

Distinguishing Informational Cascades from Herd Behavior in the Laboratory*

Boğaçhan Çelen[†]
New York University

Shachar Kariv[‡]
University of California, Berkeley

October 22, 2003

Abstract

This paper reports an experimental test of how individuals learn from the behavior of others. By using techniques only available in the laboratory, we elicit subjects' beliefs. This allows us to distinguish informational cascades from herd behavior. By adding a setup with continuous signal and discrete action, we enrich the ball-and-urn observational learning experiments paradigm of Anderson and Holt (1997). We attempt to understand subjects' behavior by estimating a model that allows for the possibility of errors in earlier decisions. (*JEL* C92, D8).

*This research was supported by the Center for Experimental Social Sciences (C.E.S.S.) and the C. V. Starr Center for Applied Economics at New York University. We are grateful to Andrew Schotter for his guidance and to three anonymous referees for their comments. We benefited from the expositional suggestions of William J. Baumol. We also acknowledge helpful discussions of Colin F. Camerer, Liran Einav, Xavier Gabaix, Douglas Gale, Charles Holt, David Laibson, and Matthew Rabin. This paper has also benefited from suggestions by the participants of the 2001 ESA North America Regional Conference and seminars at several universities.

[†]Department of Economics, New York University, 269 Mercer St., 7th Floor, New York, NY, 10003 (e-mail: bc319@nyu.edu, url: <http://home.nyu.edu/~bc319>).

[‡]Department of Economics, University of California, Berkeley, Evans Hall # 3880, Berkeley, CA 94720 (e-mail: kariv@econ.berkeley.edu, url: <http://socrates.berkeley.edu/~kariv/>).

1 Introduction

In many social and economic situations, individuals are influenced by the decisions of others. The commonest examples occur in everyday life, as in choosing a fashionable restaurant or a popular movie. But it has also been suggested that similar influences affect technology adoption and asset market decisions. For rational choice theory, however, the important question is why rational maximizing individuals should behave in this way.

In recent years, a great deal of attention has been paid to the *social learning* literature introduced by Banerjee (1992), and Bikhchandani, Hirshleifer, and Welch (1992), describing situations in which individuals learn by observing the behavior of others. This literature analyzes an economy where a sequence of Bayesian individuals make a once-in-a-lifetime decisions under incomplete and asymmetric information¹. The typical conclusion is that, despite the asymmetry of information, eventually every individual imitates her predecessor, even though she would have chosen differently if she had acted on her own information alone. In this sense, individuals rationally ignore their own information and follow the herd.

Two phenomena that have elicited particular interest are *informational cascades* and *herd behavior*, which can arise in a wide variety of economic circumstances². These phenomena have been deemed pathological because erroneous outcomes may occur despite individual rationality, and they may in fact be the norm in certain circumstances. While the terms informational cascade and herd behavior are used interchangeably in the literature, Smith and Sørensen (2000) emphasize that there is a significant difference between them. An informational cascade is said to occur when an infinite sequence of individuals ignore their private information when making a decision, whereas herd behavior occurs when an infinite sequence of individuals make an identical decision, not necessarily ignoring their private information.

In other words, when acting in a herd, individuals choose the same action, but they may have acted differently from one another if the realization of their private signals had been different. In an informational cascade, an individual considers it optimal to follow the behavior of her predecessors without regard to her private signal since her belief is so strongly held that no signal can outweigh it. Thus, an informational cascade implies a herd but a herd is not necessarily the result of an informational cascade.

The practical importance of the distinction between herds and cascades is that in a cascade social learning ceases since individual behavior becomes purely imitative and hence is uninformative. In a herd, in contrast, individuals become more and more likely to imitate but their actions still may provide information. Thus, the distinction is related to the social welfare properties rather than the informational properties per se. When acting in a herd, a group settles on a

¹For surveys see: Gale (1996), and Bikhchandani, Hirshleifer, and Welch (1998).

²For examples see, Finance: Scharfstein and Stein (1990); Welch (1992); Avery and Zemsky (1998); Welch (2000). Auctions: Neeman and Orosel (1999). Political Economy: Morton and Williams (1999). Industrial Organization: Kennedy (2002).

single pattern of behavior and, at the same time, the behavior is fragile in the sense that a strong signal may cause behavior to shift suddenly and dramatically. In contrast, a cascade is stable, i.e., no signal can cause a change in the pattern of behavior. Hence, the distinction between herds and cascades, so far not addressed by the experimental literature, sheds light on questions such as why mass behavior is so fragile and prone to fads.

Apparently, informational cascades, which are defined in terms of (unobservable) beliefs, are much harder to identify than herds, which are defined in terms of (observable) actions. In market settings, we observe behavior but not beliefs or private information. In the laboratory, in contrast, we can elicit subjects' beliefs and control their private information. A novel setup and an elicitation technique enable us to distinguish informational cascades from herd behavior.

The paper reports an experimental test of a model based on Çelen and Kariv (2002)³. We employ a design in which a sequence of subjects draw private signals from a uniform distribution over $[-10, 10]$. The decision problem is to predict whether the sum of all subjects' signals is positive or negative and to choose an appropriate action, A or B . A is the profitable action when this sum is positive and B if it is not. However, instead of choosing action A or B directly, after being informed about the history of actions of others and before observing their own private signals, subjects are asked to select a cutoff such that action A will be chosen if the signal received is greater than the cutoff and action B otherwise. Only after a subject reports her cutoff, is she informed of her private signal, and her action is recorded accordingly.

As there is a one-to-one relation between subjects' cutoffs and their beliefs about the true state of the world, cutoff data enable us to determine which subjects exhibit *cascade behavior*, i.e., acting irrespective of her private signal. Such a subject is one who reports a cutoff -10 or 10 . In contrast, a subject who joins a herd but does not follow a cascade behavior is one who reports a cutoff in the interval $(-10, 10)$, indicating that for some signal she is willing to make either decision, but when her private signal is realized she acts as her predecessors did. Hence, cascade behavior is identifiable by the choice of a cutoff while joining of a herd is identifiable from the realized action.

In the laboratory, we find that herd behavior occurs frequently as do cascades. However, not all observed herds are cascades. Since the theoretical result predicts that an informational cascade is impossible, we explain why cascades often arise in the laboratory as a particular type of deviation from Bayes rationality. For this purpose, we generalize the Bayesian model by allowing the possibility that subjects make error and that they incorporate the possibility that others are making errors into their beliefs. Our results suggest that Bayesian models properly generalized to take account of human error successfully predict subjects' behavior in the laboratory and provide an explanation for cascade behavior.

The paper is organized as follows. The next section discusses continuum-

³This paper, focuses on observational learning under imperfect information. The conventional perfect information setup is analyzed as a benchmark.

signal and discrete-signal setups. Section 3 describes the experimental design and procedures, and section 4 outlines the underlying decision problem. Section 5 summarizes the experimental results. Section 6 provides a modification of the model by introducing the possibility of noise in the model and its estimation, and section 7 concludes.

2 Continuum versus Discrete Setups

The most comprehensive study on social learning is provided by Smith and Sørensen (2000). One of their departures from Bikhchandani, Hirshleifer and Welch (1992) is the assumption that individuals may observe general signals. So while Bikhchandani, Hirshleifer and Welch (1992) find that a cascade and, consequently, a herd occurs, Smith and Sørensen (2000) show that with a continuous signal space herd behavior arises, yet there need be no informational cascade. That is, even during herding, when making a decision everyone may take private signals into account. If the signals were different, the individuals' actions might also change.

In a seminal paper, Anderson and Holt (1997) investigate social learning experimentally⁴. Their design is based on the binary-signal-binary-action model of Bikhchandani, Hirshleifer and Welch (1992)⁵. In their setup there are two decision-relevant events, say A and B , equally likely to occur *ex ante* and two corresponding signals a and b . Signals are informative in the sense that there is a probability higher than $1/2$ that a signal matches the label of the realized event. The decision to be made by the experimental subject is a prediction of which of the events will take place, basing the forecast on a private signal and the history of past decisions. In such a binary signal structure, whenever two consecutive decisions coincide, say both predict A , the subsequent individual should also choose A even if her private signal is different, b . Anderson and Holt (1997) identify rational cascades as a case in which a subject observes two consecutive identical decisions and despite her contrary private information, chooses the same action. In their setup, this is the only way one can detect informational cascades. They report that rational cascades formed in most rounds and that about half of the cascades were incorrect.

In our setup, unlike Anderson and Holt (1997), while there are two events which, *ex ante*, are equally likely to occur, there is a continuous signal space. The following example illustrates the importance of this difference. Suppose that a sequence of individuals who act alike is followed by someone who deviates. What may successors conclude by observing the deviation? In Anderson and Holt (1997)'s discrete signal world, such a deviation is impossible when individuals are Bayes-rational, but in a continuous signal world, successors might instead

⁴Following Anderson and Holt (1997), a number of experimental papers analyzed several aspects of social learning. Among others, Hung and Plott (2001), and Kübler and Weizsäcker (2003) extend Anderson and Holt (1997) to investigate further possible explanations for cascade behavior.

⁵Anderson and Holt (1996) describe a simple classroom setting of such an experiment.

infer that the deviator has private information that is so convincing that it leads her to deviate. That is, the deviator’s private signal is so strong that it provokes a rational deviation.

To summarize, in the discrete-signal-discrete-action setup all herds are cascades since once two consecutive decisions coincide no signal can lead to a deviation. In contrast, our continuous-signal-discrete-action setup, along with our belief elicitation method, enables us to distinguish cascades and herds completely.

3 Experimental Design

The experiment was run at the Experimental Economics Laboratory of the Center for Experimental Social Sciences (C.E.S.S.) at New York University. The 40 subjects in this experiment were recruited from undergraduate economics classes at New York University and had no previous experience in social learning experiments. In each session eight subjects participated as decision-makers. After subjects read the instructions (the instructions are available upon request) they were also read aloud by an experimental administrator⁶. The experiment lasted for about one and a half hours. A \$5 participation fee and subsequent earnings for correct decisions, which averaged about \$22, were paid in private at the end of the session. Throughout the experiment, we ensured anonymity and effective isolation of subjects⁷ in order to minimize any interpersonal influences that could stimulate uniformity of behavior.

Each experimental session entailed 15 independent rounds, each divided into eight decision-turns. In each round, all eight subjects took decisions sequentially in a random order. A round started by having the computer draw eight numbers from a uniform distribution over $[-10, 10]$. The numbers drawn in each round were independent of each other and of the numbers in any of the other rounds. Each subject was informed only of the number corresponding to her turn to move. The value of this number was her private signal. In practice, subjects observed their signals up to two decimal points.

Upon being called to participate, a subject first observed the history of the actions taken her predecessors in that round. After this and before being informed of her private signal, each subject was asked to select a number between -10 and 10 (a cutoff), for which she would take action A if her signal was above the cutoff and action B if it was not. Action A was profitable if and only if the sum of the eight numbers was positive and action B otherwise. Only after submitting her decision, the computer informed her of the value of her private

⁶At the end of the first round, subjects were asked if there were any misunderstandings. No subject reported any problems with understanding the procedures or using the computer program.

⁷Participants’ working-stations were isolated by cubicles making it impossible for participants to observe other’s screens or to communicate. We also made sure that all remained silent throughout the session. At the end of a session, participants were paid in private according to the number of their working-stations.

signal. Then, the computer recorded her decision as A if the signal was higher than the cutoff she selected. Otherwise, the computer recorded her action as B .

After all subjects had made their decisions, the computer informed everyone what the sum of the eight numbers actually was. All participants whose decisions determined A as their action earned \$2 if this sum was positive (or zero) and nothing otherwise. Similarly, all whose decisions led to action B earned \$2 if this sum was negative and nothing otherwise. This process was repeated in all rounds. Each session was terminated after all 15 rounds were completed.

4 Some Theory

4.1 The Bayesian Solution

To formulate the Bayesian solution of the decision problem underlying our experimental design, suppose that each individual $n \in \{1, \dots, 8\}$ receives a private signal θ_n drawn from a uniform distribution with support $[-10, 10]$. Assume that private signals are *i.i.d.* across individuals. Each individual n has to make a binary decision $x_n \in \{A, B\}$ in a sequential order where action A is profitable if and only if $\sum_{i=1}^8 \theta_i \geq 0$, and action B is the profitable one otherwise⁸. All decisions are announced publicly and therefore known to all successors.

Note that the decision problem involves incomplete and asymmetric information. That is, individuals are uncertain about the underlying decision-relevant event, $\sum_{i=1}^8 \theta_i \geq 0$ or $\sum_{i=1}^8 \theta_i < 0$, and the information about it is shared asymmetrically among them. Further, there is no private signal which can enable any individual to resolve the uncertainty by herself. This is referred in the theoretical literature as a case of *bounded beliefs*.

The optimal decision rule of individual n can be summarized as

$$x_n = A \text{ if and only if } \mathbb{E} \left[\sum_{i=1}^8 \theta_i \mid \theta_n, (x_i)_{i=1}^{n-1} \right] \geq 0,$$

Since no one has any information about her successors' signals, we get

$$x_n = A \text{ if and only if } \mathbb{E} \left[\sum_{i=1}^n \theta_i \mid \theta_n, (x_i)_{i=1}^{n-1} \right] \geq 0.$$

Hence,

$$x_n = A \text{ if and only if } \theta_n \geq -\mathbb{E} \left[\sum_{i=1}^{n-1} \theta_i \mid (x_i)_{i=1}^{n-1} \right].$$

It readily follows that the optimal decision, as a function of the realized history of actions, follows *cutoff strategy*

$$x_n = \begin{cases} A & \text{if } \theta_n \geq \hat{\theta}_n, \\ B & \text{if } \theta_n < \hat{\theta}_n, \end{cases} \quad (1)$$

⁸A note of clarification: From a technical point of view, this setup is different from the standard herding models, as private signals are conditionally dependent, i.e., conditional on the state of the world, signals are negatively correlated. In general, correlated signals make the model very hard to solve; in our case, however, the reverse is true. Our results by and large do not depend on the conditional dependence assumption.

where

$$\hat{\theta}_n = -\mathbb{E} \left[\sum_{i=1}^{n-1} \theta_i \mid (x_i)_{i=1}^{n-1} \right] \quad (2)$$

is the optimal history-contingent cutoff.

Note that $\hat{\theta}_n$ inherits all the information that individual n learns from the history of actions. As such, it determines the minimum private signal for which she optimally decides to choose action A . Hence, $\hat{\theta}_n$ is sufficient to characterize individual n 's behavior, and thus, the process of cutoffs $\{\hat{\theta}_n\}$ characterizes the social behavior. That is why we take it as the object of the experimental design and analysis.

4.2 An Illustration

To provide some intuitive interpretation, we discuss the Bayesian reasonings of the first few individuals. The first individual's decision must be based solely on her private signal. Thus the expected value of any of her successors' signals conditional on her information is zero. Hence, her cutoff is $\hat{\theta}_1 = 0$ and she takes action A if and only if $\theta_1 \geq 0$ and action B otherwise. Since the second individual observes the first's action, she conditions her decision on x_1 . Thus, according to (2)

$$\hat{\theta}_2 = \begin{cases} -5 & \text{if } x_1 = A, \\ 5 & \text{if } x_1 = B. \end{cases}$$

To clarify, if for example $x_1 = A$ then $E[\theta_1 \mid x_1 = A] = E[\theta_1 \mid \theta_1 \geq 0] = 5$ and thus it is optimal for the second individual to take action A if and only if $\theta_2 \geq -5$. Similarly, if $x_1 = B$ it is optimal for her to take action A if and only if $\theta_2 \geq 5$.

Note that the second individual may imitate the first even though she would have made a contrary decision had she based her decision solely on her own signal. Moreover, any deviation of the second individual reveals that her private signal is contrary to and stronger than the expected value of the first's private signal. Therefore, when the third individual observes a deviation, her cutoff is more sensitive to the second's action. By (2), a simple computation yields the third individual's cutoff rule:

$$\hat{\theta}_3 = \begin{cases} -7.5 & \text{if } x_1 = A, x_2 = A, \\ -2.5 & \text{if } x_1 = B, x_2 = A, \\ 2.5 & \text{if } x_1 = A, x_2 = B, \\ 7.5 & \text{if } x_1 = B, x_2 = B. \end{cases}$$

Proceeding with the same analysis, we find that if the first three individuals choose A , the fourth individual will choose A as long as $\hat{\theta}_4 \geq -8.75$; if the first four individuals choose A , the fifth individual will choose A as long as $\hat{\theta}_4 \geq -9.375$; and so on. Thus, the longer the sequence of individuals who choose A , the harder it is for a single individual to choose action B , even if her private signal is very negative.

An important concept discussed in Smith and Sørensen (2000) is the *overturning principle*. This asserts that even if many individuals have acted alike, following a rational deviation, the information revealed from the history to that point (almost) cancels out. For successors must conclude that the final deviant received a strong signal favoring the contrary action, and unconditional on any new signals, they should infer this fact, and follow suit.

[Figure 1 here]

For example, if the fourth individual chooses B after her three predecessors choose A , her action reveals that her signal is in the interval $[-10, -8.75]$. In such a case, according to (2), the fifth individual will choose A as long as $\hat{\theta}_5 \geq 0.625$. In sum, the longer the sequence of individuals acting alike, the larger the asymmetry between the information revealed by an imitation and a deviation. Moreover, no matter how many individuals have acted alike, it is always possible that one individual who receives an extreme signal will not follow the historic pattern to that point. Figure 1 illustrates the sequences of cutoffs for two histories. In sequence one, all individuals choose action A . Therefore, as time passes individuals become more confident that action A is the profitable one. In sequence two, the first two individuals take action A and all subsequent individuals take action B . Thus, the third individual deviates, revealing information indicating strongly that her signal favors action B . Because of this newly revealed information, the fourth individual's cutoff is very close to zero but yet favors action B .

4.3 Some Definitions

Next, we define some key concepts to which we refer throughout the paper. We identify a subject who engages in *cascade behavior* as one who reports a cutoff -10 or 10 , and thus takes either action A or B , no matter what her private signal is.

We say that an *informational cascade* occurs in the laboratory when beginning with some subject, either all report cutoffs -10 , or all report cutoff 10 , and *herd behavior* occurs when, beginning with some subject, all take the same action.

Therefore, a subject who joins a herd but does engage in cascade behavior is one who reports a cutoff in the interval $(-10, 10)$, indicating that there are some signals that can lead her to choose action A , some to choose B but when her private signal is realized she will act as her predecessors did.

4.4 A Note on Herds and Cascades

Since any history of actions is public information shared by all successors, all the information revealed by the history $(x_i)_{i=1}^{n-2}$ is already accumulated in individual $(n-1)$'s cutoff. Therefore, individual n 's cutoff is altered only by the new information revealed by individual $(n-1)$'s action. To be exact, $\hat{\theta}_n$ is different

from $\hat{\theta}_{n-1}$ only by $\mathbb{E}[\theta_{n-1} \mid x_{n-1}, \hat{\theta}_{n-1}]$. As a result, the cutoff rule (2) exhibits the following recursive structure:

$$\hat{\theta}_n = \hat{\theta}_{n-1} - \mathbb{E}[\theta_{n-1} \mid \hat{\theta}_{n-1}, x_{n-1}] \quad (3)$$

with the following updating rule,

$$\mathbb{E}[\theta_{n-1} \mid \hat{\theta}_{n-1}, x_{n-1}] = \begin{cases} \frac{10 + \hat{\theta}_{n-1}}{2} & \text{if } x_{n-1} = A, \\ \frac{-10 + \hat{\theta}_{n-1}}{2} & \text{if } x_{n-1} = B. \end{cases} \quad (4)$$

Substituting (4) in (3) we see that the cutoff dynamics follows the stochastic process

$$\hat{\theta}_n = \begin{cases} \frac{-10 + \hat{\theta}_{n-1}}{2} & \text{if } x_{n-1} = A, \\ \frac{10 + \hat{\theta}_{n-1}}{2} & \text{if } x_{n-1} = B, \end{cases} \quad (5)$$

where $\hat{\theta}_1 = 0$.

The impossibility of an informational cascade follows immediately from (5) since for every n , $-10 < \hat{\theta}_n < 10$. That is, in making a decision, everyone takes her private signal into account in a non-trivial way. Thus, from a theoretical point of view informational cascades are mistakes. Also, the cutoff dynamics (5) captures the distinction between herd behavior and informational cascades. Note that when all choose action A (B) the cutoff process tends to move rapidly toward -10 (10). As a result, since the probability of imitation increases, herd behavior becomes more likely.

Technically, the cutoff process $\{\hat{\theta}_n\}$ has the martingale property, i.e., $\mathbb{E}[\hat{\theta}_n \mid \hat{\theta}_{n-1}] = \hat{\theta}_{n-1}$. So, by the Martingale Convergence Theorem it converges to a random variable $\hat{\theta}_\infty$ almost surely as $n \rightarrow \infty$. In particular, $\hat{\theta}_\infty = -10$ or $\hat{\theta}_\infty = 10$ with probability one – namely, the two fixed points of (5). Further, since convergence of the cutoff process implies convergence of actions by (4), the behavior cannot be overturned forever. Thus, the behavior settles down in some finite time and is consistent with limit learning. In conclusion, we have rendered informational cascades an observable behavioral phenomenon in this model, and they ought not occur. Still, a herd must arise.

5 Experimental Results

Over all sessions, herd behavior of at least five subjects, was observed in 27 of the 75 rounds (36 percent). As Table 1 summarizes, of the 27 herds, 13 (48 percent) involved all eight subjects acting alike. Herding also occurred in 24 of the 37 (64.8 percent) rounds in which it is predicted by Bayes rule⁹. Moreover, all herds, except one turned out to be on the correct action. In contrast, Anderson and Holt (1997) report that about half of the herds turned out to be on the incorrect action. Theoretically, however, the difference between

⁹In other words, if the subjects were Bayesians, playing the game according to (5), given the realization of the signals, we should have observed herding in 37 rounds.

the probabilities of an incorrect herd in this setup and in Anderson and Holt (1997) is negligible¹⁰.

[Table 1 here]

Table 1 shows the cutoffs in rounds in which all eight subjects acted alike. Note that when a subject observed a history in which all previous actions were identical, she typically favored joining the herd for a larger set of private signals by setting her cutoff far from zero in a direction consistent with her predecessors' behavior. Since the cutoff strategy is symmetric around zero, we take a mirror image transformation¹¹ and use the average of the transformed cutoffs to get an idea of the average trend. The result suggests that as subjects observed more identical past actions, they became more confident about the profitability of the herded action. Moreover, we can readily notice that cascade behavior was, in general, increasingly likely to occur towards the last turns of a round. However, an informational cascade was not a necessary condition for herd behavior. For instance, subjects in the last turns of what we have labeled session/round 2/7 (the seventh round in the second session), 4/9, 5/5 and 5/9, participated in herds but still set their cutoffs in the interval $(-10, 10)$.

Perhaps our most unexpected result, at least from the view point of theory, is that informational cascades were observed in 26 rounds (34.7 percent). Of these 26 rounds, in one round the last four subjects, in four rounds the last three subjects, in 11 rounds the last two subjects and in 10 rounds the last subject followed cascade behavior¹². In addition, cascade behavior was observed 32 times outside of informational cascades¹³. Table 2 shows the rounds in which the longest informational cascades occurred.

[Table 2 here]

Finally, only 7.2 percent of all decisions, excluding the first decision turn in each round, were inconsistent with the observed history in the sense that the sign of the cutoff actually employed is opposite to what it should be according to (5) given the history of past actions. For example, in round 2/13 (see Table 2), the second subject favored action A even though the action she observed

¹⁰In Anderson and Holt (1997)'s setup with signal precision $2/3$, simple calculations yield that the probabilities of a correct herd, or an incorrect herd, are 70.6 percent and 28.3 percent respectively. In this setup, it cannot be found analytically since, conditional on the true state of the world, private signals are negatively correlated. However, with the help of simulations, we find that the probabilities of a correct herd, or an incorrect herd of at least five individuals are 75.7 percent and 23.7 percent respectively.

¹¹We call decisions made by subjects concurring if the sign of their cutoff is as they should be according to theory. Otherwise, we say that the decisions are contrary. We transform the data by taking the absolute value of cutoffs in concurring decision points and negative of the absolute value of cutoffs at contrary decision points.

¹²Of all 40 subjects only two followed cascade behavior in all 15 rounds in which they participated.

¹³We identify a subject who engages in cascade behavior which was not a part of any cascade, as one who reports a cutoff -10 (10) and there exist a subsequent subject who reports a cutoff in the interval $(-10, 10]$ ($[-10, 10)$).

was B . Similarly, the third subject favored action B , where past actions should have led her to favor action A .

6 An Econometric Analysis

The frequency of herd behavior, and the fact all herds, except one, turned out to be correct, suggests that subjects process the information revealed by others' actions, and attempt a best response based on this information. In so doing subjects must presumably estimate the errors of others and consider this in processing the information revealed by their predecessors' actions. Like Anderson and Holt (1997), we attempt to formulate this by estimating a recursive model that allows for the possibility of errors in earlier decisions. This approach enables us to evaluate the degree to which Bayes rationality explains behavior in the laboratory. However, while Anderson and Holt (1997) use subjects' expected payoffs, our cutoff elicitation allows us to estimate recursively the process of cutoff determination adjusted for decision errors and independent shocks.

To incorporate the possibility that individuals can make mistakes, we modify the original model, relaxing the assumption of rationality. To be precise, we assume that at each decision turn n , with probability p_n an individual is Bayesian and rationally computes her cutoff, and with probability $1 - p_n$, she is noisy, in the sense that her cutoff is a random draw from a distribution function G_n with support $[-10, 10]$ and mean $\tilde{\theta}_n$. Suppose that others cannot observe whether an individual's behavior is noisy, but the sequences $\{p_n\}$ and $\{G_n\}$ are common knowledge among individuals.

At any turn $n > 1$, a rational individual makes her decision based on the information revealed by the history $(x_i)_{i=1}^{n-1}$ taking the noise, $(p_i)_{i=1}^{n-1}$ and $(G_i)_{i=1}^{n-1}$, into consideration. Since all the information revealed by the history $(x_i)_{i=1}^{n-2}$ is already accumulated in the cutoff of the rational individual at turn $(n-1)$, the cutoff rule of rational individuals exhibits the following recursive structure:

$$\hat{\theta}_n = \hat{\theta}_{n-1} - p_{n-1} \mathbb{E}[\theta_{n-1} \mid \hat{\theta}_{n-1}, x_{n-1}] - (1 - p_{n-1}) \mathbb{E}[\theta_{n-1} \mid G_{n-1}, x_{n-1}]$$

where

$$\mathbb{E}[\theta_{n-1} \mid G_{n-1}, x_{n-1} = A] = \int_{-10}^{10} \frac{10+x}{2} dG(x) = \frac{10 + \tilde{\theta}_n}{2},$$

and

$$\mathbb{E}[\theta_{n-1} \mid G_{n-1}, x_{n-1} = B] = \int_{-10}^{10} \frac{-10+x}{2} dG(x) = \frac{-10 + \tilde{\theta}_n}{2}.$$

Hence, the cutoff dynamics of rational individuals follow the process

$$\hat{\theta}_n = \hat{\theta}_{n-1} - \begin{cases} \frac{10+(1-p_{n-1})\tilde{\theta}_{n-1}+p_{n-1}\hat{\theta}_{n-1}}{2} & \text{if } x_{n-1} = A, \\ \frac{-10+(1-p_{n-1})\tilde{\theta}_{n-1}+p_{n-1}\hat{\theta}_{n-1}}{2} & \text{if } x_{n-1} = B, \end{cases} \quad (6)$$

where $\hat{\theta}_1 = 0$.

In addition, we assume that rational individuals could tremble, in the sense that their cutoff embodies uncorrelated small computation or reporting mistakes. To be precise, a rational individual in turn n reports cutoff $\hat{\theta}_n + \phi_n$ where ϕ_n is distributed normally with mean 0 and variance σ_n^2 . It is important to note, however, that the mistakes of the rational individuals differ from the behavior of noisy individuals since the former is a tremble from the rational cutoff, i.e., has mean $\hat{\theta}_n$, whereas the later is simply random behavior.

Under these assumptions, at any decision turn n and round i , the expected cutoff is

$$y_n^i = (1 - p_n)\tilde{\theta}_n + p_n\hat{\theta}_n^i + p_n\phi_n^i$$

and in matrix form

$$\mathbf{y}_n = (1 - p_n)\tilde{\theta}_n\mathbf{1} + p_n\hat{\boldsymbol{\theta}}_n + p_n\boldsymbol{\phi}_n \quad (7)$$

where $\mathbf{y}_n, \mathbf{1}, \hat{\boldsymbol{\theta}}_n$, and $\boldsymbol{\phi}_n$ are vectors whose components are $y_n^i, 1, \hat{\theta}_n^i$ and ϕ_n^i respectively.

This leads the following econometric specification:

$$\mathbf{y}_n = \alpha_n\mathbf{1} + \beta_n\mathbf{z}_n + \boldsymbol{\varepsilon}_n \quad (8)$$

where

$$\alpha_n = (1 - p_n)\tilde{\theta}_{n-1}, \beta_n = p_n, \boldsymbol{\varepsilon}_n = p_n\boldsymbol{\phi}_n$$

and for any round i , $\mathbf{z}_1 = \mathbf{0}$ and for any $n > 1$

$$z_n^i = z_{n-1}^i - \begin{cases} \frac{10 + (\hat{\alpha}_{n-1} + \hat{\beta}_{n-1}z_{n-1}^i)}{2} & \text{if } x_{n-1}^i = A, \\ \frac{-10 + (\hat{\alpha}_{n-1} + \hat{\beta}_{n-1}z_{n-1}^i)}{2} & \text{if } x_{n-1}^i = B, \end{cases} \quad (9)$$

is the i^{th} component of vector \mathbf{z}_n ¹⁴.

Notice that the parameters are estimated recursively. That is, the estimated parameters for the first decision-turn, $\hat{\alpha}_1$ and $\hat{\beta}_1$, are employed in estimating the parameters for the second turn, α_2 and β_2 , and so on. So, at each turn n , the estimates for the previous turn $\hat{\alpha}_{n-1}$ and $\hat{\beta}_{n-1}$ are used to calculate an estimate of the optimal cutoff for each decision $\hat{\theta}_n^i$, denoted by z_n^i , according to (9), which, in turn, constitutes the independent variable in the estimation (8) for that turn. This is the sense in which the updating rule given by (9) is optimally adjusted to previous decision errors in a Bayesian way. The error-adjusted updating rule (9) suggests that rational individuals estimate average errors in earlier decisions and take them into account in making their decisions. This is the behavioral interpretation of the recursive econometric method.

As to the interpretation of the parameters, coefficient β is the probability that a subject participating in decision-turn n is rational, which can be interpreted as a parameterization of the average weights given to the information

¹⁴Note that a cutoff, which is based on the observed histories of decisions adjusted to previous decision errors, may escape the support of private signals, i.e. go outside the interval $[-10, 10]$. In such cases, we set the cutoff at the corresponding boundary. That is, whenever $z_n^i < -10$ ($z_n^i > 10$) we set it to -10 (10).

revealed by the history of actions. On the other hand, coefficient α can be interpreted as a parameterization of the information processing bias such as a blind tendency toward a particular action. For example, since $\hat{\theta}_n = \alpha_n / (1 - \beta_n)$, when $\beta_n < 1$, any $\alpha_n < 0$ ($\alpha_n > 0$) indicates that subjects participating in turn n are biased toward action A (B). When the information processing biases diminish, i.e., $\alpha_n \rightarrow 0$, and $\beta_n \rightarrow 1$ (and $\sigma_n^2 \rightarrow 0$), the behavior tends to become Bayesian. That is, when $\alpha_n = 0$ and $\beta_n = 1$ for all n , according to (8), the laboratory decision-making conforms perfectly with the optimal history-contingent cutoff process given by (5). Similarly, the behavior tends to be random as $\alpha_n \rightarrow 0$ and $\beta_n \rightarrow 0$. Notice that when $\alpha_n = \beta_n = 0$ (and $\sigma_n^2 \rightarrow 0$), equation (8) requires expected cutoff to be zero, which is simply a choice based on private information. In general, any $\beta_n < 1$ indicates that the population of subjects in turn n underweights the information revealed by the history of others' actions relative to their private information. This is a plausible response to the belief that others can make errors in their decisions. In a simulation that was carried out taking $\hat{\mathbf{y}}_n$ as the benchmark cutoff and the corresponding realized signals in the experiment, our estimation accurately predicts 467 (77.8 percent) of the 600 actions in the experiment. Table 3 reports the results¹⁵.

[Table 3 here]

As Table 3 reports, we fail to reject the hypothesis that $\hat{\alpha}_n = 0$ in all turns, which is strong support for the inference that noisy subjects do not have any systematic bias towards a particular action, A or B , and the $\hat{\beta}_n$ coefficients are bounded away from zero and one. Note that on average a subject acting as a second decision-maker tends to undervalue sharply the first subjects' decision, $\hat{\beta}_2 = 0.22$. Thus, our econometric results strongly suggest that subjects who take decisions early have a substantial tendency to determine their cutoffs randomly instead of employing Bayesian reasoning. Also noteworthy is the obvious upward trend in the $\hat{\beta}_n$ coefficients, which indicates that over time subjects tend to approach to Bayesian updating more closely. Put differently, along the line of subjects the information revealed by the history of actions is relied upon more and subjects become increasingly likely to imitate their predecessors.

Note that when noisy individuals ignore history and make decisions solely on the basis of private information, by simply setting cutoffs at zero, put side by side with a rational individual, a noisy individual reveals more of her private information. Thus, our empirical results indicate that, in Bayesian terms, subjects weigh their own information too heavily and give too little weight to public information. Furthermore, subjects who are early decision-makers tend to rely more heavily on their own information in the learning process, which then becomes available to late decision-makers, who tend more to be Bayesian.

This finding explains why informational cascades arise in the laboratory. This is because someone who overweights private information, reveals more in-

¹⁵The GLS random-effects (mixed) estimators and robust variance estimators for independent data and clustered data (data not independent within subjects but independent across subjects) yield similar results.

formation about the private signal, and, as a result, may stimulate an informational cascade. To illustrate this, note that since $\hat{\alpha}_1 = \hat{\alpha}_2 = \hat{\alpha}_3 = 0$, $\hat{\beta}_2 = 0.22$ and $\hat{\beta}_3 = 0.48$ whenever the first three individuals take the same action, say A, A, A , a simple calculation using (6) shows that the cutoff of a subsequent rational individual escapes the support of private signals, $\hat{\theta}_4 < -10$, and thus she chooses A no matter what her private signal is.

Next, we turn our attention to the following question: How well do our econometric results predict cascade behavior?

To answer this question, for each decision turn, we first find the number of rounds for which the estimation predicts that a rational individual would engage in cascade behavior, i.e., according to (9), either $z_n^i = 10$ or $z_n^i = -10$. Then, we multiply this number by the estimated probability that an individual is rational, $\hat{\beta}_n$, to obtain the expected number of occurrences of cascade behavior. As a goodness-of-fit measure, the histograms in Figure 2 compare, turn by turn, the proportion of rounds in which cascade behavior was observed in the laboratory (Black) to the proportion of rounds it was predicted by the model (Gray). Figure 2 shows that in decision turns 4-8 the estimation adequately predicts cascade behavior in the laboratory: overall there is no significant difference from what the estimation predicts. However, in turns 1-3, even though the estimation predicts the impossibility of rational cascade behavior, it still arises in the laboratory. We attribute the cascade behavior in early decision turns to noisy individuals who are more populated in these turns.

[Figure 2 here]

So far, we have focused on the frequency of cascade behavior as a goodness-of-fit measure between the observed behavior and the one prescribed by the estimation. This tells only part of the story as it ignores how well the estimation fits the data decision point by decision point. Note that we can organize the data in terms of cascade behavior into four possible cases: the estimation predicts that a rational individual would engage in cascade behavior and cascade behavior is observed or not observed (cases I and II respectively), and the estimation predicts that a rational individual would not engage in cascade behavior and cascade behavior is observed or not observed (cases III and IV respectively). Figure 3 summarizes these cases.

[Figure 3 here]

In case I, our prediction attributes the observed cascade behavior to rational subjects. By contrast, in case II, even though the prediction of our estimation suggests that rational subjects would engage in cascade behavior, we do not observe such behavior. This observation leads us to conclude that the decisions in case II originate from noisy subjects. Similarly, we can identify the decisions in case III as noisy since the estimation predicts that a rational individual would not engage in cascade behavior and cascade behavior is observed. A notable difference between cases I and II is that, for case II these decisions might be

attributable to a preference for conformity¹⁶. This type of cascade behavior was observed many years ago in social psychology experiments. This literature suggests several alternative explanations to the rational view of cascade behavior. In particular, individuals inherently wish to conform with the behavior of others because this inclination to conform is a natural property of individual preferences, i.e., preferences for conformity for its own sake¹⁷. Lastly, in case IV, the estimation remains silent about the composition of rational and noisy decisions. This is because the prediction that a rational subject would engage in cascade behavior does not necessarily imply that the observed behavior comes from a rational subject.

In order to put the observed behavior into perspective, for each turn n let C_n^k denote the number of occurrences in case k and consider the ratios

$$r_n^1 := \frac{C_n^I}{C_n^I + C_n^{III}} \text{ and } r_n^2 := \frac{C_n^{II}}{C_n^{II} + C_n^{IV}}.$$

Hence, for decision points where cascade behavior was observed (not observed) in the data, r_n^1 (r_n^2) is the fraction that the model predicts that a rational subject would engage in cascade behavior. Table 4 summarizes the percentages of each case and the ratios, r_n^1 and r_n^2 by turn. Note that r_n^1 is higher than r_n^2 in late decision-turns, indicating that the estimation is not only predicting the right frequencies of rational and noisy cascade behavior, but also predicting it in the right occurrences. Furthermore, r_n^1 also provides an upper bound on the fraction of cascade occurrences in turn n which, according to the estimation, might have resulted from rational behavior.

[Table 4 here]

Finally, it will be illustrative to compare the predictions of the modified model with the Bayesian model. For this purpose Figure 4 compares the theoretical Bayes' cutoff process and the estimated error-adjusted expected process $\hat{\mathbf{y}}_{\mathbf{n}}$ for two histories. When all choose action A (sequence one), the estimated expected cutoffs are far above their theoretical counterparts in early decisions, meaning a relative predisposition of subjects to follow their private information. However, over time, the gap between the theoretical and estimated cutoffs diminishes, which suggests that the Bayesian solution, as given by cutoff strategy (5), adequately predicts the behavior of a large portion of subjects in the laboratory¹⁸. One may argue that this result is not robust since a history in

¹⁶Hung and Plott (2001) manipulate the payoff structure of Anderson and Holt (1997) to investigate further possible explanations for cascade behavior in binary-signal-binary-action setup. They reject preference for conformity and non-equilibrium Bayesian behavior as explanations, in favor of Bayesian equilibrium behavior.

¹⁷For early literature, see Asch (1958). Anderson, et al. (1997) review this psychology literature and provide additional references. In the economics literature, Bernheim (1994) formulates some relevant concepts.

¹⁸Note that when all choose action A , \hat{z}_n , which is the estimate of the rational cutoff $\hat{\theta}_n$, equals -10 at late decision turns but $\hat{\mathbf{y}}_{\mathbf{n}} = \hat{\alpha}_n + \hat{\beta}_n \hat{z}_n$ does not escape the support of private signals.

which all subjects act alike is a very special case. However, Bayes rationality also performs well as a predictor of the behavior observed in the laboratory in case of an overturn. When the first two subjects take action A and all subsequent subjects take action B (sequence two), the estimated cutoffs show that an average successor follows the deviation, as the theoretical findings indicate, although she values the newly revealed information a little less than the theory predicts. This is an obvious reaction to the possibility that her predecessor's deviation is erroneous.

[Figure 4 here]

To conclude, the subjects' behavior can best be characterized as a mixture of bounded rationality and rationality. Taken as a whole, the estimated cutoffs are properly modified to take these traits into account and this permits successful prediction of the subjects' behavior in the laboratory.

7 Concluding Remarks

Social learning models are easily adapted to an experimental setting and this has provided a valuable opportunity to test theoretical predictions. In addition to testing the theory, by using a novel setup, this paper discovers behavior patterns about which the existing theory has little to say. This paper offers two contributions to method: First, our experiment shows how a continuous-signal social learning model can be tested, theoretically yielding a behavior richer than the one of the simple binary-signal models tested by Anderson and Holt (1997). Second, it shows the use of a cutoff elicitation technique to elicit subjects' beliefs. This enables us to distinguish herd behavior experimentally from the important behavioral phenomenon, the informational cascades, and leads us to examine how well Bayes rationality approximates the actual behavior observed in the laboratory.

Our results are summarized as follows. First, we find that herd behavior develops frequently (36 percent) in the laboratory and that all herds, except one, turned out to be correct. This is particularly interesting since a prediction of the theory, which was matched in many experiments, is that mass behavior is likely to be erroneous. Moreover, as apposed to the impossibility of informational cascades prediction of the theory, we find that cascades often arise (34.7 percent). Thus, we conclude that although cascades are not a theoretical possibility, they are a reality. Second, we find that in the laboratory subjects give excessive weight to their private information relative to the public information revealed by the behavior of others, but, over time, they tend towards Bayesian updating. We have used this result to help explain why the cascade behavior observed in the laboratory may, after all, be rational.

The message of the paper is, therefore, that its novel setup, along with its elicitation method, enriches the social learning paradigm and, at the same time, provide an effective explanations of mass behavior. Special interest is merited by the sequential decision problem that permits patterns of mass behavior to

be fragile and easily overturned after a deviation, a type of episodic instability that is characteristic of social behavior in the real world. In this way, we find the experimental results helpful both for understanding and improvement of the theory of social learning. The experimental techniques and results that we have developed provide some tools promising for future work in this area, which is certainly needed for fuller understanding of the economic impact of social learning.

References

- [1] Anderson, L. and C. Holt (1996) "Classroom Games: Information Cascades." *Journal of Economic Perspectives*, 10(4), pp. 187-93.
- [2] Anderson, L. and C. Holt (1997) "Information Cascades in the Laboratory." *American Economic Review*, 87(5), pp. 847-62.
- [3] Aronson, Eliot, T. Wilson and R. Akert. *Social psychology*. New York: Addison Wesley Longman, 1997.
- [4] Asch, S. E. "Effects of Group Pressure upon the Modification and Distortion of Judgements." in Eleanor E. Maccoby, Theodore M. Newcomb, and E. L. Hartley, eds., *Readings in social psychology*. New York: Holt, Rinehart & Winston, 1958, pp. 174-83.
- [5] Avery, C. and P. Zemsky (1998) "Multidimensional Uncertainty and Herd Behavior in Financial Markets." *American Economic Review*, 88(4), pp. 724-48.
- [6] Banerjee, A. (1992) "A Simple Model of Herd Behavior." *Quarterly Journal of Economics*, 107(3), pp. 797-817.
- [7] Bernheim, D.(1994) "A Theory of Conformity." *Journal of Political Economy*, 102(5), pp. 841-877.
- [8] Bikhchandani, S., D. Hirshleifer and I. Welch (1992) "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascade." *Journal of Political Economy*, 100(5), pp. 992-1026.
- [9] Bikhchandani, S., D. Hirshleifer and I. Welch (1998) "Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades." *Journal of Economic Perspective*, 12(3), pp. 151-70.
- [10] Çelen, B. and S. Kariv (2003) "Observational Learning under Imperfect Information." *Games and Economic Behavior*, forthcoming.
- [11] Gale, D. (1996) "What Have We Learned from Social Learning?" *European Economic Review*, 40(3-5), pp. 617-28.

- [12] Hung, A. and C. Plott (2001) "Information Cascades: Replication and an Extension to Majority Rule and Conformity-Rewarding Institutions." *American Economic Review*, 91(5), pp. 1508-20.
- [13] Kennedy, R. (2002) "Strategy Fads and Competitive Convergence: An Empirical Test for Herd Behavior in Prime-time Television Programming." *Journal of Industrial Economics*, 50(1), pp. 57-84.
- [14] Kübler, D. and G. Weizsäcker (2003) "Limited Depth of Reasoning and Failure of Cascade Formation in the Laboratory." *Review of Economic Studies*, forthcoming.
- [15] Morton, R. and K. Williams (1999) "Information Asymmetries and Simultaneous versus Sequential Voting." *American Political Science Review*, 93(1), pp. 51-68.
- [16] Neeman, Z. and G.. Orosel (1999) "Herding and the Winner's Curse in Markets with Sequential Bids." *Journal of Economic Theory*, 85(1), pp. 91-121.
- [17] Scharfstein, D. and J.. Stein (1990) "Herd Behavior and Investment." *American Economic Review*, 80(3), pp. 465-79.
- [18] Smith, L. and P. Sørensen (2000) "Pathological Outcomes of Observational Learning." *Econometrica*, 68(2), pp. 371-398.
- [19] Welch, I. (1992) "Sequential Sales, Learning and Cascades." *Journal of Finance*, 47(2), pp. 695-732.

Table 1: Data for rounds in which all eight subjects acted alike

Session/ Round*	Action herded	Cutoff by turn							
		1	2	3	4	5	6	7	8
1/11	<i>B</i>	0	4.25	10	5	10	10	10	10
2/7	<i>A</i>	4	1	-2	1	-6.1	-10	0	-9.4
2/1	<i>B</i>	2	3	9	10	10	9.8	10	10
2/11	<i>A</i>	-6.1	-2	-6	-10	-8.8	0	-10	-10
2/14	<i>A</i>	0	-7	-10	-10	-4	-9.9	-10	-10
3/1	<i>A</i>	0	1.5	-0.01	-2	-10	0	0	-10
4/3	<i>A</i>	-5	5	0	-4	-10	-9.87	-10	-10
4/9	<i>B</i>	5.4	10	8.69	6.4	3	9	10	8.5
4/12	<i>A</i>	0	0	-9	-10	-9.13	-5	-10	-10
4/15	<i>B</i>	10	9.99	5	0	9.9	10	10	10
5/3	<i>B</i>	0	4	-2	10	2	10	0	10
5/5	<i>A</i>	-1	0	-10	-8	-5	-10	5	0
5/9	<i>A</i>	0	0	-10	-10	-3	-10	-6	-9
Average**		2.0	2.5	6.3	6.5	7.0	8.0	6.2	9.0

* - (Session/Round). For example, 1/11 is the eleventh round in the first session.

** - Average of the cutoffs absolute values.

Table 2: Data for the rounds with the longest informational cascades

Subject number : Action

Cutoff

Private signal

Session/ Round	1	2	3	4	5	6	7	8	Sum of Signals
1/9	8: <i>B</i> 0 -5.41	6: <i>B</i> 0 -7.12	4: <i>B</i> 5 -3.72	2: <i>B</i> 5 -1.89	1: <i>A</i> 0 3.59	7: <i>B</i> 10 5.59	5: <i>B</i> 10 1.76	3: <i>B</i> 10 -7.95	-15.2
1/1	8: <i>B</i> 0 -0.35	6: <i>B</i> 0 -4.71	1: <i>A</i> * -10 -0.34	3: <i>A</i> 8.6 9.17	4: <i>A</i> 0 0.63	2: <i>A</i> -10 8.69	5: <i>A</i> -10 -7.61	7: <i>A</i> -10 -4.00	1.5
1/11	4: <i>B</i> 0 -1.44	6: <i>B</i> 4.25 -2.71	8: <i>B</i> 10 0.74	2: <i>B</i> 5 -4.76	5: <i>B</i> 10 1.87	7: <i>B</i> 10 -7.94	1: <i>B</i> 10 4.80	3: <i>B</i> 10 -1.06	-10.5
2/13	2: <i>B</i> 0 -1.50	1: <i>A</i> * -1 4.11	5: <i>A</i> * 1.5 4.11	7: <i>A</i> -5.9 1.35	3: <i>A</i> -6.6 6.42	6: <i>A</i> -10 -5.71	4: <i>A</i> -10 6.04	8: <i>A</i> -10 8.71	23.5
4/15	2: <i>B</i> 10 -6.45	3: <i>B</i> 9.99 -4.56	4: <i>B</i> 5 -6.57	1: <i>B</i> 0 -1.82	6: <i>B</i> 9.9 6.05	5: <i>B</i> 10 0.8	7: <i>B</i> 10 -8.19	8: <i>B</i> 10 1.29	-19.5

 - Cascade behavior.

* - Cutoffs inconsistent with the observed history.

Table 3: The econometric results by turn

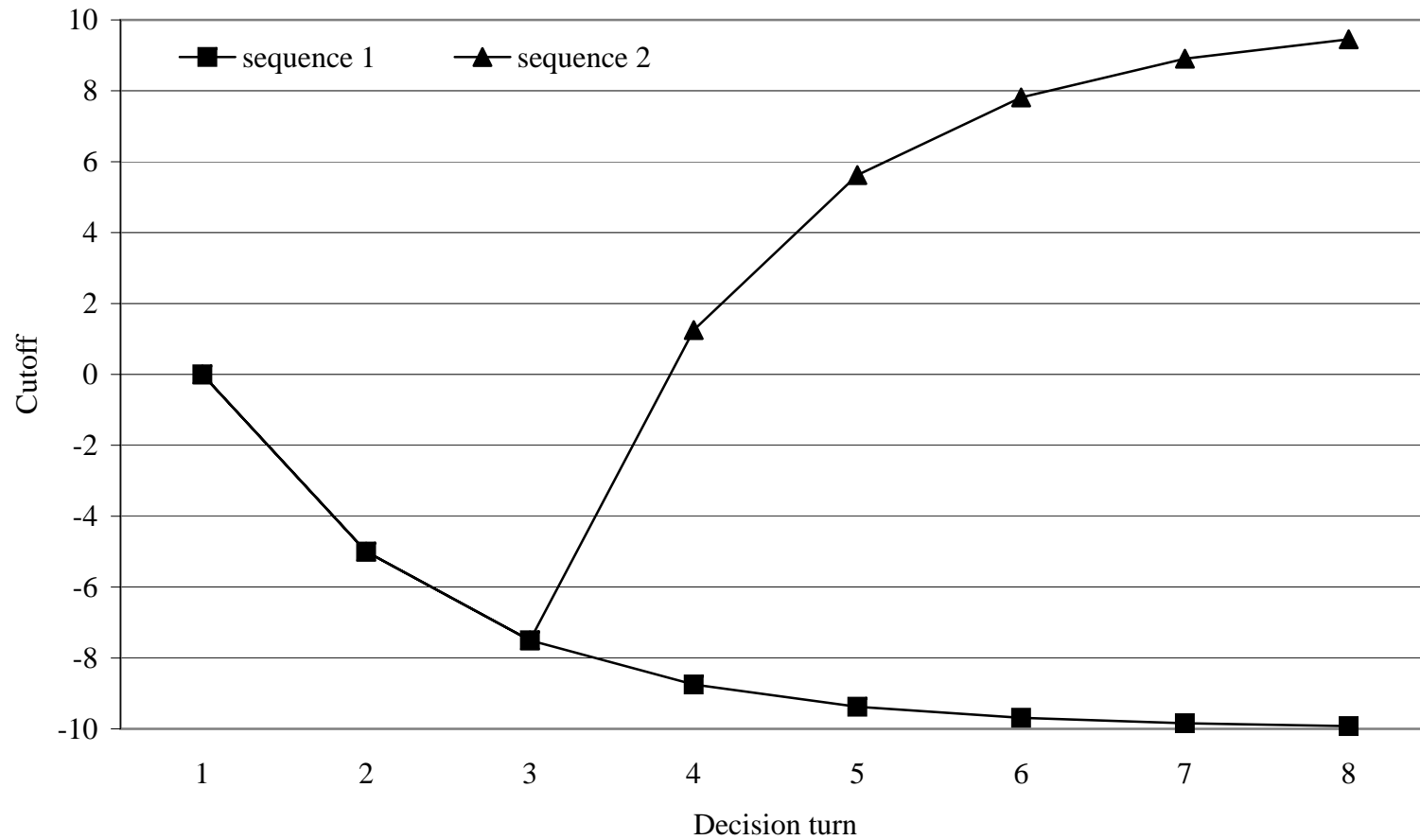
Turn	1	2	3	4	5	6	7	8
# of obs.	75	75	75	75	75	75	75	75
$\hat{\alpha}$	-0.41 (0.53)	0.96 (0.46)	0.02 (0.56)	0.16 (0.56)	-0.02 (0.48)	0.39 (0.59)	-0.05 (0.63)	0.27 (0.67)
$\hat{\beta}$	--	0.22 (0.09)	0.48 (0.07)	0.49 (0.07)	0.59 (0.06)	0.60 (0.07)	0.59 (0.08)	0.62 (0.08)
R^2		0.07	0.31	0.39	0.51	0.47	0.45	0.45

(standard errors)

Table 4: The percentages of each case of cascade behavior by turn

Case	Decision turn							
	1	2	3	4	5	6	7	8
I	0	0	0	16.0	10.7	17.3	22.7	32.0
II	0	0	0	29.3	33.3	29.3	30.7	25.3
III	14.7	5.3	18.7	2.7	6.7	10.7	14.7	6.7
IV	85.3	94.7	81.3	52.0	49.3	42.7	32.0	36.0
r_n^1	0	0	0	0.86	0.62	0.62	0.61	0.83
r_n^2	0	0	0	0.36	0.40	0.41	0.49	0.41

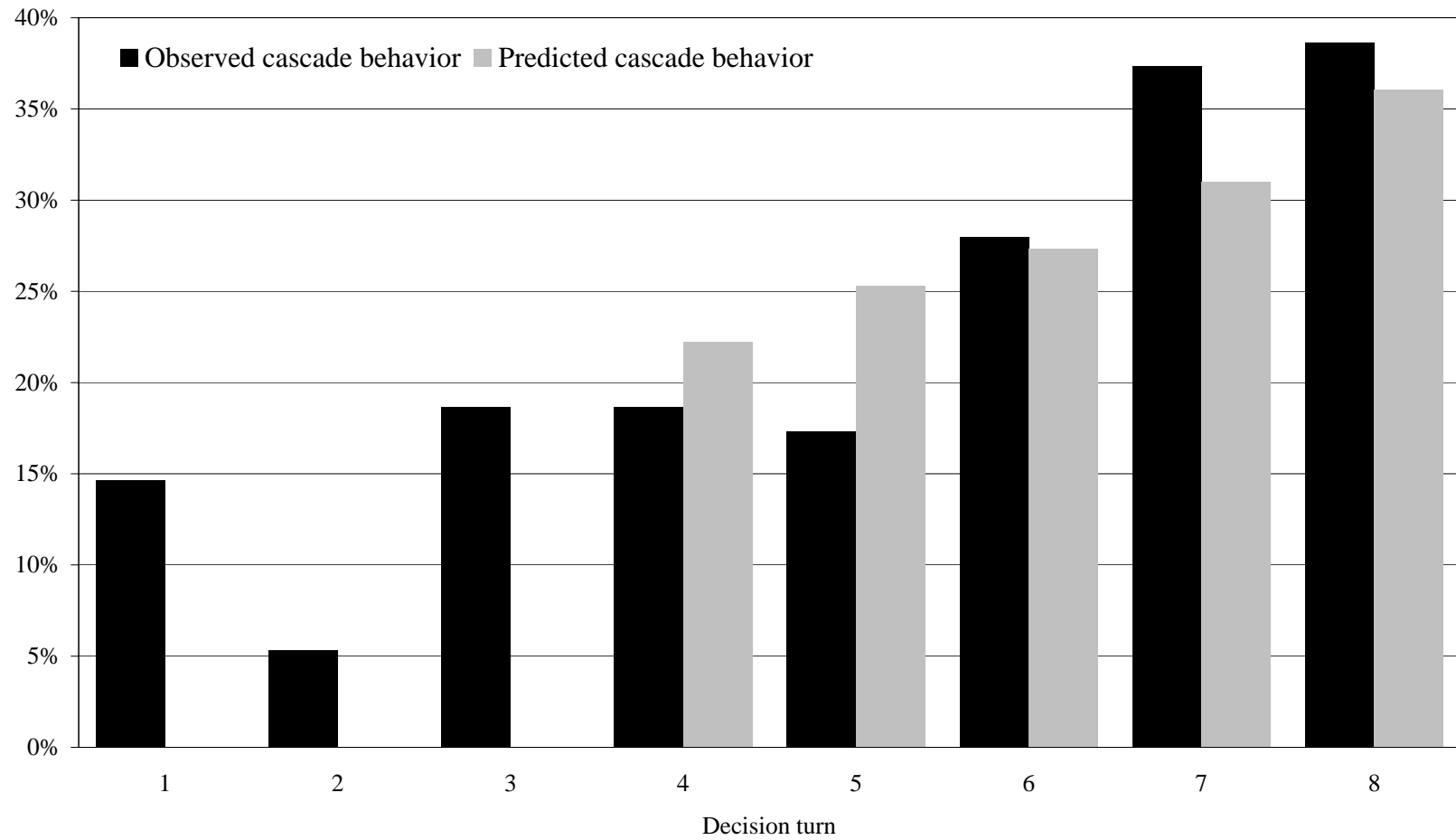
Figure 1: Sequences of cutoffs for two histories



Sequence 1 - all individuals choose action A .

Sequence 2 - the first two individuals take action A and the third individual deviates.

Figure 2: Cascade behavior

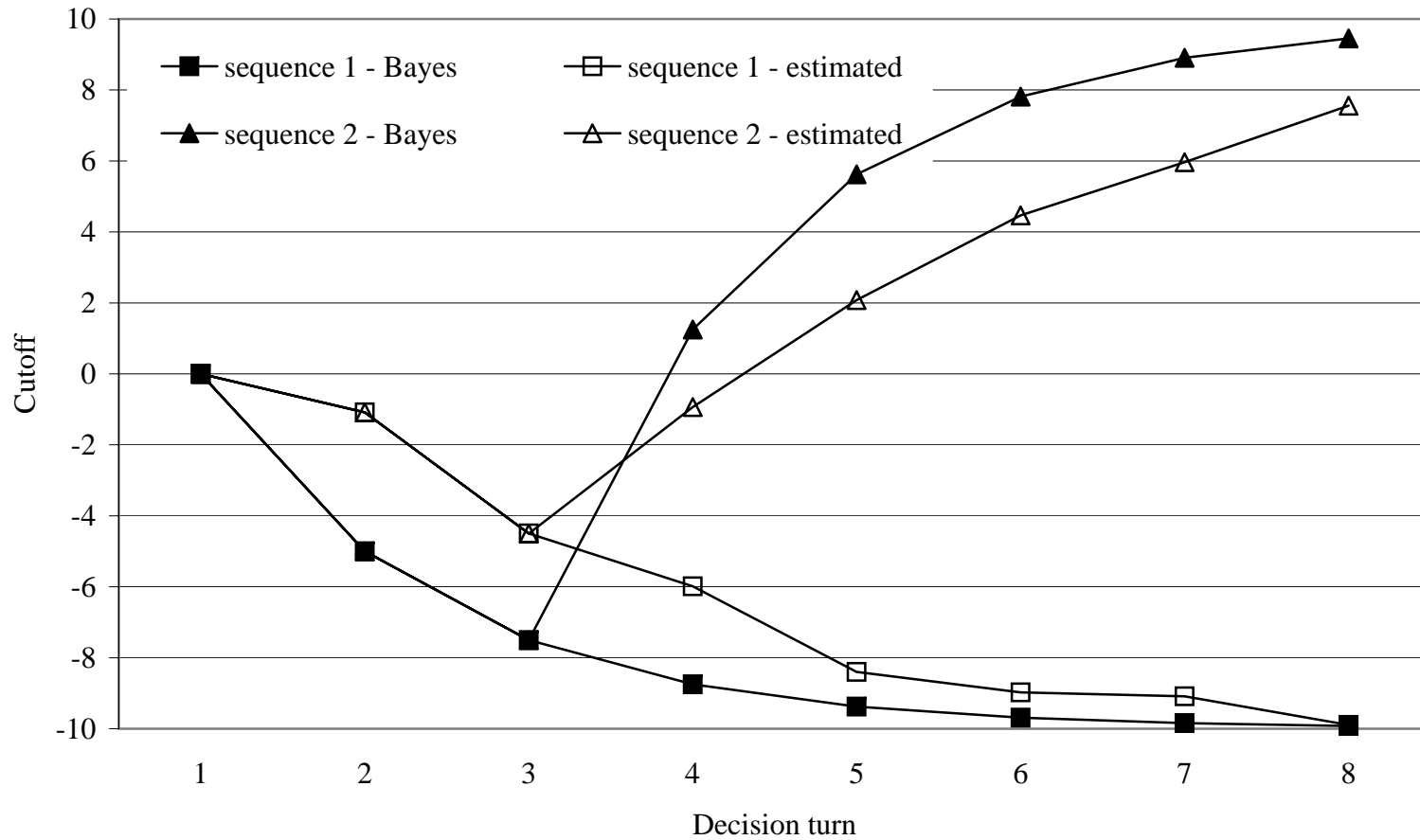


The proportion of rounds in which cascade behavior was observed in the laboratory (Black) and the proportion of rounds it was predicted by the model (Gray)

Figure 3: Cases of cascade behavior - observed and predicted

		Observed	
		Cascade	No cascade
Predicted	Cascade	I	II
	No cascade	III	IV

Figure 4: Sequences of cutoffs for two histories - theory and estimated



Sequence 1 - all individuals choose action A .

Sequence 2 - the first two individuals take action A and the third individual deviates.