Do We Learn From Mistakes of Others? A Test of Observational Learning in the Bandit Problem

Igor Asanov*

University of Kassel, D-34109, Kassel, Germany

December 9, 2016

Abstract

I experimentally investigate observational (social) learning in the simple two-armed bandit framework where the models based on Bayesian reasoning and non-Bayesian reasoning (count heuristics) have different predictions. The results contradict the predictions of the Bayesian rationality e.g. Bayesian Nash Equilibrium, Naïve herding model (BRTNI): Subjects follow the choices that contain no information about the state of the world, follow the coinciding choices of others (though it is empirically suboptimal), sustain losses making every first choice and cascade early then Bayesian-based models predict, but not in a random way. In addition, the Quantal Response Equilibrium is tested and the robustness of the theory questioned.

JEL Classification: C92, D82, D83

Keywords: Observational learning; information cascade; experiment

---

*Adress: Mönchebergstrae 17, Department of Economics, University of Kassel, Kassel, Germany. E-mail: igor.asanov@uni-kassel.de. I am grateful to Oliver Kirchkamp, Uwe Cantner, Arno Riedl, Werner Güth for their useful suggestions and support. I would like to thank Thomas De Haan, Paolo Crosetto, Alexia Gaudel, Kyle Hyndman and the participants of the 29th European Economic Association (EEA) Congress, the M-BEES conference 2013, the METEOR seminar in Maastricht, and Max Planck Institute Brown Bag seminar in Jena for their insightful comments and great discussion. I am also grateful to German Sciene Foundation (DFG) for financial support.

1 Introduction

The fact that perfectly rational agents act like they are ignoring their private information - falling into the information cascade (Banerjee, 1992; Bikhchandani et al., 1992) - raises an important issue for various economic activities where information accumulation plays a key role: technology diffusion, practice adoption, financial markets. The problem appears to be more pronounced with experimental evidence corroborating the theoretical predictions (Anderson and Holt, 1997; Hung and Plott, 2001). This experimental evidence, however, initiates a discussion on whether the observed behavior can be attributed to Bayesian reasoning or if it is an artifact of the information cascade environment (Huck and Oechssler, 2000; Eyster and Rabin, 2010).

The experiments that increase the complexity of information structure in the game find that behavior is inconsistent with Bayesian Nash Equilibrium and support the count heuristics explanation of behavior (Huck and Oechssler, 2000; Noth and Weber, 2003; Ziegelmeyer et al., 2009; March et al., 2012). Nevertheless, Bayesian-based models that allow for errors in decision making and over-weighting of public or private information – Quantal Response Equilibrium with base rate neglect or Naïve herding model – also fit the experimental data fairly well (Kübler and Weizsäcker, 2004; Goeree et al., 2007; Eyster et al., 2015). To disentangle these possible explanations I experimentally investigate observational learning in the simple two-armed bandit framework where the models based on Bayesian reasoning and count heuristics have sharply different predictions.

In the experiment, subjects face a choice between two alternatives. Each of the alternatives can bring the reward with certain probability, which depends on the equally probable state of the world. Subjects act in a sequence. When it is the turn of a subject, (s)he has to choose one of the alternatives two times in a row. Each time the subject makes a choice (s)he observes the outcomes of the choice. In addition, each subject observes all choices of their predecessors.

In this environment each of the two observed choices provides information only about one outcome since each subject makes the first choice before (s)he observes the outcome. Nevertheless, count heuristics suggest considering each of the choices as informative and follow the coinciding choice of predecssor(s) more often. As a consequence, simple counting prescribes earlier cascading behavior than rational theory does and lower expected payoff for every first (odd) choice.

The experiment based on the bandit cascade environment has few advantages. First, it is simple: Subjects have to make two binary choices in an environment where two states of the world are equally probable and the outcomes are equally informative. Therefore, myopic optimization is sufficient

(Bradt et al., 1956) and subjects in principle can mimic rational strategy (as in information cascade setting) by using counting strategies. Second, the bandit cascade model shares most of the features of the classical information cascade setting in pure Bayesian Nash Equilibrium. Nevertheless, predictions of other theories differ in this model. Third, the first choice of the first player in the game brings no information about the state of the world that allows me to directly confront predictions of count heuristics with the models based on Bayesian reasoning, e.g. Naïve herding model (BRTNI), Quantal Response Equilibrium.

Subjects, indeed, report that the experiment is very simple, however, that is not reflected in their payoffs: They sustain significant losses making first (odd) choices. Subjects act like they ignore their outcomes (cascade) earlier than the Bayesian Nash Equilibrium and Naïve herding prescribe, but not in a random way. More importantly, the subjects are significantly more likely to follow coinciding choice of others though it contradicts not only perfect rationality but is also empirically suboptimal. In addition, the choices of others that contain no information about the state of the world influence the subject's choices.

These facts are hard to reconcile with Bayesian reasoning. They conflict with the Quantal Response Equilibrium and Naïve herding model, explaining their poor predictive power. These results question the assumptions of the rational theory demanding for the development of non-Bayesian based theories in the domain of the observational learning. They also shed light on the observational learning in case of experimentation with alternatives of unknown merit e.g. technology, practice adoption or venture capital investment.

The rest of the paper is organized as follows. In Section 2, I introduce the bandit cascade model and discuss the theoretical predictions. Section 3 contains the experimental design and procedures. Section 4 describes the results and Section 5 concludes.

2 The Bandit Cascade Model

The infinite set of agents act in a sequence. Each agent i faces the choice between the two alternatives x, y. Each of the alternatives can bring the reward r. The probability of receiving the reward depends on the state of the world which has equal probability to be either X or Y.

In the state X, the alternative x brings the reward with probability $q \in (1/2, 1)$ and no reward 0 with $1 - q$, whereas the alternative y brings reward with probability $1 - q$ and no reward with q. In the state of the world Y, alternative y brings the reward with probability q, while alternative x brings the reward only with $1-q$. The state of the world is common for all agents, but the agents do not know the actual state of the world. They only have the prior public belief that $\theta_i^X = Pr(X) = Pr(Y) = 1/2$.

Agents make two choices in a row $c_{i,n}$ indexed as $n = \{1,2\}$. They privately observe the outcomes of their choices $O_{i,n} = \{r_{i,n}^x, r_{i,n}^y, c_{i,n}^x, c_{i,n}^y\}$. However, their choices are public – each agent observes a public history of all choices of their predecessors $H_{i-1} = (c_{1,1}, \ldots, c_{k,n}, \ldots, c_{i-1,2})$. The agents do not observe the outcomes of others' choices and the actions of others produce only information externalities.

2.1 Bayesian Nash Equilibrium

Bayesian Nash Equilibrium assumes that agents update their beliefs about the state of the world based on the outcomes of their choices as well as on the public history of choices of others. Suppose the overall number of agent i choices of the alternatives x and y is $C_{i,n}^x = \sum_1^n c_{i,n}^x$ and $C_{i,n}^y = \sum_1^n c_{i,n}^y$, respectively. The overall number of rewards received from these choices is $R_{i,n}^x = \sum_1^n r_{i,n}^x$ and $R_{i,n}^y = \sum_1^n r_{i,n}^y$. Without loss of generality in the state of the world X the probability of observing a history of outcomes $O_{i,n}$ is

$$Pr(O_{i,n}|X) = q^{R_{i,n}^x}(1-q)^{C_{i,n}^x - R_{i,n}^x}(1-q)^{R_{i,n}^y} q^{C_{i,n}^y - R_{i,n}^y} \quad (1)$$

Using Bayes' rule the agent i forms a private belief that the state of the world is X:

$$\pi(X|O_{i,n}, H_{i-1}) = \frac{Pr(O_{i,n}|X)\theta_i^X}{Pr(O_{i,n}|X)\theta_i^X + Pr(O_{i,n}|Y)\theta_i^Y} \quad (2)$$

Given the private beliefs, the agent's expected utility from taking the action x is:

$$\mu_{i,n}^x = \mu(x|O_{i,n}, H_{i-1}) = \pi(X|O_{i,n}, H_{i-1})q + \pi(Y|O_{i,n}, H_{i-1})(1-q) \quad (3)$$

Since any outcome of any choice of an agent is equally informative, the optimal strategy prescribes myopic optimization (Bradt et al., 1956). That is, the optimal decision in pure strategies of the agent i is simply to choose the alternative with the highest expected utility, $\max[\mu_{i,n}^x, \mu_{i,n}^y]$.

The expected utility from the first choice x of the first agent $i = 1$ is equal to 1/2, $\mu_{1,1}^x = 1 - \mu_{1,1}^y = 1/2$, and the agent is equally likely to choose either alternative x or y. The expected utility of the next choice of the agent $i = 1$ and the choice itself is fully determined by the private outcomes. The choice of the following agents, however, is affected by additional information, namely, the choice history of their predecessors.

A rational agent $i > 1$ takes into account the public history of preceding agents' choices H_{i-1} and forms the public belief about the state of the world with respect to it. In order to estimate this belief one has to calculate the transitional probabilities of the observed choice histories conditional on the possible states of the world. The transitional probability of choice $c_{i-1,n}$ in the world X is

$$\tau(c_{i-1,n}|X) = \mu(c_{i-1,n}|R^+, H_{i-2})Pr(R^+|X) + \mu(c_{i-1,n}|R^0, H_{i-2})Pr(R^0|X) \quad (4)$$

where R^+ and R^0 denote increase and no change in the number of the received rewards given the choice $c_{i-1,n-1}$, respectively.

Similarly, the transitional probability of the observed choice conditional on state of the world Y is

$$\tau(c_{i-1,n}|Y) = \mu(c_{i-1,n}|R^+, H_{i-2})Pr(R^+|Y) + \mu(c_{i-1,n}|R^0, H_{i-2})Pr(R^0|Y) \quad (5)$$

Now, if one combines these two transitional probabilities applying Bayes' rule, it is possible to obtain the public belief:

$$\theta_i^X = \theta(X|c_{i-1,n}) = \frac{\tau(c_{i-1,n}|X)\theta_{i-1}^X}{\tau(c_{i-1,n}|X)\theta_{i-1}^X + \tau(c_{i-1}|Y)\theta_{i-1}^Y} \quad (6)$$

The public belief dictates the agent's $i > 1$ first choice. When it is in favor of world X (Y), $\theta_i^X > 1/2$ ($\theta_i^Y > 1/2$), the first choice of the agent is x (y). The next choice depends on the interplay between the public and private beliefs. Thus, in order to describe equilibrium it is necessary to provide analysis of beliefs' properties.

Proposition 1. [1] For all $q \in (1/2, 1)$ the public belief θ_i has the following properties in pure strategy Bayesian Nash Equilibrium if the number of agents i goes to infinity:

(a) There is an agent i whose public belief exceeds q. $\exists i : \theta_i > q$.

(b) If $\theta_i > q$ agent choice is independent of the private outcomes. $\forall(\theta_i > q)$: $Pr(c_{i,n}|O_{i,n}) = Pr(c_{i,n})$.

The property of the public belief to exceed q is shared with the classical information cascade model (Banerjee, 1992; Bikhchandani et al., 1992).[2] As a consequence, the model exhibits the feature of cascading behavior – the agents tend to fall into the information cascade.

[1] The proof follows the logic similar to the proof in classical information cascade setting and is provided in appendix A.

[2] See Smith and Sorensen (2000) for an exhaustive study of the information cascade properties in two state cases and Park and Sabourian (2011) for the multi-state version.

Definition. The information cascade is the choice sequence when the agents act as if they ignore their own information and follow the behavior of their predecessor(s).

Similar to the classical cascade setting, in the bandit framework the agents eventually fall into the cascade. Sooner or later more then two agents make the same second choice and the public belief exceeds q. Thus, the agents' private outcome cannot overturn their beliefs and the learning stops.

Proposition 2. For all $q \in (1/2, 1)$ if the number of agents i goes to infinity the information cascade arises with probability 1. $\forall q \in (1/2, 1) : \lim_{i\to\infty} Pr(\theta_i > q) = 1$.

2.2 Naïve Herding

Naïve herding theory supposes that agents neglect the correlation between the actions of their predecessors (Eyster and Rabin, 2010). They naïvly believe that each action only reflects the agents' private information. That is, they simply think that other agents do not update their public beliefs and have public belief equal to the prior one. Since the prior beliefs assume that both of the states of the world are equally probable $\theta^\alpha = \theta^\beta$, the naïve agent i thinks that the private belief of the preceding agent is simply $\pi^x_{i-1} = 1 - \pi^y_{i-1} = Pr(O_{i-1,n}|\alpha)$.

As in pure strategies Bayesian Nash Equilibrium, the agents can fall into the information cascade starting from the third agent. However, when the naïve agents are in cascade they do not stop updating their beliefs. They continue to consider the actions of the agents that are in cascade as informative. Therefore, the public belief continues to grow.

2.3 Simple Count Heuristics

Unlike the previously discussed theories, the count heuristics reasoning departs completely from the Bayesian logic. Instead of sophisticated calculations, the agents simply count choices of others and combine this with the outcomes of their own actions. Thus, for agent i the value of choice x is

$$V^x_i = (\sum_1^{i-1} c^x - \sum_1^{i-1} c^y) + (R^x_{i,n} - R^y_{i,n} + ((C^y_{i,n} - R^y_{i,n}) - (C^x_{i,n} - R^x_{i,n}))) \qquad (7)$$

and the value of alternative y is $V^y_i = -V^x_i$. The agents choose the alternative with the highest (positive) value $\max[V^x, V^y]$. If $V^x = V^y$ they randomize uniformly between the alternatives.

The count heuristic model prescribes that the agent considers all of the predecessors' choices as informative, however, only the second choice contains useful information. As a consequence, the counting dictates suboptimal actions. For instance, agent i motivated by count heuristics is indifferent between alternatives if (s)he observes the switching choice of the predecessor $c_{i-1,1} \neq c_{i-1,2}$. The choices cancel each other out in counting and the agent fails to extract the useful information from these actions.

Agents using counting heuristics can also exhibit cascading behavior. Moreover, it takes even more severe forms. With the exception of the first one $i > 1$, all agents have a positive probability of acting as if they ignoring their outcomes, because they can easily double count the choices of their predecessors.

Claim 1. For all $q \in (1/2, 1)$ and all agents $i > 1$ there is a positive probability that the agent falls into the information cascade.

2.4 Theories and Predicted Expected Utility

The dissimilarity of the theories is reflected in the expected utility and can be illustrated with figure 1.

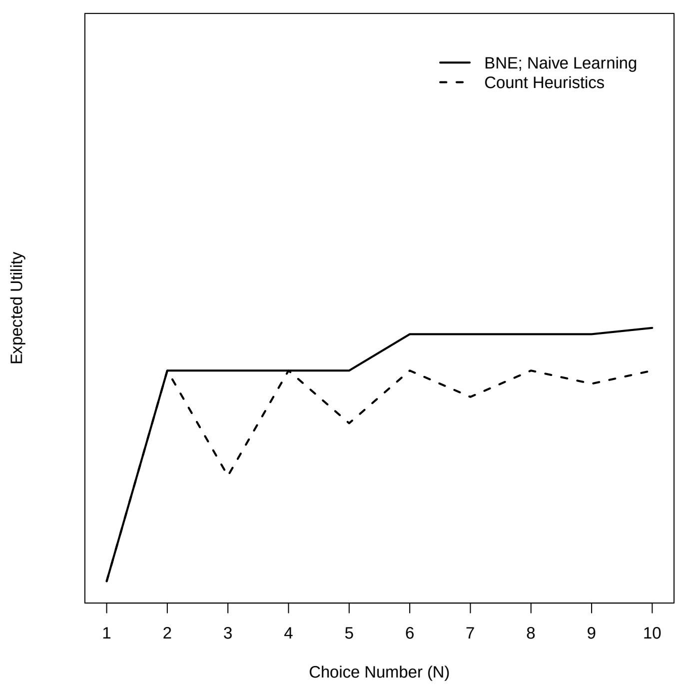


Figure 1: Expected Utility

Bayesian Nash Equilibrium as well as the Naïve herding model assume an increase in the expected utility for the second choice of the agent if the number of the predecessors is even. Under these circumstances, the agents either (1) neglect their own information and fall into the cascade or (2) follow the combination of private and public information.

If the agents are in cascade, they simply replicate their previous choice and the expected value does not change. If the agents do not fall into cascade, they perfectly infer the outcome of others and, taking into account their own outcome, on average correctly deduce the actual state of the world. Hence, they more often make a correct choice and the expected utility grows.

The utility shape in the count heuristic model is drastically different. First of all, agents that use counting rules sustain losses every first choice they make. They misinterpret switching, non-matching choices of others and randomize more then optimal.

Secondly, the agents' average benefit stays constant when they choose the second time. Consider that the difference between opposing choices of other agents is either equal to zero or to an even number. Therefore, when the agents make their second choice they either follow only their own outcome or replicate the last non-switching, matching choice of one of their predecessors. The agents have additional gain in neither case.

3 The Experiment

The experiment replicates the theoretical model. In the experiment, subjects face a game in which they have to make a choice between two alternatives x and y. The alternatives produce the reward ($r = 1$ game point) with a certain probability. The probability of the reward depends on the state of the world. The world can be in two equally probable states.

In one of the states the alternative x produces the reward with a probability q and the alternative y with probability $1 - q$. In the other state of the world, the alternative y produces the reward with probability q and the alternative x with $1 - q$.

The experiment involves 30 repetitions of the game. At the beginning of the first game, subjects are randomly matched in groups of five. In each game the actual state of the world is randomly selected for each group and subjects from the same group start to act in a sequence. The order in which they act is randomly determined in each new game.

When it is the turn of a subject, (s) he has to choose one of the alternatives two times in a row. Each time the subject makes a choice (s)he observes the outcomes of the choice. In addition, each subject observes all choices of their predecessors in the same game.

The study uses two treatments. The treatments are applied between subjects and differ in the probability of the reward, specifically q takes values 6/7 and 7/9. In each of the treatments the probabilities are expressed in frequencies and presented as balls of different color to facilitate subjects' comprehension.

In both treatments subjects are paid a flat fee of €3 for participation and each game point has a value of €3. To avoid the income effect, subjects are paid only for three of their choices. These choices are selected at random.[3]

The experiment was conducted in January and February 2013 in the laboratory of the Friedrich-Schiller-University in Jena.[4] [5] In total the experiment consists of 10 sessions with 10 participants in each. The participants were recruited on-line using ORSEE (Greiner, 2004).

On average each subject earned €8.52 for the participation in the one-hour session. The average age of the participants was 24.18 and the sample was almost perfectly balanced on gender (see table 1). The reported understanding of instruction was fairly high with a mean 3.83 out of a possible 5 and perceived task difficulty was low with a mean 3 out of 10.

Table 1: Participant characteristics

Statistic	N	Mean	St. Dev.	Min	Max
Age	100	24.180	3.899	18	41
Share of Females (♀)	100	0.520	0.502	0	1
Exp. Interesting	100	2.760	1.248	1	5
Exp. Length	100	2.420	0.741	1	5
Exp. Understandable	100	3.830	1.173	1	5
Task difficulty	100	3.000	1.735	1	9

4 Results

4.1 Average Payoffs

First, I consider the empirical average payoff (see figure 2). The visual inspection indicates that the shape of the empirically estimated payoff function

[3]See the theoretical analysis of the incentives in the experiments of Azrieli et al. (2012).

[4]The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007).

[5]The instructions of the experiment as well as a game screen are provided in appendix B.

contradicts both the predictions of Bayesian Nash Equilibrium and Naïve herding model. Instead, it has a shape predicted by count heuristics – every first choice of the player (odd choice number) is characterized by a lower average payoff than the second one (even) and the average payoffs for the second choice (even choice number) stay relatively constant.

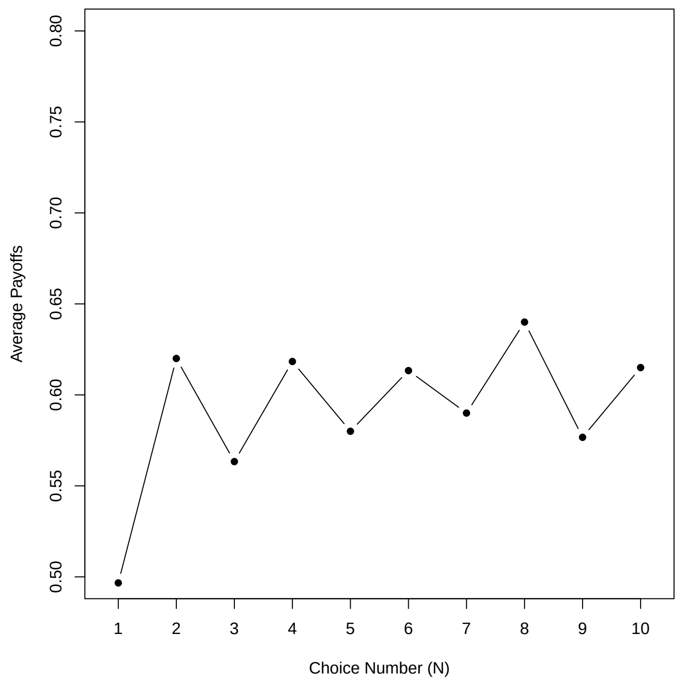


Figure 2: Average payoffs in both treatments.

Difference in payoffs between first and second choice. To assess the observation that the first choice is associated with losses in payoffs, I use the mixed-effect logistic regression model.[6] This estimates how the probability of receiving positive payoff $\mu = Pr(O = 1)$ depends on the other variables. Specifically, I run the main following regression:

$$\mu = Pr(O = 1) = \mathcal{L}(\beta_0 + \beta_{\odot} \odot_{i,r} + \beta_r r_i + \beta_T T_i + \beta_{♀} ♀_i + \beta_M M_i + \upsilon_s + \upsilon_i) \quad (8)$$

where $\mathcal{L}$ is the standard logistic function, $\odot$ is the first (odd) choice dummy, r – repetition number, T is a treatment dummy that is equal to one for treatment $q = 6/7$, ♀ – dummy equal to one for female subject and M subject reported math skills. υ_s and υ_i are random effects for session s and subject i, respectively.

[6]I provide all estimations within the R programming environment.

Table 2: Subject's Payoff – estimation of equation 8 and robustness check

	Dependent variable:			
	Payoff (μ)			
	(1)	(2)	(3)	(4)
First choice ($\odot$)	-0.188^{***}	-0.178^{***}	-0.188^{***}	-0.188^{***}
	(0.059)	(0.061)	(0.059)	(0.059)
Repetition (r)	-0.001	-0.001	-0.001	-0.001
	(0.003)	(0.003)	(0.003)	(0.003)
Treatment (T)	0.215^{***}			
	(0.078)			
Choice number (N)		0.010		
		(0.013)		
Female (♀)	-0.019	-0.025	-0.026	-0.005
	(0.078)	(0.080)	(0.079)	(0.080)
Math skills (M)	0.038^{**}	0.033^{**}	0.033^{**}	
	(0.016)	(0.016)	(0.016)	
Constant	0.173	0.249	0.318^{**}	0.520^{***}
	(0.138)	(0.164)	(0.133)	(0.092)
Observations	4,800	4,800	4,800	4,800
Log Likelihood	$-3{,}208.770$	$-3{,}210.920$	$-3{,}211.170$	$-3{,}213.290$
Akaike Inf. Crit.	6,431.530	6,437.840	6,436.340	6,438.580
Bayesian Inf. Crit.	6,476.870	6,489.650	6,481.680	6,477.440

Note: *p<0.1; **p<0.05; ***p<0.01

The results are reported in table 2. The null hypothesis that there is no difference between the odd and even choices is rejected, the p-value associated with the odd choice is below 0.01 [7] and the sign of the odd choice coefficient is negative, which is predicted only by the count heuristic theory.[8]

Given that the probability of receiving a positive payoff depends not only on the choice but also on the stochastic component of the reward realization, the regularity of the difference strikes. Moreover, the non-parametric exact paired Wilcoxon test applied across aggregated averages over the sessions also rejects the hypothesis that there is no difference between odd and even choices ($p < 0.01$), suggesting that the results are robust.[9]

Interestingly, the value of the $\beta_{\odot}$ is close to the value predicted by the count heuristics theory: Estimated $\beta_{\odot} = -0.188$ ($e^{\beta_{\odot}} = 0.828615$) with 95% confidence interval between -0.304 and -0.071, while the counting theory predicts $\beta_{\odot}^{Count} = -0.221$ ($e^{\beta_{\odot}^{Count}} = 0.801717$).[10] Correspondingly, estimating the regression 8 under assumption that $\beta_{\odot} = \beta_{\odot}^{Count}$ I cannot reject the hypothesis that the value predicted by the count heuristics theory and estimated value are equal ($p = 0.573$). This result supports the count heuristic explanation of behavior.

Stability of the average payoffs. The simple counting assumes that the average payoffs remain constant when subjects make their second choice, whereas Bayesian Nash equilibrium and the Naïve herding model predict growth of the payoffs with the choice number. I examine these predictions estimating whether the probability of receiving the payoff μ making the second (even) choice depends on the choice number N_i.

First, I apply a Friedman test on the session averages, assessing the payoffs disparity between each of the second (even) choices. However, I do not find any difference at any conventional significance level ($p = 0.938$). Second, I use the following regression on the subset of the subjects second (even) choices:

$$\mu = \mathcal{L}(\beta_0 + \beta_N N_{i,r} + \beta_r r_i + \beta_T T_i + \beta_{♀} ♀_i + \beta_M M_i + \upsilon_s + \upsilon_i) \tag{9}$$

[7]This result is robust to different specifications and estimation methods including cluster and fixed effect regression analysis.

[8]The regression is estimated on the subset which excludes the choice of the first player in the repetition since the theoretical predictions for these choices are not different. Clearly, since the first choice of the first player assumes low expected payoff, the inclusion of the choices of the first player increases the magnitude and significance of the effect, however does not allow a selecting of the theory.

[9]Throughout the paper the estimations of the exact paired Wilcoxon test are based on the Shift Algorithm by Streitberg and Röhmel (1986).

[10]$\beta_{\odot}^{Count}$ is (log) odds ratio of the expected utility for the first and second choices if the agents use count heuristics. The code for the calculation of this value is available upon request.

Table 3: Subject's Payoff for the Second (Even) Choice – estimation of equation 9 and robustness check

	Dependent variable:			
	Payoff (μ)			
	(1)	(2)	(3)	(4)
Choice number (N)	0.003	0.002	0.002	0.002
	(0.013)	(0.013)	(0.013)	(0.013)
Repetition (r)	0.001	0.001	0.001	0.001
	(0.004)	(0.004)	(0.004)	(0.004)
Treatment (T)	0.373***			
	(0.097)			
Female (♀)	−0.012	−0.022		−0.001
	(0.096)	(0.101)		(0.102)
Math skills (M)	0.043**	0.036*	0.036*	
	(0.019)	(0.020)	(0.020)	
Constant	0.017	0.251	0.243	0.479***
	(0.186)	(0.184)	(0.181)	(0.136)
Observations	3,000	3,000	3,000	3,000
Log Likelihood	−1,970.740	−1,976.420	−1,976.440	−1,978.020
Akaike Inf. Crit.	3,957.490	3,966.840	3,964.880	3,968.050
Bayesian Inf. Crit.	4,005.540	4,008.880	4,000.920	4,004.090
Note:		*p<0.1; **p<0.05; ***p<0.01		

The regression analysis indicates (see table 3) that the coefficient of the choice number is close to zero ($\beta_N = 0.003$; $e^{\beta_N} = 1.003005$) and insignificant ($p = 0.851$).[11] [12] It suggests that the average payoffs do not significantly grow with the choice number. No significant difference in average payoffs for even choices is in line with the predictions of count heuristics and questions the predictive power of both Bayesian Nash Equilibrium and Naïve herding.

4.2 Actual Choice and Count Heuristics

The payoff realization indirectly points to simple counting as the explanation for the subjects' choice. However, the experimental data allows a direct comparison of the counting rules with the other theories by examining the actual choice of the subjects.

Counting prescribes that the subjects' first choice follows the prevailing choice of their predecessors, whereas Bayesian Nash Equilibrium as well as Naïve herding suggest that the subjects' first choice is based solely on the second choice of others and must be independent of the others' first choice. Hence, one can compare the theories by estimating whether the empirical probability of making the first choice according to the Bayesian or Naïve logic is independent of the first choice of the predecessors.

The first choice of the second player $c_{2,1}$ is the best candidate for such an analysis. Subjects make this choice observing only the first and second choices of the first player, where the first choice of the first player is independent of the state of the world. Thus, in Bayesian Nash Equilibrium and Naïve herding model the second players simply follow the second choice of the first player.

They make this choice regardless of whether the first and second choice of the first player coincide or not, $Pr(c_{2,1} = c_{1,2}|c_{1,2} = c_{1,1}) = Pr(c_{2,1} = c_{1,2}|c_{1,2} \neq c_{1,1})$. By contrast, subjects that use counting rules follow prevailing choice and they imitate the predecessor's second choice with higher probability if both choices of the first player are the same, $Pr(c_{2,1} = c_{1,2}|c_{1,2} = c_{1,1}) > Pr(c_{2,1} = c_{1,2}|c_{1,2} \neq c_{1,1})$.

I estimate how the probability of making the same choice as the first player second choice depends on the matching choice of the first player using the next logistic regression:

$$Pr(f = 1) = \mathcal{L}(\beta_0 + \beta_{\ddot{c}}\ddot{c}_r + \beta_r r_i + \beta_T T_i + \beta_{♀} ♀_i + \beta_M M_i + \upsilon_s + \upsilon_i), \qquad (10)$$

where f equals 1 if the second player follows the predecessor's second choice $c_{2,1} = c_{1,2}$ and 0 otherwise, $c_{2,1} \neq c_{1,2}$; $\ddot{c}$ takes the value of 1 when both choices of the second player coincide, $c_{1,2} = c_{1,1}$, and 0 if not, $c_{1,2} \neq c_{1,1}$.

[11]The results hold when I separately apply the same regression for each of the treatments.

[12]If I use the choice number dummies there is no significant difference in the payoffs either.

Table 4: Actual choice and counting – estimation of equation 10 and robustness check

	Dependent variable:		
	Follow the choice, $Pr(f = 1)$		
	(1)	(2)	(3)
Coincide ($\ddot{c}$)	0.898***	0.927***	0.893***
	(0.248)	(0.249)	(0.248)
Repetition (r)	0.069***	0.068***	0.069***
	(0.015)	(0.015)	(0.015)
Treatment (T)	0.664*	0.459	0.642*
	(0.372)	(0.356)	(0.369)
Female (♀)	0.221		
	(0.412)		
Math skills (M)		0.230***	
		(0.071)	
Constant	−0.131	−1.429**	−0.008
	(0.430)	(0.564)	(0.364)
Observations	600	600	600
Log Likelihood	−268.097	−263.357	−268.236
Akaike Inf. Crit.	550.195	540.714	548.472
Bayesian Inf. Crit.	580.973	571.492	574.854

Note: *p<0.1; **p<0.05; ***p<0.01

The results are shown in the table 4. I find strong evidence ($p = 0.000302$) that subjects are more likely to follow their predecessor's second choice if it is the same as the first choice ($\beta_{\ddot{c}} = 0.898$; $e^{\beta_{\ddot{c}}} = 2.454$). Moreover, applying the exact paired Wilcoxon test across aggregated averages over the sessions I can reject the null-hypothesis that there is no difference in probability of following predecessors' second choice if the choices of the first player coincide ($p = 0.00195$).

These results are remarkable – despite the fact that the first choice of the first player does not contain any information about the state of the world it influences the second player's choice. This observation is hard to explain by any type of Bayesian reasoning even if one allows for randomness in behavior, base rate neglect, neglected correlation or another form of Naïve inference.[13] Indeed, the result is in favor of the count heuristics theory.

Interestingly, the regression analysis reported in table 4 reveals that the repetition of game r has a highly significant though very small positive effect on the probability of following the predecessor's choice ($p = 0.000002$;$\beta_{\ddot{c}} = 0.069$; $e^{\beta_{\ddot{c}}} = 1.072$). It suggests that the subjects learn to follow the second choice of the first player over the course of the experiment, but they learn this too slowly.

4.3 Empirically Optimal Choice

The results corroborate counting as an explanation of the subjects' choice. However, it can be also empirically optimal to use counting rules. In this case, the subjects that mimic counting behavior act rationally since they use available information efficiently. Thus, one needs to analyze this possibility to understand the actual level of the subjects' rationality.

Consider using counting rules is empirically optimal. Without loss of generality the probability that the outcome of the first choice $O_{i,1}$ is in favor of state of the world X is more likely if both second and first choices are x than when the first choice is y but the second is x, $Pr(O_{i,1} = 1|c_{i,1} = x, c_{i,2} = x) > Pr(O_{i,1} = 0|c_{i,1} = y, c_{i,2} = x)$. Clearly, if it is not empirically optimal: $Pr(O_{i,1} = 1|c_{i,1} = x, c_{i,2} = x) \leqslant Pr(O_{i,1} = 0|c_{i,1} = y, c_{i,2} = x)$.

To assess this relation I run the two main regressions:

$$Pr(O^X = 1) = \mathcal{L}(\beta_0 + \beta_{\ddot{y}}\ddot{y}_{i,r} + \beta_N N_{i,r} + \beta_r r_i + \beta_T T_i + \beta_{♀}♀_i + \upsilon_s + \upsilon_i) \quad (11)$$

$$Pr(O^Y = 1) = \mathcal{L}(\beta_0 + \beta_{\ddot{y}}\ddot{y}_{i,r} + \beta_N N_{i,r} + \beta_r r_i + \beta_T T_i + \beta_{♀}♀_i + \upsilon_s + \upsilon_i) \quad (12)$$

[13] One could argue that the first choice of the first player can contain payoff-relevant information since it reveals the first player preferences for the first choice. However, in this case, the rational second player should be more likely to follow non coinciding choices of the first player since it realizes more information.

where O^X (O^Y) is a dummy variable that is equal to 1 if the subject receives positive outcome making the choice x (y) or did not receive anything making the choice y (x), $O_{i,1} = 1 \wedge c_{i,1} = x \vee O_{i,1} = 0 \wedge c_{i,1} = y$ ($O_{i,1} = 1 \wedge c_{i,1} = y \vee O_{i,1} = 0 \wedge c_{i,1} = x$). The variable $\ddot{x}$ ($\ddot{y}$) takes value of 1 if $c_{i,1} = x \wedge c_{i,2} = x$ ($c_{i,1} = y \wedge c_{i,2} = y$) and zero if $c_{i,1} = y \wedge c_{i,2} = x$ ($c_{i,1} = x \wedge c_{i,2} = y$).

Table 5: Empirically Optimal Choice – estimation of equation 11, 12

	Dependent variable:	
	Outcome in favor of X $Pr(O^X)$	Outcome in favor Y $Pr(O^Y)$
Coincide x ($\ddot{x}$)		-0.744^{***}
		(0.167)
Coincide y ($\ddot{y}$)	-0.510^{***}	
	(0.145)	
Choice number (N)	-0.065^{***}	-0.032
	(0.024)	(0.022)
Repetition (r)	0.041^{***}	0.020^{***}
	(0.008)	(0.007)
Treatment (T)	0.671^{***}	-0.025
	(0.223)	(0.289)
Female (♀)	0.306	0.017
	(0.223)	(0.178)
Constant	0.752^{***}	1.665^{***}
	(0.264)	(0.300)
Observations	1,460	1,540
Log Likelihood	-776.831	-825.361
Akaike Inf. Crit.	1,569.660	1,666.720
Bayesian Inf. Crit.	1,611.950	1,709.440
Note:		*p<0.1; **p<0.05; ***p<0.01

The results (see table 5) do not show that it is empirically optimal to follow the choices that coincide. On the contrary, the coefficients of coinciding choices are significantly negative both if the choices are x ($p = 0.0004; \beta_{\ddot{x}} = -0.51; e^{\beta_{\ddot{x}}} = 0.6$) and if they are y ($p = 0.000008; \beta_{\ddot{y}} = -0.744; e^{\beta_{\ddot{y}}} = 0.475$).[14] A possible explanation for this finding is that subjects react more strongly to the zero outcome than to the positive one since it negatively reinforces their choice.

[14] If I apply the same regression to each player (from 1 to 5) in sequence separately the coefficients remain always negative, but sometimes lose significance.

It indicates that the subjects that follow the coinciding choice of others' more often than the non coinciding one depart completely from rationality. Their behavior cannot be justified by payoff maximization, but rather corresponds to counting strategies.

4.4 Information Cascades

Now, let's consider the cascading behavior of subjects. Bayesian Nash Equilibrium and the Naïve herding model predict the information cascades from the choice number 6. Nevertheless, cascading behavior is observed in both treatments beginning from the 4th choice – second choice of the second player. In fact, in both treatments the subjects exhibit cascading behavior from the first moment of the interaction between private and public information (see figure 3).

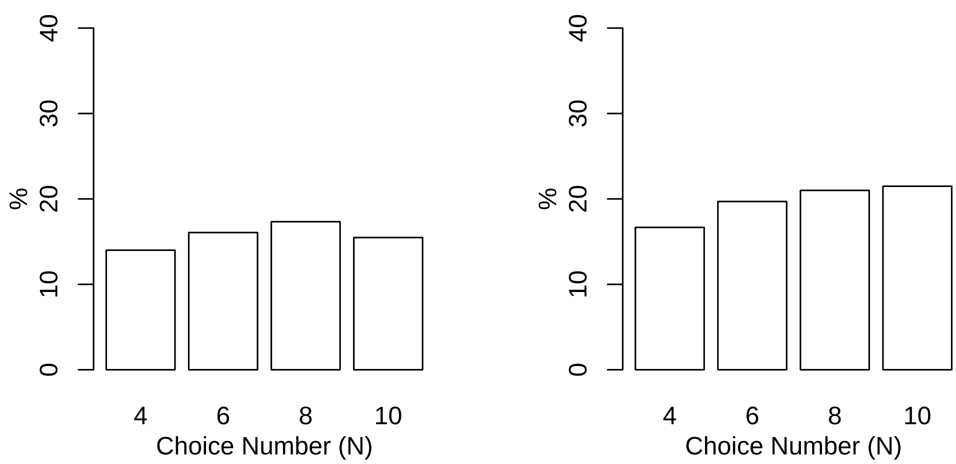


Figure 3: Frequency of cascades in treatment 6/7 (left) and 7/9 (right).

This type of behavior is predicted by the count heuristics theory. Nevertheless, the cascading of the second player as well as cascading in general can be attributed to randomness in choice as well. To disentangle these two explanations I analyse the choice history preceding the cascade of the second player.

Assuming that subjects randomize, the second player is equally likely to cascade regardless of whether the first player's second choice coincides with the first choice of the first player or not. However, the subjects that use counting are more likely to cascade if the first player makes the same choices.

I assess whether the probability to cascade depends on the choice history of the first player with help of the next regression:

$$Pr(\kappa = 1) = \mathcal{L}(\beta_0 + \beta_{\ddot{c}}\ddot{c}_r + \beta_r r_i + \beta_T T_i + \beta_{♀}♀_i + \beta_M M_i + \upsilon_s + \upsilon_i) \tag{13}$$

where $\kappa = 1$ if the second player cascades, that is (s)he follows the second choice of the first player against thier own information $c_{2,2} = c_{1,2} \wedge \mu(c_{2,2}|O_{2,1}) < 0.5$ and $\ddot{c}$ is dummy variable that is equal to 1 if the first player's choices coincide.

The estimation shows that there is positive significant association between cascading behavior of the second player and coincidence of the first two choices of the first player ($p = 0.0249; \beta_{\ddot{c}} = 0.577; e^{\beta_{\ddot{c}}} = 1.78$). This implies that the cascading behavior cannot be simply attributed to the randomness in choice. And furthermore, the subjects exhibit a tendency to comply with the prevailing choice of their predecessors.

In addition, the observed pattern of information cascades frequency highly resembles the patterns in the classical information cascade setting: The cascade frequency tends to gradually grow with choice number.[15] It indicates relatedness of the observed phenomena, suggesting potential generalization of the results.

4.5 Quantal Response Equilibrium

The examination of the subjects' choice demonstrates that simple randomness in the behavior cannot produce the observed patterns. However, it can still be interesting to analyse how the models that incorporate the errors in choice are able to fit the data. The best candidate for such analysis is the logit Quantal Response Equilibrium. The logit Quantal Response Equilibrium (McKelvey and Palfrey, 1995) incorporates errors in behavior using logit specification where the probability of choice x is given by

$$Pr(c^x_{i,n}) = \frac{e^{\lambda\mu^x_{i,n}}}{e^{\lambda\mu^x_{i,n}} + e^{\lambda\mu^y_{i,n}}} \tag{14}$$

and the probability of choice y is $Pr(c^y_{i,n}) = 1 - Pr(c^x_{i,n})$. The parameter λ denotes the level of randomness in the actions of agents. The actions are completely random if $\lambda = 0$, while if $\lambda \to \infty$ the logit Quantal Response Equilibrium approaches the Bayesian Nash Equilibrium.

Despite the fact that Quantal Response Equilibrium does not impose falsifiable restrictions (Haile et al., 2008), it is possible to compare the observed behavior in different games. In the classical information cascade $\lambda \approx 6$ (Goeree

[15] See also the discussion of consistency of the information cascade replication in meta-study of Weizsäcker (2010).

et al., 2007). This value, of course, already suggests a very random behavior. However, using standard maximum likelihood estimation of the λ parameter based on the bandit cascade experimental data (see table 6) I find that $\lambda = 2.886$ ($s.e. = 0.08$, $logL = -3568.154$)[16]. It can be interpreted in following way: the Quantal Response Equilibrium assumes almost completely random behavior and is not rigid to the change in the environment.

Table 6: Parameter estimates for the QRE and QRE-BRF models

	QRE	QRE-BRF
λ	2.886	3.331
s.e.	0.08	0.208
α		0.734
s.e.		0.087
Observations	6,000	6,000
Log Likelihood	-3,568	-3,565

Perhaps, the variation of λ can be attributed to some cognitive bias that gains influence in the different setting and introduces additional distortion in the data not allowing for direct comparison.

To my knowledge the main cognitive bias that is discussed in the literature is the Base Rate Fallacy. In the information cascade setting the Base Rate Fallacy implies that subjects over-weight their private information relative to the public one. Applied to the bandit cascade it can be formalized by the addition of another parameter α in the beliefs function and for the beliefs in the world X it is:

$$\pi^{\alpha}(X|O_{i,n}, H_{i-1}) = \frac{Pr^{\alpha}(O_{i,n}|X)\theta_i^X}{Pr^{\alpha}(O_{i,n}|X)\theta_i^X + Pr^{\alpha}(O_{i,n}|Y)\theta_i^Y} \tag{15}$$

and by analogy it applies to Y, where $\alpha \in (0, \infty)$. The rational agent has $\alpha = 1$, $\alpha > 1$ denotes that subjects relatively over-weight the private information and $\alpha < 1$ that they over-weight the public information. Following (Goeree et al., 2007) I estimate α introducing the new beliefs function in the Quantal Response model and measuring new λ_{α}.

The results of the estimation produce two interesting insights. First, despite the fact that the model employs more parameters than the Quantal Response model, the difference in the likelihood is negligible ($logL_{\alpha} = -3564.648$ vs. $logL = -3568.154$). It suggests no considerable advantage of using the more complicated model. Second, while the model assumes the slightly lower

[16]Appendix C contains R code of the estimations.

rate of randomness in behavior $\lambda_\alpha = 3.331$ ($s.e. = 0.208$) the α parameter is significantly lower than 1 ($\alpha = 0.734$, $s.e. = 0.087$). [17]

The subjects seem to over-weight the public information relatively to private information. This goes against the usual estimations in the information cascade literature, where typically $\alpha > 1$. The observed feature, however, is consistent with the counting rules – counting requires considering each choice of others informative and, hence, to over-weight it as opposed to a private outcome.

The fact that the predictions of the different parametric theories are unstable calls for the theories that have higher predictive power and, ideally, they are simpler. The theory based on simple counting rules appears to be a good candidate.

5 Conclusion

Despite that already in the seminal paper on experimental test of information cascades Anderson and Holt (1997) discuss if cascading behavior can be attributed to the Bayesian or non-Bayesian reasoning, the debates still continues (Eyster et al., 2015; Penczynski, 2016). The design of standard information cascade setting is not well suited to confront various theories in the domain of observational learning (Eyster and Rabin, 2010), therefore, I design and test the model of the bandit cascade.

This model is simple and replicates most of the features of the classical information cascade setting, but allows the behavior based on the counting strategies to be disentangled from the choice motivated by Bayesian reasoning. I reproduce the model in the experiment to reveal the actual subjects' behavior.

The results of the experiment show that the observed behavior can be explained fairly well by counting, but not by Bayesian reasoning. The subjects sustain losses making every first choice and start to ignore the information earlier than Bayesian models predict, but not in a random way. I find that subjects follow the coinciding choice of others, though this behavior is empirically sub-optimal. Moreover, they follow and cascade on choices of others, when these choices contain no information about the state of the world or behavior of others. These facts barely can be reconciled with perfect rationality or even with Bayesian-based models that account for base rate or redundancy neglect or other forms of Naïve inference. It urges to incorporate non-Bayesian based reasoning in the theory of observational learning.

The results of the bandit cascade experiment shed light on the human behavior in the situations when one experiment with alternatives of unknown

[17] A simple t test reject the hypothesis that $\alpha = 1$ with t-statistic of -3.06 ($p = 0.002$).

merit in presence of others actions e.g. technology adoption, managerial decisions, venture capital investment. Since it is demanding to understand the actual motivation for the choice of others (pure experimentation or informed decision), people simply follow prevailing choice disregarding its informativeness. These results provides additional explanation for the failures in the technology adoption (Cowan, 1991), inefficient practice prevalence (Hirshleifer and Welch, 2002) and investment decisions (Dixit, 1992).

References

Anderson, L. R. and Holt, C. A. (1997), 'Information Cascades in the Laboratory', *The American economic review* **87**(5), 847–862.
URL: *http://www.jstor.org/stable/10.2307/2951328*

Azrieli, Y., Chambers, C. and Healy, P. (2012), 'Incentives in experiments: A theoretical analysis', *Working paper* pp. 1–39.
URL: *http://econ.ucsd.edu/ cpchambers/combineddecisions_acts_120425.pdf*

Banerjee, A. (1992), 'A simple model of herd behavior', *The Quarterly Journal of Economics* **107**(3), 797.
URL: *http://qje.oxfordjournals.org/content/107/3/797.short*

Bikhchandani, S., Hirshleifer, D. and Welch, I. (1992), 'A theory of fads, fashion, custom, and cultural change as informational cascades', *The Journal of Political Economy* **100**(5), 992–1026.
URL: *http://www.jstor.org/stable/2138632*

Bradt, R., Johnson, S. and Karlin, S. (1956), 'On sequential designs for maximizing the sum of n observations', *The Annals of Mathematical Statistics* **27**(4), 1060–1074.
URL: *http://www.jstor.org/stable/10.2307/2237195*

Cowan, R. (1991), 'Tortoises and Hares: Choice Among Technologies of Unknown Merit', *The Economic Journal* **101**(407), 801.
URL: *http://www.jstor.org/stable/2233856?origin=crossref*

Dixit, A. (1992), 'Investment and Hysteresis Author (s): Avinash Dixit Published by : American Economic Association Stable URL : http://www.jstor.org/stable/2138376', **6**(1), 107–132.

Eyster, E. and Rabin, M. (2010), 'Naive herding in rich-information settings', *American Economic Journal: Microeconomics* **2**(4), 221–243.
URL: *http://www.aeaweb.org/articles.php?doi=10.1257/mic.2.4.221*

Eyster, E., Rabin, M. and Weizsäcker, G. (2015), 'An Experiment on Social Mislearning'.
URL: *http://scholar.harvard.edu/files/rabin/files/erw2.pdf?m=1444731345*

Fischbacher, U. (2007), 'z-Tree: Zurich toolbox for ready-made economic experiments', *Experimental economics* **10**(2), 171–178.
URL: *http://link.springer.com/10.1007/s10683-006-9159-4*

Goeree, J. K., Palfrey, T. R., Rogers, B. W. and McKelvey, R. D. (2007), 'Self-Correcting Information Cascades', *Review of Economic Studies* **74**(3), 733–762.
URL: *http://restud.oxfordjournals.org/lookup/doi/10.1111/j.1467-937X.2007.00438.x*

Greiner, B. (2004), 'An online recruitment system for economic experiments', pp. 79–93.
URL: *http://mpra.ub.uni-muenchen.de/13513/*

Haile, P., Hortaçsu, A. and Kosenok, G. (2008), 'On the empirical content of quantal response equilibrium', *The American Economic Review* **98**(1), 180–200.
URL: *http://www.jstor.org/stable/10.2307/29729968*

Hirshleifer, D. and Welch, I. (2002), 'An Economic Approach to the Psychology of Change: Amnesia, Inertia, and Impulsiveness', *Journal of Economics & Management Strategy* **11**(3), 379–421.
URL: *http://www.ingentaselect.com/rpsv/cgi-bin/cgi?ini=xref&body=linker&reqdoi=10.1162/105864002320272549*

Huck, S. and Oechssler, J. (2000), 'Informational cascades in the laboratory: Do they occur for the right reasons?', *Journal of Economic Psychology* **21**(6), 661–671.
URL: *http://www.sciencedirect.com/science/article/pii/S0167487000000258*

Hung, A. A. and Plott, C. C. (2001), 'Information cascades: Replication and an extension to majority rule and conformity-rewarding institutions', *The American economic review* **91**(5), 1508–1520.
URL: *http://www.jstor.org/stable/2677936*

Kübler, D. and Weizsäcker, G. (2004), 'Limited depth of reasoning and failure of cascade formation in the laboratory', *Review of Economic Studies* **71**(2), 425–441.
URL: *http://restud.oxfordjournals.org/content/71/2/425.short*

March, C., Krügel, S. and Ziegelmeyer, A. (2012), 'Do we follow private information when we should? Laboratory evidence on naive herding', *Jena Economic Research Papers* **N2012**(002), 0–26.
URL: *https://opus.zbw-kiel.de/dspace/handle/10419/56873*

McKelvey, R. and Palfrey, T. (1995), 'Quantal response equilibria for normal form games', *Games and economic behavior* **10**(1), 6–38.
URL: *http://www.sciencedirect.com/science/article/pii/S0899825685710238*

Noth, M. and Weber, M. (2003), 'Information Aggregation with Random Ordering: Cascades and Overconfidence*', *The Economic Journal* **113**(484), 166–189.
URL: *http://doi.wiley.com/10.1111/1468-0297.00091*

Park, A. and Sabourian, H. (2011), 'Herding and Contrarian Behaviour in Financial Markets', *Econometrica* **79**(4), 973–1026.
URL: *http://www.jstor.org/stable/4132758*

Penczynski, S. P. (2016), 'The Nature of Social Learning : Experimental Evidence', (June).
URL: *http://www.penczynski.de/attach/Sociallearning.pdf*

Smith, L. and Sorensen, P. (2000), 'Pathological outcomes of observational learning', *Econometrica* **68**(2), 371–398.
URL: *http://www.jstor.org/stable/2999431*

Streitberg, B. and Röhmel, J. (1986), 'Exact distributions for permutations and rank tests: An introduction to some recently published algorithms', *Statistical Software Newsletter* **12**(1), 10–17.
URL: *http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Exact+distributions+f*

Weizsäcker, G. (2010), 'Do we follow others when we should? A simple test of rational expectations', *The American Economic Review* **100**(5), 2340–2360.
URL: *http://www.jstor.org/stable/10.2307/41038765*

Ziegelmeyer, A., Koessler, F., Bracht, J. and Winter, E. (2009), 'Fragility of information cascades: an experimental study using elicited beliefs', *Experimental Economics* **13**(2), 121–145.
URL: *http://www.springerlink.com/index/10.1007/s10683-009-9232-x*

A Proofs

A.1 Proof of Proposition 1

(a) Suppose the first two agents in a sequence choose the same alternative when they make choice. That is, without loss of generality the choice history is $H_2 = (x, x, x, x)$. Then, the public belief of the third agent $i = 3$ that the state of the world is X:[18] $\theta_3^X = \frac{q^2}{2q^2-2q+1} > q$ if $q \in (1/2, 1)$ □

(b) Without loss of generality let the public belief be in favor of the state X, $\theta_i^X > 1/2$. If an agent makes the first choice $n = 1$, $O_{i,1} = \varnothing$, hence, $\pi_i^X = \theta_i^X > 1/2$ and the agent chooses alternative x. If an agent makes the second choice $n = 2$, then there are two cases: the private and public beliefs are in favor of the same world, $\theta_i^X > 1/2 \wedge \pi_i^X > 1/2$, or they are not, $\theta_i^X > 1/2 \wedge \pi_i^Y > 1/2$. The agents expected utility from taking the action x is:

$$\mu_{i,2}^X = \begin{cases} \frac{\theta_i^X q^2}{\theta_i^X q+(1-\theta_i^X)(1-q)} + \frac{(1-\theta_i^X)(1-q)^2}{\theta_i^X q+(1-\theta_i^X)(1-q)} & \text{if } \pi_i^X > 1/2 \\ \frac{(1-\theta_i^X)q^2}{(1-\theta_i^X)q+\theta_i^X(1-q)} + \frac{\theta_i^X(1-q)^2}{(1-\theta_i^X)q+\theta_i^X(1-q)} & \text{if } \pi_i^Y > 1/2 \end{cases} \tag{16}$$

If $\theta_i^X > q$, $\mu_{i,2}^X > 1/2$ in both cases and, hence, the agent chooses alternative x in both cases as well. Thus, $Pr(c_{i,n}|\pi_{i,n}) = Pr(c_{i,n}) \Leftrightarrow Pr(c_{i,n}|O_{i,n}) = Pr(c_{i,n})$ □

A.2 Proof of Proposition 2

Note that $\forall(i > 2)$: $Pr(\theta_i \leqslant q) = (q(1-q))^i \Leftrightarrow \forall(i > 2)$: $Pr(\theta_i > q) = 1 - (q(1-q))^i$. Thus, $\lim_{i\to\infty} Pr(\theta_i > q) = \lim_{i\to\infty}(1 - (q(1-q))^i) = 1$ if $q \in (1/2, 1)$ □

B Instructions

B.1 Treatment I

Welcome to the experiment!

Thank you very much for participating. We hope that you feel comfortable. We ask you to remain quiet and do not communicate with any other players. Please understand that in case you communicate with other players we will have to exclude you from the experiment without payment. If you have

[18] Note that in pure strategies Bayesian Nash Equilibrium the agent choose the alternative with highest expected utility with probability 1.

any questions please raise your hand and wait for the experimenter to come to you.

The experiment is on decision-making. Your earnings will depend partly on your decisions and partly on chance. The earnings will be calculated in points. At the end of the experiment **three decisions of you** will be randomly chosen. **The earnings from these decisions** will be summed up and converted to Euros with the following rate:

1 point = 3 euro

In addition to the earnings from your decisions, you will receive 3 euro as a compensation for showing up on time.

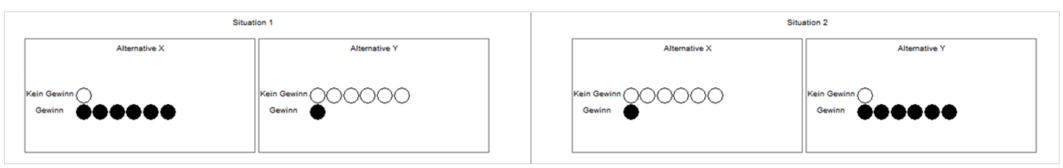
During the experiment you will play **the same simple game 30 times**. Each time you play the game you will have to make **2 decisions.**

Each time you make a decision, you will have to choose between two alternatives: **X, Y**. Each of the alternatives each time it is chosen can bring the reward equal to **1 point**.

In the beginning of each game a computer will toss a coin. The coin has **equal chance** to fall on each side. The side on which it falls will determine in which of the two possible situations you **and other players** are.

In the situation 1, the alternative X will bring the reward (1 point) in 6 out of 7 cases and alternative Y will bring the reward (1 point) in 1 out of 7 cases.

In the situation 2, the alternative Y will bring the reward (1 point) in 6 out of 7 cases and alternative X will bring the reward (1 point) in 1 out of 7 cases. The following scheme summarizes these two situations:



You and other players will make your choices in **a predetermined order**. Your place in the order is **determined randomly** in the beginning of each game and does not depend on your choice. In each game you will observe **the choices that others have already made** in this game.

B.2 Treatment II

Welcome to the experiment!

Thank you very much for participating. We hope that you feel comfortable. We ask you to remain quiet and do not communicate with any other players. Please understand that in case you communicate with other players we will have to exclude you from the experiment without payment. If you have any questions please raise your hand and wait for the experimenter to come to you.

The experiment is on decision-making. Your earnings will depend partly on your decisions and partly on chance. The earnings will be calculated in points. At the end of the experiment **three decisions of you** will be randomly chosen. **The earnings from these decisions** will be summed up and converted to Euros with the following rate:

1 point = 3 euro

In addition to the earnings from your decisions, you will receive 3 euro as a compensation for showing up on time.

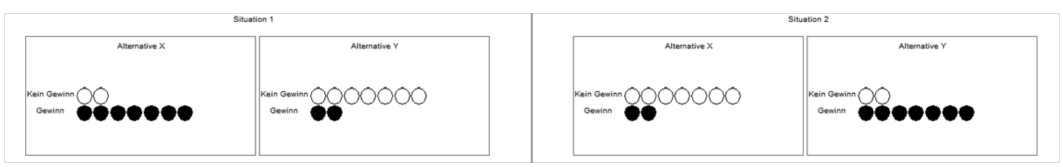
During the experiment you will play **the same simple game 30 times**. Each time you play the game you will have to make **2 decisions.**

Each time you make a decision, you will have to choose between two alternatives: **X, Y**. Each of the alternatives each time it is chosen can bring the reward equal to **1 point**.

In the beginning of each game a computer will toss a coin. The coin has **equal chance** to fall on each side. The side on which it falls will determine in which of the two possible situations you **and other players** are.

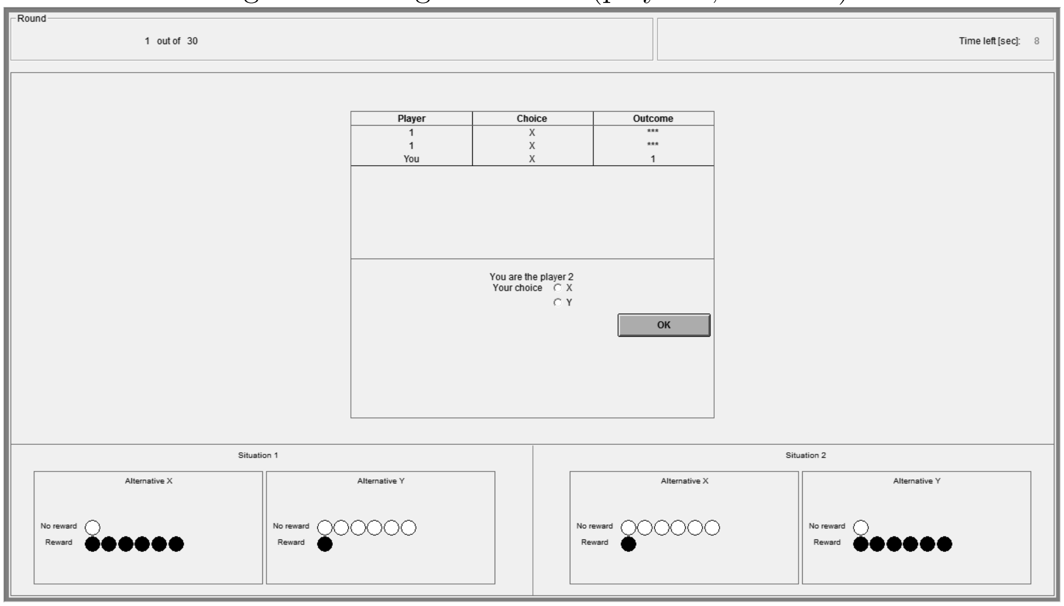
In the situation 1, the alternative X will bring the reward (1 point) in 7 out of 9 cases and alternative Y will bring the reward (1 point) in 2 out of 9 cases.

In the situation 2, the alternative Y will bring the reward (1 point) in 2 out of 9 cases and alternative X will bring the reward (1 point) in 7 out of 9 cases. The following scheme summarizes these two situations:



You and other players will make your choices in **a predetermined order**. Your place in the order is **determined randomly** in the beginning of each game and does not depend on your choice. In each game you will observe **the choices that others have already made** in this game.

Figure 4: The game screen (player 2, choice 2)



C Estimation Program

In R function "logLQRE" I use the experimental data that is stored in the data frame called "data". The coding is as follows: x choices are labelled by a 1 and y choices by a 0; Reward is denoted by a 1 and no reward by a 0; "Nchoice" variable is a choice number in one game. The function estimate log-likelihood given the randomness parameter "lambda".

```
logLQRE <- function( lambda, data ){

  logL = 0
  m = 1
  #Number of choices in one game
  T = max ( data$Nchoice )
  #Number of repetitions of the game
  M = length ( data$Nchoice ) / max ( data$Nchoice )


  while ( m <= M ){
    p = 1/2 # Note: Prior belief that the state of the world X
    t = 1   # Note: Counter

    while ( t <= T ){

      Nchoice = data$Nchoice[ ( m - 1 ) * T + t ]
```

```
#Note: q - the probability to recieve the reward
q = data$q[ ( m - 1 ) * T + t ]

#Note: First choice of a player

if ( Nchoice %% 2 == 1 ){

  #Note: BNE probabilty to choose action x or y
  PSx = p * q + ( 1 - p ) * ( 1 - q )
  PSy = p * ( 1 - q ) + ( 1 - p ) * q

  Pchoicex = 1 / ( 1 + exp( lambda * ( 1 - 2 * PSx ) ) )
  Pchoicey = 1 - Pchoicex

  choice_1 = data$choice[ ( m - 1 ) * T + t ]

  if (choice_1 == 0) {p = p; logL = logL + log ( Pchoicex )}
  if (choice_1 == 1) {p = p; logL = logL + log ( Pchoicey )}


}

#Note: Second choice of a player
if ( Nchoice %% 2 == 0 ){

  #Note: Beliefs about the state of the world
  #given outcome of the choice
  PX_SxorNSy = q * p / ( q * p + ( 1 - q ) * ( 1 - p ) )
  PY_SxorNSy = ( 1 - q ) * ( 1 - p )/(q * p + ( 1 - q ) * ( 1 - p ) )

  PX_NSxorSy = ( 1 - q ) * p / ( ( 1 - q ) * p + q * ( 1 - p ) )
  PY_NSxorSy = q * ( 1 - p ) / ( ( 1 - q ) * p + q * ( 1 - p ) )

  PX_SyorNSx = ( 1 - q ) * p / ( ( 1 - q ) * p + q * ( 1 - p ) )
  PY_SyorNSx = q * ( 1 - p ) / ( ( 1 - q ) * p + q * ( 1 - p ) )

  PX_NSyorSx = q * p / ( q * p +( 1 - q ) * ( 1 - p ) )
  PY_NSyorSx = ( 1 - q ) * ( 1 - p ) / ( q * p + ( 1 - q ) * ( 1 - p ) )

  #Note: BNE probabilty to choose action x or y
  PSx_SxorNSy = PX_SxorNSy * q + PY_SxorNSy * ( 1 - q )
  PSx_NSxorSy = PX_NSxorSy * q + PY_NSxorSy * ( 1 - q )

  PSy_SyorNSx = PX_SyorNSx * ( 1 - q ) + PY_SyorNSx * q
  PSy_NSyorSx = PX_NSyorSx * ( 1 - q ) + PY_NSyorSx * q


  #Note: Probabilty to choose x or y action
  #given the level of randmoness parameter lambda.
```

```
Pchoicex_SxorNSy = 1 / ( 1 + exp( lambda * ( 1 - 2 * PSx_SxorNSy ) ) )
Pchoicex_NSxorSy = 1 / ( 1 + exp( lambda * ( 1 - 2 * PSx_NSxorSy ) ) )

Pchoicey_SxorNSy = 1 - Pchoicex_SxorNSy
Pchoicey_NSxorSy = 1 - Pchoicex_NSxorSy

Pchoicex_X = Pchoicex_SxorNSy * q + Pchoicex_NSxorSy * ( 1 - q )
Pchoicex_Y = Pchoicex_SxorNSy * ( 1 - q )+Pchoicex_NSxorSy * q

Pchoicey_X = Pchoicey_SxorNSy * ( 1 - q ) + Pchoicey_NSxorSy * q
Pchoicey_Y = Pchoicey_SxorNSy * q + Pchoicey_NSxorSy * ( 1 - q )

#Note: Plugin transition probabilties in the Beleif function

PBeliefX_x =

  p * Pchoicex_X  / ( p * Pchoicex_X+( 1 - p ) * Pchoicex_Y )

PBeliefX_y =

( 1 - p ) * Pchoicex_Y  / ( p * Pchoicex_X+( 1 - p ) * Pchoicex_Y )

outcome_1 = data$Out[ ( m - 1 ) * T + t - 1 ] #Outcome
choice_1  =  data$choice[ ( m - 1 ) * T + t - 1]
choice_2  =  data$choice[ ( m - 1 ) * T+t]

if ( choice_1 == 0 ){

  if ( outcome_1 == 1 & choice_2 == 0 )
  { p = PBeliefX_x; logL = logL + log ( Pchoicex_SxorNSy ) }

  if ( outcome_1 == 0 & choice_2 == 0)
  { p = PBeliefX_x; logL = logL + log ( Pchoicex_NSxorSy ) }

  if ( outcome_1 == 1 & choice_2 == 1)
  { p = PBeliefX_y; logL = logL + log ( Pchoicey_SxorNSy ) }

  if (outcome_1 == 0 & choice_2 == 1)
  { p = PBeliefX_y; logL = logL + log ( Pchoicey_NSxorSy ) }

}

if ( choice_1 == 1 ){

  if ( outcome_1 == 0 & choice_2 == 0 )
  { p = PBeliefX_x; logL = logL + log ( Pchoicex_SxorNSy ) }

  if ( outcome_1 == 1 & choice_2 == 0)
```

```
        { p = PBeliefX_x; logL = logL + log ( Pchoicex_NSxorSy ) }

        if ( outcome_1 == 0 & choice_2 == 1 )
        { p = PBeliefX_y; logL = logL + log ( Pchoicey_SxorNSy ) }

        if ( outcome_1 == 1 & choice_2 == 1 )
        { p = PBeliefX_y; logL = logL + log ( Pchoicey_NSxorSy ) }


      }

    }

    t = t + 1
   }
   m = m + 1
  }

  return ( logL )

}
```