feedback under the presence of ambiguity and a particular heuristic used by people.

References

- [1] Weinsten A. N. The ambiguity aversion literature a critical assessment. *Economics and Philosophy*, pages 249–284, 2009.

- [2] Colin Camerer. *Behavioral Game Theory*. Princeton University Press, 2003.
- [3] Gary Charness and Dan Levin. When optimal choices feel wrong: a laboratory study of bayesian updating, complexity, and affect. *The American Economic Review*, 41(4):1300–1309, 2005.
- [4] Colin Cramer. Recent developments in modeling preferences: Uncertainty and ambiguity. *Journal of Risk and Uncertainty*, pages 325–370, 1992.
- [5] Colin Cramer. Experience-Weighted Attraction Learning in Normal form games. *Econometrica*, 67, 1999.
- [6] Daniel Ellsberg. Risk, ambiguity and the savage axioms. *The Quarterly Journal of Economics*, pages 643–661, 1961.
- [7] Larry Epstein and Martin Schneider. Learning under ambiguity. *Review of Economic Studies*, 74:1275–1303, 2007.
- [8] Ido Erev and Alvin Roth. Predicting how people play games : Reinforcement learning in experimental games with unique mixed strategy. *The American Economic Review*, 88(4):848–881, 1998.
- [9] Zeelenberg E.V. The discounting of ambiguous information in economic decision making. *Journal of Decision making*, 2003.
- [10] Mark A. Gluck, Daphna Shohamy, and Catherine Myers. How do people solve the weather prediction task: Individual strategies for probabilistic category learning. *Learning and Memory*, 9(6):408–418, 2002.
- [11] Philippe Jehiel and Samet Dove. Learning to play extensive games by valuation. *Journal of Economics Theory*, 121(557):129–148, 2005.
- [12] Philippe Jehiel and Samet Dove. Valuation equilibrium. *Theoretical Economics*, pages 163–185, 2007.
- [13] Train Kenneth. *Discrete Choice Methods and Simulations*. Cambridge University Press, 2003.

- [14] Mark J. Machina and Marciano Siniscalchi. *The Handbook of Economics of Risk and Uncertainty*, chapter 13. Elsevier North-Holland, 2013.
- [15] Richard D McKelvey and Thomas R. Palfrey. Quantal response equilibria for normal form games. *Games and Economic Behavior*, 10:6–38, 1995.
- [16] Joel Sobel. Economist model of Learning. *Journal of Economic Theory*, pages 241–261, 2000.
- [17] Amos Tversky. Thinking through uncertainty: Non consequential reasoning and choice. *Cognitive Psychology*, 1992.
- [18] Amos Tversky and Daneil Kahneman. Judgment under uncertainty: Heuristics and biases. *Science, New Series*, 185(4157):1124–1131, 1974.
- [19] Amos Tversky and Daneil Kahneman. Prospect theory, an analysis of decision under risk. *Econometrica*, 47(2):263–291, 1979.

9 Appendix

Instruction sheet for the players (In the lab the instructions were handed out in French):

Control Group:

Welcome to the experiment and I thank you for your participation. Please listen to these instructions carefully. If you have any questions kindly raise your hand and it shall be addressed. You receive 5 euros for participating and then your payoff depends on your performance in the experiment.

The Experiment:

The experiment consists of 70 rounds. It is a simple decision task. There are two situations you may face referred to as states 1 and 2. In each state, you have to choose one of two urns. Each urn is composed of ten balls either black or white in color. When you choose an urn, one of the balls in the urn is drawn at random (by the computer) and it is immediately replaced after the computer has noted the color of the ball. If the ball drawn is Black, you can receive extra payment (see below for details) whereas if the ball drawn is White you receive no payment.

The two urns available in state 1 are Blue and Red, respectively. The two urns available in state 2 are Green and Red, respectively. While the compositions of the various urns remain the same throughout the experiment, note that the compositions of the Red urn in state 1 need not be the same as the composition of the Red urn in state 2. These are two different urns.

As the experiment goes, on your computer screen, you will be informed whether you have to make a choice of urns in state 1 (Blue or Red) or in state 2 (Green or Red). The sequence of choices from states 1 or 2 is decided randomly by the computer. Your task is to choose one urn out of the two in each state.

Note: We drew 100 times a ball (replacing the ball in the urn after each draw) out of the Blue and Green urn. We obtained the following composition

Blue	30 Black	70 White
Green	68 Black	32 White

At the beginning of the experiment:

- Your terminal is randomly assigned a State of the world. If in State 1, you choose between a Red and Blue urn. If in State 2, you choose between a Red and Green urn.
- After you choose the color of the urn that you want to pick, you click on the screen. A ball (the color of which could be either Black or White) will be drawn from that urn by the computer. You will not know the color of the ball drawn. This implies you will not have the information of your choice.
- Once all participants have made their choice, we provide you with some feedback. The total no. of black and white balls drawn in previous rounds by all subjects according to the color of the urn (Blue, Red1, Red2, Green).
- Following the feedback, your terminal is randomly assigned a state of the world again. The state may vary from the previous round or remain the same.
- We then repeat the same experiment again until the completion of the 70 rounds.

For determining your payoff, two of the rounds will be randomly chosen at the end of the experiment. If one of your balls in these two rounds is Black, you will get an extra 5 euros. If both

of your balls in these two rounds are Black, you will have an extra 10 euros. Otherwise (if both balls are White), you will have no extra return. So if you have no questions let us begin!

Treatment Group:

Welcome to the experiment and I thank you for your participation. Please listen to these instructions carefully. If you have any questions kindly raise your hand and it shall be addressed. You receive 5 euros for participating and then your payoff depends on your performance in the experiment.

The Experiment:

The experiment consists of 70 rounds. It is a simple decision task. There are two states of the world each containing two urns. One of the urns is Red and the other is either Blue or Green depending on the state (Note: The composition of the Red urn across states is not the same). On your computer screen, you will play different states of the world randomly. Your task is to choose one urn out of the two in each state. Each urn consists of 10 balls either black or white in color. The Black gives you a payoff of 5 euro whereas the White gives no payoff .

Note: We drew 100 balls randomly out of the Blue and Green urn which gave us the composition

Blue	30 Black	70 White
Green	68 Black	32 White

At the beginning of the experiment:

- Your terminal is randomly assigned a State of the world. If in State 1, you choose between a Red and Blue urn. If in State 2, you choose between a Red and Green urn.
- After you choose the color of the urn that you want to pick, you click on the screen. A ball (color of which could be either Black or White) will be picked up from that urn. You will not know the color of the ball drawn. This implies you will not have the information of your choice
- Once every participant has made their choice, we provide you with the feedback. The no. of black and white balls drawn from each colored urn(Blue, Red, Green) across states based on only the previous round draw is reported.

- Following the feedback, your terminal is randomly assigned a state of the world again. The state may vary from the previous round or remain the same. Note that the composition of the urn is however fixed throughout the experiment.
- We then repeat the same experiment again till we complete 70 rounds.

For determining your payoff, two of the rounds will be randomly chosen at the end of the experiment. If you have picked up B in that particular round, you end up with 5 euros more for each B otherwise no returns. So if you have no questions let us begin!