Distinguishing Informational Cascades from Herd Behavior in the Laboratory[1]

Boğaçhan Çelen
Department of Economics, New York University
269 Mercer St., 7th Floor, New York, NY, 10003-6687.
E-mail: bc319@nyu.edu
Url: http://home.nyu.edu/~bc319

and

Shachar Kariv
Department of Economics, New York University
269 Mercer St., 7th Floor, New York, NY, 10003-6687.
E-mail: sk510@nyu.edu
Url: http://home.nyu.edu/~sk510

Version: October 10, 2001

This paper reports an experimental test of how individuals learn by observing the behavior of others. By using techniques only available in the laboratory, we elicit subjects' beliefs. This elicitation allows us to distinguish informational cascades from herd behavior. By using a continuous-signal-discrete-action setup, along with our elicitation method, we are able to enrich the ball-and-urn observational learning experiments paradigm of Anderson and Holt (1997). To interpret our results, we test a model that explains subjects behavior as a form of modified Bayesian behavior that incorporates limits on the rationality of others. We find strong evidence that this type of Bayes rationality adequately explains the behavior in the laboratory.

Journal of Economic Literature Codes: C92, D8.

Key Words: Asymmetric Information, Herd Behavior, Informational Cascade.

[1] This research was supported by the C. V. Starr Center for Applied Economics at New York University. We are grateful to Andrew Schotter for his guidance. We also acknowledge helpful discussions of Liran Einav, Xavier Gabaix, Douglas Gale, Charles Holt, David Laibson, Wieland Müller, and Matt Rabin. This paper has also benefited from suggestions by the participants of The Inter-university Conference at Columbia and seminars at Cornell, NYU and Yale. In addition, we would like to thank Anne Stubing for her valuable assistance, and Roi Zemmer for writing the experimental computer program.

1. INTRODUCTION

In recent years there has been a great deal of interest paid to *observational learning*, which occurs in any situation in which decision-makers (*dm*s) learn by observing the behavior of others. Two phenomena that have drawn particular interests are that of *informational cascades* and *herd behavior*. The theoretical works of Banerjee (1992) and Bikhchandani, Hirshleifer and Welch (1992) (BHW), demonstrate that such phenomena can arise in a wide variety of empirically relevant economic contexts[2] and that, despite the available information, observational learning may lead to an inefficient uniform social behavior. As such, informational cascades and herd behavior give insight into phenomena such as manias, fads, fashions, crashes and booms. Following the leading work by Anderson and Holt (1997) (AH), a number of experimental papers[3] analyzed several aspects of observational learning.

While the terms informational cascade and herd behavior are used interchangeably in the literature, there is a significant difference between them. From a theoretical point of view, when a sequence of *dm*s make identical once-in-a-lifetime decisions based on private information about an underlying payoff-relevant event as well as all predecessors' decisions, it is said that

> an informational cascade occurs when an infinite cluster of *dm*s ignore their private information when making a decision, whereas
>
> herd behavior occurs when an infinite cluster of *dm*s make an identical decision, not necessarily ignoring their private information.

In other words, a herd occurs when after some point in time, all *dm*s make identical decisions, while an informational cascade occurs when, after some point in time, all *dm*s happen to make the same decision and would do so no matter what private signals they receive. Thus, an informational cascade implies a herd but a herd is not necessarily the result of an informational cascade. In a herd every *dm* chooses the same action, but they might have chosen a different sequence of actions if the realization of their private signals had been different. In an informational cascade, there is no private signal that could lead *dm*s to do anything else but follow the herd since their beliefs are so strongly held that no signal can outweigh them. Put differently, we can regard an informational cascade as convergence of beliefs and herd behavior as convergence of actions. From this distinction it is clear that informational cascades, which are an unobservable phenomenon defined with respect to beliefs, is much harder to verify

[2] For applications in business strategy, consumer marketing, politics and more, see Bikhchandani, Hirshleifer and Welch (1998).

[3] See, e.g. Allsopp and Hey (1999), Anderson (1999) and Huck and Oechssler (2000).

than is a herd, which is an observable phenomenon defined with respect to actions.

This paper reports an experimental test based on an earlier theoretical paper by Çelen and Kariv (2001) (ÇKa)[4]. We employ a design in which a finite sequence of *dm*s draw private signals from a bounded and continuous support which is symmetric around zero. The decision problem is to predict whether the sum of all *dm*s' 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 otherwise. However, instead of choosing action A or B directly, after being informed about the history of actions of the subjects before them and before observing their private signals, subjects are asked to state a cutoff such that action A would be chosen if the signal received is greater than the cutoff and action B if it is less. Only after a subject submits her cutoff, is she informed of her private signal, and her action is recorded accordingly. Eliciting cutoffs enables us to identify a subject who exhibits a *cascade behavior*, i.e. acts irrespective of her private signal, as someone who chooses the same action as her predecessor for any realization of her private signal. In contrast, a subject who joins a herd is one who, given the realization of her private signal, acts as her predecessors did, but might have acted differently if the realization of her signal had been different.

In the laboratory, we find that herd behavior develops frequently as do cascades. However, not all observed herds are cascades. Since our theoretical model predicts that an informational cascade is impossible, we explain why cascades often arise in the laboratory as a specific type of deviation from Bayes rationality. More precisely, from the cutoffs that we elicit, we perceive the extent to which subjects are making errors, and how they incorporate the possibility that others are making errors into their beliefs, i.e. incorporating limits on the rationality of others. Our results suggest that Bayesian models properly modified to account for human error successfully predict subjects behavior in the laboratory.

The reminder of this paper is organized as follows. In the next section, we review the related theoretical and experimental observational learning literature. In section 3, we describe the experimental design and procedures, and in section 4 we outline the underlying decision problem. We discuss our results in section 5, and provide an econometric analysis in section 6. We conclude in the closing section.

[4] In this paper, which is built on Gale (1996), we focus on observational learning under imperfect information where each *dm* learns only from her immediate predecessor's decision. The conventional perfect information setup is analyzed as a benchmark. We provide an experimental test of observational learning under imperfect information in a separate paper, Çelen and Kariv (2001) (ÇKb).

2. BACKGROUND

When a sequence of Bayes-rational *dm*s make identical once-in-a-lifetime decisions after viewing a private signal as well as all predecessors' decisions, they may be rationally tempted to revise their prior beliefs given the information revealed from the history of decisions. This situation, in which the behavior of an individual indirectly affects the welfare of others via revealed information, is referred as an *information externality*.

The above framework is introduced by Banerjee (1992) and BHW to explore the problem of observational learning. The basic concern of these papers is to understand the behavior of Bayes-rational *dm*s in the presence of a *pure informational externality, i.e.,* where a *dm*'s payoff is independent of others' actions. Their common conclusion is that eventually everyone would make an identical decision and those decisions may not be the best possible. That is, uniform behavior might be inefficient. To clarify, when a *dm* observes a history in which all the previous decisions are identical, she may either deduce that each of her predecessors has private information that favors this particular decision, or some may have contradictory signals but not convincing enough to lead them to deviate to another decision. Therefore, she will join the herd, unless her private information is very convincing against the information revealed from the decisions she observes.

The most comprehensive study of observational learning is provided by Smith and Sørensen (2000) (SS). One of their points of departure from BHW is to assume that *dm*s have a continuous rather than a discrete signal space. Therefore, with discrete actions, past decisions may never perfectly reveal the signals upon which they were based upon. As a result, while BHW find that a cascade, and consequently a herd arises, SS show that with a continuous signal space a herd arises, yet an informational cascade need not arise. That is, even during a herd, when making a decision everyone takes her private signal into account in a non-trivial way. Hence, if the signals were different so might the *dm*s' actions.

In a seminal paper AH investigate observational learning experimentally. Their design is based on the binary-signal-binary-action model of BHW[5]. In their setup there are two decision-relevant events, say A and B, which are *ex ante* equally likely to occur and two corresponding signals a or 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 problem is to predict which of the events takes place based on a private signal and the history of decisions. AH implement the above setup by putting 2 balls labeled a and 1 ball labeled b in an urn A and *vice versa* in an urn B[6]. Thus, given a signal a, the posterior probability of event A is 2/3.

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

[6] Besides this symmetric design in which Bayes' rule corresponds to a simple counting heuristic, AH report an asymmetric design in which counting can be distinguished from Bayesian behavior.

With such a discrete signal structure, whenever two consecutive decisions coincide, all successors should imitate the decisions of their predecessor regardless of their private signals. That is to say that Bayes' rationality leads to an informational cascade. To clarify, note that when the two initial decisions coincide, say A, A, all successors should also choose A even if their private signals are different, b. AH report that rational cascades formed in most rounds in which initial decisions coincide and that about half of the cascades were incorrect. That is, resulting a pattern of mistakes.

In our setup, unlike AH, while there are two events, which are *ex ante* equally likely to occur, there is a continuous signal space. The following example illustrates the importance of this difference. Suppose that a cluster of *dm*s acting alike is followed by someone who deviates. Then, what may successors conclude from observing the deviation? In an AH discrete signal world, such a deviation is either irrational in the Bayesian sense or a tremble. On the other hand, outside the discrete signal world, successors could also deduce 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 since it is more decisive than the information revealed from past decisions. Note that in AH's discrete-signal-discrete-action setup all herds are cascades since once two consecutive decisions coincidence beliefs are such that no discrete signal can lead to a deviation. In contrast, our continuous-signal-discrete-action setup, along with our belief elicitation method, enables us to completely distinguish cascades and herds.

3. EXPERIMENTAL DESIGN

The experiment was run at the Experimental Economics Laboratory of the C.V. Starr Center for Applied Economics at New York University. The 40 subjects in this experiment were recruited from undergraduate economics classes at New York University. In each session eight subjects participated as *dm*s. After subjects read the instructions (see Appendix) they were also read aloud by an experimental administrator. 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 assured anonymity and an effective isolation of subjects[7] in order to minimize any interpersonal factors[8] that may cause a superfluous tendency towards uniform behavior.

[7] Participants' 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 silently throughout the session. At the end of a session, participants were paid in private according to the number of their working-stations.

[8] Other approaches that have been proposed for uniform social behavior are: (*i*) preference for conformity (see Jones (1984)), (*ii*) sanctions on deviants, (*iii*) payoff complementarities, and (*iv*) communication.

A session consisted of 15 independent rounds[9], each divided into eight turns. In all 15 rounds of the experiment, the following process has been repeated. In each round, all 8 subjects took decisions sequentially in a random order. A round started by having the computer select randomly eight numbers from the set of real numbers $[-10, 10]$. The numbers selected in each round were independent of each other and of the numbers selected 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 by his or her predecessors in that round. After observing these actions and before being informed of her private signal, each subject was asked to state a number between -10 and 10 (a cutoff), for which she would take action A if her signal is above the cutoff and action B if it is 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 signal. Then, the computer recorded her decision as A if the signal was higher than the cutoff she stated. 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. Everyone whose decision determines their action as A earned \$2 if this sum was positive (or zero) and nothing otherwise. On the other hand, everyone whose decision determines their action as 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. Hence, all 5 experimental sessions consisted of 15 rounds of 8 decisions each.

By eliciting cutoffs, we can recognize a subject who is following a cascade behavior, i.e. acting irrespective of her private signal, as one who reports a cutoff -10 or 10. That is, if her reported cutoff is -10 (10), her belief is that the sum of all signals is positive (negative), and accordingly she wishes to take action A (B), for any realization of her private signal. In contrast, a subject who joins a herd but does not follow a cascade behavior is one who reports a cutoff in the open interval $(-10, 10)$, indicating that for some signal she is willing to make either decision, and given the realization of her private signal she acts like her predecessors did. However, she might have acted differently if the realization of her signal had been different. Hence, cascade behavior is identifiable by the choice of a cutoff while joining a herd is identifiable by the realized action. Furthermore, we shall say that in the laboratory

[9] At 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.

a cascade occurs when from some subject on all follow a cascade behavior, and

herd behavior occurs when from some subject on all take the same action.

Note that subjects might benefit from the observation of past decisions simply because they reveal partial information on what the sum of the eight numbers is, which is the underlying experimental decision-relevant event. An observed history of decisions determines a cutoff such that if a private signal is greater than the cutoff the subject should take action A and if it is less she should take action B. Moreover, there is a one-to-one relation between subjects' cutoffs and their beliefs about what the realization of the uncertain decision-relevant event is.

4. SOME THEORY

4.1. The Bayesian solution

The Bayesian solution of the decision problem underlying our experimental design is outlined by ÇKa[10]. In order to formulize it, let us suppose that each *dm* $n \in \{1, ..., 8\}$ receives a private signal θ_n drawn from a uniform distribution with support $[-10, 10]$[11]. Assume that private signals are *i.i.d.* across *dm*s. Each *dm* 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 otherwise[12]. All decisions are announced publicly and therefore known to all successors. Formally, *dm* n's optimal decision rule is

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

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

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

Hence,

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

[10] This section is based on the perfect information part of ÇKa.

[11] Our results described below are robust whenever private signals are distributed with *c.d.f.* F which is symmetric, *i.e.*, $F(-\theta) = 1 - F(\theta)$ for any $\theta \geq 0$, and smooth on a compact support.

[12] In ÇKa, we assume that each *dm* makes a once-in-a-lifetime decision, to invest or not to invest, and that the payoff-relevant state of the world is determined by the uncertain value of the investment, $\sum_{n=1}^{8} \theta_n$. That is, we let a *dm* get the value of the investment if she invests and zero otherwise. For the experiment, this payoff schedule would have lead subjects to be risk averse although there is no difference in the decision problem between this payoff schedule and the one that was actually used.

It readily follows that the optimal decision, as a function of the realized history of decisions, takes the form of the following *cutoff strategy*,

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

where

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

is the optimal history-contingent cutoff which inherits all the information that *dm* 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 *dm* n's behavior, and thus it is the object of our analysis.

Note that the signal structure is informative in the sense that if, for example, the sum of all signals is positive, one is more likely to receive a positive signal. Yet, for any population greater than two, there is no private signal for which a *dm* can resolve the uncertainty by herself. This is referred in the theoretical literature as *bounded beliefs*. Moreover, the decision problem is under incomplete and asymmetric information. That is, *dm*s are uncertain what is the underlying the decision-relevant event, $\sum_{i=1}^{8} \theta_i \geq 0$ or $\sum_{i=1}^{8} \theta_i < 0$, and the information regarding it is shared asymmetrically among them. Moreover, note that *dm*s may benefit from observing past decisions solely because it reveals some information about the private signals on which they were based upon.

In order to gain some intuition, in what follows we discuss the Bayesian reasonings of the first few *dm*s. The first *dm*'s decision is based solely on her private signal, thus her 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 *dm* observes the first's decision, she conditions her decision on x_1. Thus, according to (2),

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

To clarify, if for example $x_1 = A$ then $E\left[\theta_1 \,|x_1 = A\,\right] = E\left[\theta_1 \,|\theta_1 \geq 0\,\right] = 5$ and thus it is optimal for the second *dm* 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 *dm* might 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 *dm* reveals that her private signal is contrary to and stronger than the expected value of the first's private signal. Therefore, when the third *dm* observes a deviation, her cutoff is more sensitive to the second's action. By

(2), a simple computation yields that the third dm's cutoff rule is,

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

Proceeding with the same analysis, we find that if the first three dms choose A, the fourth dm will choose A as long as $\hat{\theta}_4 \geq -8.75$; if the first four dms choose A, the fifth dm will choose A as long as $\hat{\theta}_4 \geq -9.375$; and so on. Thus, the longer the sequence of dms who choose A, the harder it is for a single dm to choose action B, even if her private signal is very negative. That is, the longer the cluster of dms who have acted alike, the harder it is for a successive dm to deviate, even if her information is contrary and very decisive. Moreover, each dm who joins the crowd reveals less information about her own signal and makes it harder for successors not to join as well.

Another concept discussed in ÇKa is the *overturning of behavior*. That is, even if many dms have acted alike, it is possible that, because of a rational deviation, the information revealed from the history up until that point cancels out. Suppose that a cluster of dms acting alike is followed by a dm who deviates since she receives a decisive signal favoring the contrary action. Since such a deviation reveals fairly clear-cut information regarding the signal it has been based upon, the newly revealed information slightly dominates the information that had been already accumulated. That is, all the accumulated information favoring a particular action is offset by the newly revealed information. Thus, a deviator induces her successor dm to be slightly in favor of joining the deviation. For example, if the fourth dm chooses B after her three predecessors choose A, her decision reveals that her signal is in the interval $[-10, -8.75)$. In such a case, according to (2), the fifth dm will choose A as long as $\hat{\theta}_5 \geq 0.625$. In summation, the longer the cluster of dms acting alike, the bigger the asymmetry between the information revealed by the same decision and a deviation. Moreover, no matter how many dms have acted alike, it is always possible that one dm endowed with an extreme signal will cancel up the history up until that point.

In Figure 1, we illustrate the processes of cutoffs for various histories. In the top left panel, all dm choose action A. Therefore, along time dms become more confident that action A is the profitable one. Other panels illustrates the dynamics in the presence of a deviation. For instance, in the top right panel the first five dms take action A and dm six deviates. Consequently, her deviation reveals considerable information that her signal is in favor of action B. As a result of this newly revealed information, dm seven's cutoff is very close to zero but yet in favor of action B.

Figure 1

4.2. A note on herds and cascades

Recall that any history of actions is public information shared by all successors. Thus, all the information revealed by any history is already accumulated in the following *dm*'s cutoff. As a result, the cutoffs of neighboring *dm*s alter only by the new information revealed from the predecessor's action. To be exact, $\hat{\theta}_n$ is different from $\hat{\theta}_{n-1}$ only by $\mathbb{E}[\theta_{n-1} \mid \hat{\theta}_{n-1}, x_{n-1}]$. As a result, the cutoff rule exhibits the following recursive structure,

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

with the following updating rule,

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

Substituting (4) in (3) yields that the cutoff dynamics abides by the following stochastic process,

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

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

Recall that an informational cascade takes place when from some *dm* on everyone's decision is independent of their private signal and thus does not provide any additional information for the next *dm* who will make a decision. The impossibility of an informational cascade follows immediately from (5) since for every *dm* n, $-10 < \hat{\theta}_n < 10$. That is, in making a decision, every *dm* takes her private signal into account in a non-trivial way. Thus, from a theoretical point of view informational cascades are a mistake[13]. Behaviorally, however, they occur when one sets a cutoff of -10 or 10. Hence, cascades are a behavioral phenomenon and not a theoretical one.

Furthermore, an informational cascade implies a herd but herd is not necessarily the result of an informational cascade since in a herd every *dm* chooses the same action, but they might have chosen a different sequence of actions if the realization of their private signals had been different. As to the probability of herd behavior, the cutoff dynamics given in (5) captures the theoretical distinction between herd behavior and informational cascade. Note that when all choose action A (B) the cutoff process tends fast to -10 (10). As a result, an informational outcome in which all base their decision almost entirely upon their predecessor's action arises, and imitation becomes more likely along the line of *dm*s. Precisely, let p_n denote the

[13] Note that in BHW's discrete signal setup a cascade, referring to an infinite cluster of *dm*s acting irrespective of the content of their private signals, eventually occurs. However, generically cascades need not arise.

probability that all dms from some dm n onwards make the same decision, conditional that all predecessors to n also made the same decision. For a population arbitrary large, beginning from the first dm, we find that as $n \to \infty$

$$p_1 = \frac{1}{2}\frac{3}{4}\frac{7}{8}\cdots(1-\frac{1}{2^n})\cdots$$

and thus

$$\begin{aligned}\log p_1 &= \log\left[\prod\nolimits_{n=1}^{\infty}(1-\frac{1}{2^n})\right] \\ &= \sum\nolimits_{n=1}^{\infty}\log(1-\frac{1}{2^n}) \\ &\approx -\sum\nolimits_{n=1}^{\infty}\frac{1}{2^n} \\ &> -\infty\end{aligned}$$

Hence, there is a positive probability that herd behavior arises from the first dm. Moreover, $p_{n+1} > p_n$ for all n and in fact $p_n \to 1$ as $n \to \infty$ [14]. In other words, a herd must arise even though a cascade never arises. Summarizing,

> Although informational cascades are a behavioral phenomenon, from a theoretical point of view cascades are impossible.
>
> Notwithstanding with the theoretical impossibility of informational cascade, the theory predicts that herd behavior must arise.

5. EXPERIMENTAL RESULTS

Over all sessions[15], herds, in the sense that from some point on all subjects acted alike, of at least five subjects, were observed in 28 of the 75 rounds (37.3 percent). Herding was also the outcome in 25 of the 37 (67.6 percent) rounds in which herding should have occurred according to Bayes' rule. Moreover, all herds, except one, turned out to be on the correct decision. In contrast, AH report that about half of the herds turned out to be on the incorrect decision. A possible reason behind the difference is the relative richness of the continuous signal space, and because in our experiment subjects could fine tune their decisions by choosing a cutoff strategy instead of making a binary decision directly. As we show in Table 1, out of the 28 herds, 14 involved all 8 subjects acting alike.

[14] In ÇKa we show that the martingale property of the stochastic cutoff process (Note that $\mathbb{E}[\hat{\theta}_n \mid \hat{\theta}_{n-1}] = \hat{\theta}_{n-1}$) establishes convergence of beliefs and actions. That is, by the Martingale Convergence Theorem $\{\hat{\theta}_n\}$ converges to a random variable $\hat{\theta}_\infty$ almost surely as $n \to \infty$. Hence, it is stochastically stable in the neighborhood of the fixed points, namely -1 and 1, meaning a limit-cascade. Moreover, we argue that convergence of the cutoff process implies convergence of actions. In other words, the behavior settles down in some finite time and is consistent with the limit learning.

[15] For a complete data appendix see http://homepages.nyu.edu/~sk510/Research.html

Table 1

Note in Table 1 that when a subject observed a history in which all previous decisions were identical, she was typically in favor of joining the herd by setting her cutoff far-off from zero in a direction consistent with acting like her predecessors for a larger set of private signals. Moreover, we can readily notice that cascade behavior, i.e. setting a cutoff of -10 or 10, was in general increasingly likely to occur towards the last turns of a round. However, an informational cascade is not a necessary condition for a herd. For instance, subjects in the last turns of session/round 2.7 (the seventh round in the second session), 4.9, 5.9 and 5.5 were in herds but still set there cutoffs in the interior of their signal support, and thus did not disregard ignore their private information in their decision. The general trend of cutoffs reported in the laboratory in rounds in which all subjects acted alike is illustrated in Figure 2. Since the cutoff strategy is symmetric around zero, we use the average of the cutoffs absolute values to see that as subjects observed more identical past actions, they became more confident about the profitability of the herded action.

Figure 2

Perhaps our most unexpected result, at least from the theory perspective, is that informational cascades, in the sense that from some subject on all set a cutoff of -10 or10 and thus acted irrespective of their signals, were observed in 26 rounds (34.7 percent). Out of these 26 rounds, in 1 round the last four subjects followed a cascade behavior, in 4 rounds the last three subject, in 11 rounds the last two subjects and in 10 rounds the last subject followed a cascade behavior. In addition, a cascade behavior was observed 32 times outside of informational cascades[16]. In Table 2, we show the rounds in which the longest informational cascades occurred in any of the sessions.

Table 2

Only 7.2 percent of all decisions, excluding the first decision in each round, did not follow the inference which is consistent with the information revealed from past decisions. For example, in round 2.13 (see Table 2), the second subject was in favor of decision A even though the decision she observed was B. Moreover, the third subject favored decision B, where according to Bayes' rule or majority rule she should have favored decision A or be indifferent respectively.

Since herd behavior develops frequently, and all herds, except one, turned out to be correct, we next turn our attention to answering the following question:

[16] Out of all 40 subjects, only two subjects followed a cascade behavior in all rounds in which they participated.

How efficient were our subjects in using the information revealed from the actions of their predecessors?

To measure this efficiency, consider the following thought experiment. Say a subject was able to observe both the actual cutoffs and the actions of all predecessors. Since once predecessors may have made errors, armed with this information she could then estimate the expected value of the signal received by each of her predecessors and then set her cutoff to be a best response to this information. Call this cutoff rule the *best response cutoff rule*. Put differently, the best response cutoff strategy is a subject's best response to her predecessors' actual cutoffs. As such, the best response cutoff is as if one anticipates perfectly the errors of others and acts accordingly. On the other extreme, assume that a subject completely ignores the history and simply follows her own information. Under this circumstance, she sets her cutoff to be zero and thus her action is determined solely by her private signal. Call this cutoff rule the *follow-own-signal cutoff rule*. Since the best response cutoff rule and the follow-own-signal cutoff rule present behavior on the opposite end of the extreme, we are interested in discovering which rule better arranges the data.

To answer the question posted above we measure both the extent to which subjects did worse than choosing a best response cutoff and the extent to which they did better than choosing only according to their private information. First, we shall use a sign test of the following type. By turns, take any subject actual cutoff and compare it to the best response cutoff and follow-own-signal cutoff. Score the choice of the subject as a plus if the square difference between her actual cutoff and the best response cutoff is less than the square difference between her actual cutoff and the follow-own-signal cutoff. Otherwise, score it as a minus. Then, we test the hypothesis that the median difference between these two groups is zero, meaning that the probability of observing an actual cutoff that is closer to be a best response cutoff is exactly equal to the probability that it is closer to be a follow-own-signal cutoff. In each decision turn, with the exception of the first decision turn[17], under the null, the probability of getting a cutoff closer to the best response cutoff would be exactly equal to the probability of a cutoff closer to the follow-own-signal cutoff. In other words, the probability of observing a cutoff closer to the best response cutoff is 1/2. We then calculate the probability of observing at least as the observed number of pluses in each of the turns. By using a normal approximation, we are able reject the null hypothesis in favor of the one-tailed alternative that the actual cutoffs were closer to be follow-own-signal cutoffs in the second decision turn (p-value 0.003). Besides, we can reject it in favor of the one-tailed alternative that the actual cutoffs were closer to be best response cutoffs in the last decision turn (p-value 0.053). In all other turns, we can not reject the null hypothesis.

[17] In the first decision turn, follow-own-signal cutoff rule is the best response cutoff.

In addition to categorizing decisions in this binary fashion, we can also proceed as AH did and use subjects' expected earnings payoffs to test how efficiently subjects had processed the information revealed from the history of decisions. That is, to what extent did subjects worse than best responding to the decisions of their predecessors, and to what extent they did better than choosing only according to their private information. To this end, we let the *best response* (*follow-own-signal*) *payoff* be the expected earnings for a subject whose cutoff is the best response (follow-own-signal) cutoff. Furthermore, let the *actual payoff* be the expected earnings for the subject actual cutoff. Taking the best response payoff as a benchmark, we shall use subjects' actual payoff and follow-own-signal payoff to measure how efficiently subjects had processed the information revealed from past decisions. The sums of a subject's best response payoff, actual payoff and follow-own-signal payoff over the session she participated in will be denoted by π_{BR}, π_A and π_S[18] respectively. We shall use these payoffs to measure how efficiently subjects used the information revealed from the history to make their decisions. Let a subject's *actual* (*follow-own-signal*) *efficiency* be her actual (follow-own-signal) payoff as a percentage of her best response payoff. That is,

$$\text{actual efficiency} = \frac{\pi_A}{\pi_{BR}} 100 \tag{6}$$

and

$$\text{follow-own-signal efficiency} = \frac{\pi_S}{\pi_{BR}} 100 \tag{7}$$

Over all subjects, actual efficiency averaged 90.4 percent while follow-own-signal efficiency averaged only 74.1 percent, indicating a surprising aptitude to process the information revealed by others' decisions, and to carry out a best response based on this information. The high average actual efficiency suggests that subjects approximate very well the extent to which others make errors and consider it in processing the information revealed from their decisions. This preliminary finding leads us to provide an econometric analysis that internalizes the effects of decision errors.

6. AN ECONOMETRIC ANALYSIS

Since our analysis in the previous section indicates that subjects take into consideration the possibility that predecessors make errors. In other words, incorporating limits on the rationality of others. In what follows, we attempt to formulize this insight by recursively estimating the process of cutoff determinations given by (5) adjusted for decision errors and independent shocks. To this end, let y_n be the vector of actual cutoffs reported by subjects when they participated as $dm\ n$ in the laboratory, and let z_n be

[18] Note that, over a session, the expected earnings for a subject whose cutoff is the follow-own-signal cutoff, π_S, is \$15.

the vector of cutoffs based on the observed histories of decisions adjusted to previous decision errors. That is,

$$y_n = \alpha_n + \beta_n z_n + \varepsilon_n \tag{8}$$

where $z_1 = 0$ and for any dm $n > 1$ and round i,

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

is the error-adjustment updating rule[19].

Notice that the parameters are estimated recursively. That is, the estimated parameters for the first dm, $\hat{\alpha}_1$ and $\hat{\beta}_1$, are employed, in turn, to estimate the parameters for the second dm, α_2 and β_2, and so on. Clearly, allowing the possibility of errors in predecessors' decisions effects successors' decision problems. Thus, the estimated value of the parameters of dm n, $\hat{\alpha}_n$ and $\hat{\beta}_n$, are used to determine the optimal decision of dm $n+1$. Note the error-adjustment mechanism given in (9) suggests that subjects anticipate the average error that their predecessors make and process it in making their decisions. This recursive econometric method tests directly how well our model predicts the behavior resulted in the laboratory. Overall, out of the 600 decisions in the laboratory, our estimation predicts accurately 467 (77.8 percent) decisions. We report the results of the estimation in Table 3.

Table 3

As to α_n, notice that it parameterizes an information processing bias such as conformity preferences, i.e. an additional tendency to go along with the crowd. For example, for dm n, $\alpha_n < 0$ ($\alpha_n > 0$) points a bias toward action A (B). Another irrational bias that is captured by the α_n, is a tendency of subjects towards an alternative that was profitable in previous rounds. For instance, a subject who is in favor of action A, can constantly set a negative cutoff, since this action was the profitable one in most of the previous rounds[20]. Our econometric results show that $\hat{\alpha}_n$ is not significantly different from zero in all turns. This is also an indication that our experimental design controls for these factors successfully. As was pointed out by AH, the more complex the setting such irrational biases are more likely to occur. That is, as the decision-making is more involved, subjects are less likely to be able to act rationally. Yet, the fact that we

[19] Note 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 of the interval $[-10, 10]$. In such cases, we set it to be at the corresponding boundary. That is, whenever $z_{n,i} < 10$ ($z_{n,i} > 10$) we set it to 10 (-10).

[20] It was made clear in the instruction that the numbers selected in each decision problems are independent of the numbers selected for any of the other decision problems.

can not reject the hypothesis that $\hat{\alpha}_n = 0$ for all n is a strong support against a superfluous bias in favor of following predecessors' decisions or own successful decisions in previous rounds.

The β_n parameters measure the weights given by subjects to the information revealed from the history of decisions. When the information processing biases diminish, $\alpha_n \to 0$, as $\beta_n \to 1$ the behavior tends to be according to the best-response cutoff rule. That is, when $\alpha_n = 0$ and $\beta_n = 1$, according to (8), the laboratory decision-making conforms perfectly with the optimal history-contingent history cutoff rule given by (2). Conversely, the behavior tends to be based on the private signal as $\alpha_n \to 0$ and $\beta_n \to 0$. Notice that when $\alpha_n = \beta_n = 0$, equation (8) specifies a cutoff zero which is simply choosing according to private information. In general, any $\beta_n < 1$ indicates that subjects undervalue the information revealed from the history of decisions relative to their private information. This is plausibly a response to beliefs that others' decision errors are such that their actions provide excessively noisy information as to their private signals. In contrast, any $\beta_n > 1$ is a sign of overvaluing the information revealed from predecessors' decisions relative to their private information. In particular, this is likely to be overvaluing of the information revealed from decisions of predecessors who joins an established pattern, and thus, their actions reveal hardly any information about their signals.

As reported in Table 3 above the $\hat{\beta}_n$ parameters are bounded away from zero and one[21]. Note that when subjects assigned to be the second *dm* they tend to sharply undervalue the first's decision, $\hat{\beta}_2 = 0.22$. Thus, our econometric results strongly suggest that when subjects take decisions early they have a strong tendency to follow their own signals instead of employing Bayesian reasoning. However, as illustrated in Figure 3 below, after the second decision turn β_n exhibits an upward trend showing that along the line of *dm* subjects tend more towards engaging in Bayesian updating. Put differently, over time the information revealed from the history of actions is relied on more and subjects become increasingly likely to imitate their predecessors.

Figure 3

Next, we incorporate our findings to elucidate why, even though the theory predicts the impossibility of an informational cascade, it still arises in the laboratory. To begin with, let us illustrate, with an extreme example, how the behavior alters because of errors in previous decisions. Suppose that the second *dm* ignores the information revealed from the first's decision and follows her own signal, and that the third *dm* knows that this error has

[21] Precisely, only $\hat{\beta}_1 < 0$. However, it mainly resulted from the fact that according to the optimal history-contingent cutoff rule (2), $\hat{\theta}_1 = 0$. Moreover, as apposed to any $\hat{\beta}_n$ of other turn, which measure the weights given by subjects to the information revealed from previous decisions, $\hat{\beta}_1$ has no significant importance.

been made. Since the third *dm* observes that both her predecessors acted alike and because she identifies the nature of the second *dm*'s error, it is optimal for her to join the herd for any realization of her private signal. In other words, given that the cutoffs of both the first and second *dm*s are zero, if the third *dm* observes a history A,A (B,B), it is optimal for her to set her cutoff to -10 (10) since her expected value of both the first and the second *dm*s private signals, θ_1 and θ_2, are 5 (-5). Thus, cascade behavior may arise with error considerations. More precisely, when others make errors, the information revealed from their decision is not necessarily less informative. For instance, if one follows her own signal, her action reveals more information about her private signal, and, as a result, may induce an informational cascade.

As was argued above, our econometric results point out that, as they assigned to be late *dm*, subjects move towards being Bayesian. Moreover, when Bayes-rational *dm*s act alike, their decisions reveal partial information. In contrast, when *dm*s are non-Bayesian, as those who follow their own signal, more information is made public. Hence, since subjects who are early *dm* are more likely to act according to their own signal, more information is available to late *dm* subjects how tend more to be Bayesian. Put differently, earlier subjects bring more of their own information into the learning process. As a result, it is possible that, along the line of *dm*s, the history of decisions contains information which dominates any private information. Thus, internalizing the effects of errors in previous decisions, can induce acting irrespective of the private information. In Figure 4, we illustrate it by comparing the theoretical Bayes' cutoff process (5) and the estimated error-adjusted one for a history in which all choose action A. Note that for early *dm*s, the estimated cutoffs are a lot above the theoretical ones, meaning a relative predisposition to follow their own information. However, along the *dm*s' line, the gap between the theoretical and estimated cutoffs diminishes. Hence, we find strong support that the Bayesian solution as given by the cutoff strategy (1) and (2) adequately predicts the behavior of a large portion of subjects in the laboratory.

Figure 4

One might argue that the above result is not robust since a history in which all subjects act alike is in our favor. However, Bayes-rationality performs well as a predictor of the behavior observed in the laboratory also in cases when the behavior overturns. Recall that, in theory, as more *dm*s act alike, the bigger the asymmetry between the information revealed by an imitator and a deviator. Thus, a deviator induces her successor to be slightly in favor of joining the deviation. On the other hand, if *dm*s follow the majority, a successor to a deviator would not be tempted to join the deviation. In Figure 5 we illustrate by an example how well Bayes' rule predicts subjects' behavior where an overturn occurs. We do so by comparing the theoretical Bayes' cutoff process (5) and the estimated one for the history in which the third dm overturns the behavior. The first two dms choose action A but the third deviates. The estimated cutoffs show that all successors follow the deviation, as according to the theoretical findings, although they undervalue the revealed information a little less than what the theory predicts. Note that while theoretically, according to (5), the fourth dm is somewhat in favor of joining the deviation, i.e. $\hat{\theta}_4 = 1.25$, her estimated cutoff is -0.9. That is, she is much less enthusiastic of joining the herd but not in favor of joining the deviation. This is an obvious reaction to the possibility that her predecessors deviation is erroneous.

Figure 5

7. CONCLUDING REMARKS

Observational learning describes many economic situations in which dms learn by observing the behavior of others. In this paper we provide a test of observational learning theory by effectively minimizing potentially confounding affects. The contribution of this paper is two-fold: First, our experiment designed to test a continuous signal space and a discrete decision set model, which theoretically yields a behavior different from the one of the basic binary-signal-binary-action herding model tested by AH. Second, we elicit subjects' beliefs about the underlying decision-relevant event by having subjects stating a cutoff instead of taking an action directly. Accordingly, we are able to distinguish between two behavioral phenomena - informational cascades and herd behavior. Moreover, this leads us to address the more appealing question which is how well Bayes-rationality approximates the actual behavior observed in the laboratory.

We find that herd behavior develops frequently (37.3 percent) in the laboratory. While AH report that in their experiment frequently uniform behavior turned out to be on the incorrect decision, we find that all herds, except one, turned out to be correct. This is possibly due to the relative richness of the continuous signal space, and because subjects could fine tune their decisions by having a cutoff strategy instead of making a binary decision directly. Moreover, as apposed to the impossibility of informational cascades predicted by the theory, we find that cascades often arise (34.7 percent). Thus, we conclude that although cascades are not a theoretical phenomenon they are a behavioral one.

Furthermore, since, almost whenever occurred, uniform behavior turned out to be correct and in view of the high earnings efficiency measure (90.4 percent) we reason that subjects used the information revealed from past decisions efficiently by approximating fairly well the extent to which others are making errors. For that reason, we provide an econometric analysis which internalizes the effects of errors in previous decisions. In addition to verifying that our experimental design indeed minimizes all other incentives to go along with the crowd, our estimations provide measures for the weights given by subjects at each decision turn to the information revealed from the history of decisions. Virtually, that is disclosing what subjects believe about predecessors' decision rules or, in other words, how they incorporate limits on the rationality of others. We find that subjects, although undervalue the information revealed from past decisions, tend towards Bayesian updating. We incorporate this result to elucidate that as early subjects undervalue the information revealed from the history and thus act according to their own signal, more information is available to late subjects. As a result, the cascade behavior observed in the laboratory may after all be rational. In conclusion, we find strong evidence that modified Bayesian behavior that incorporates limits on the rationality of others adequately explains the behavior in the laboratory. We believe that this result poses a challenge for the theory of observational learning. Feeding back in to our theoretical results in ÇKa, our experimental results suggest to introduce the possibility of noise or trembling in to the model.

In a recent paper, ÇKb, we present an experimental evidence that addresses many economic situations in which individuals learn by observing only the behavior of those who recently faced their decision problem, i.e. they have *imperfect information* about the history of actions. We test the behavior under such an imperfect information structure, in particular when all *dm*s observe the same number of their immediate predecessors' decisions. Furthermore, with the intention of underlining the differences between observational learning under perfect and imperfect information, we let each subject to observe only her immediate predecessor's decision. The experimental design we employ in ÇKb is similar to the one we employed in this paper to test observational leaning under *perfect information*. Our theoretical results in ÇKa in a game where each *dm* observes only her predecessor's binary action indicate that the Bayesian outcome is considerably different from the outcome under perfect information. In particular, the behavior cycles forever but exhibits longer and longer clusters of uniform behavior, punctuated by switches that are more rare.

REFERENCES

Allsopp, L. and J. Hey (1999) "Two Experiments to Test a Model of Herd Behavior", University of York, Discussion Paper.

Anderson, L. (1999) "Payoff Effects in Information Cascade Experiments", College of William and Mary, Discussion Paper.

Anderson, L. and C. Holt, (1996) "Classroom Games: Information Cascades", *Journal of Economic Perspectives*, **10**, 187-193.

Anderson, L and C. Holt, (1997) "Information Cascades in the Laboratory", *American Economic Review,* **87,** 847-862.

Banerjee, A. (1992) “A Simple Model of Herd Behavior”, *Quarterly Journal of Economics,* **107**, 797-817.

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**, 992-1026.

Bikhchandani, S., D. Hirshleifer and I. Welch (1998) “Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades”, *Journal of Economic Perspective,* **12**, 151-170.

Çelen, B. and S. Kariv (2001), “A Model of Observational Learning Under Imperfect Information”, mimeo, NYU.

Çelen, B. and S. Kariv (2001), “An Experimental Test of Observational Learning under Imperfect Information”, mimeo, NYU.

Gale, D. (1996) “What Have We Learned from Social Learning?”, *European Economic Review,* **40**, 617-628.

Jones, S. (1984) *“The Economics of Conformism”*, Oxford: Blackwell.

Smith, L. and P. Sørensen (2000) “Pathological Outcomes of Observational Learning”, *Econometrica,* **68**, 371-398.

Appendix A: Tables

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	B	0	4.25	10	5	10	10	10	10
2.7	A	4	1	-2	1	-6.1	-10	0	-9.4
2.10	B	2	3	9	10	10	9.8	10	10
2.11	B	-6.1	-2	-6	-10	-8.8	0	-10	-10
2.12	A	0	-7	-10	-10	-4	-9.9	-10	-10
2.14	A	0	-7	-10	-10	-4	-9.9	-10	-10
3.1	A	0	1.5	-0.01	-2	-10	0	0	-10
4.3	A	-5	5	0	-4	-10	-9.87	-10	-10
4.9	B	5.4	10	8.69	6.4	3	9	10	8.5
4.12	A	0	0	-9	-10	-9.13	-5	-10	-10
4.15	B	10	9.99	5	0	9.9	10	10	10
5.3	B	0	4	-2	10	2	10	0	10
5.5	B	-1	0	-10	-8	-5	-10	5	0
5.9	A	0	0	-10	-10	-3	-10	-6	-9
Average**		2.4	3.9	6.6	6.9	6.8	8.1	7.2	9.1

* (Session.Round). For example, 1.11 is the eleventh round in the first session.

** Average of the cutoffs absolute values. Note that the cutoff strategy is symmetric around zero.

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

Subject Number: Decision
Cutoff
Private Signal

Session / Round	1	2	3	4	5	6	7	8	Sum of Signals
1.9	8:B 0 -5.41	6:B 0 -7.12	4:B 5 -3.72	2:B 5 -1.89	1:A 0 3.59	7:B 10 5.59	5:B 10 1.76	3:B 10 -7.95	-15.2
1.10	8:B 0 -0.35	6:B 0 -4.71	1:A* -10 -0.34	3:A 8.6 9.17	4:A 0 0.63	2:A -10 8.69	5:A -10 -7.61	7:A -10 -4.00	1.5
1.11	4:B 0 -1.44	6:B 4.25 -2.71	8:B 10 0.74	2:B 5 -4.76	5:B 10 1.87	7:B 10 -7.94	1:B 10 4.80	3:B 10 -1.06	-10.5
2.13	2:B 0 -1.5	1:A* -1 4.11	5:A* 1.5 4.11	7:A -5.9 1.35	3:A -6.6 6.42	6:A -10 -5.71	4:A -10 6.04	8:A -10 8.71	23.5
4.15	2:B 10 -6.45	3:B 9.99 -4.56	4:B 5 -6.57	1:B 0 -1.82	6:B 9.9 6.05	5:B 10 0.8	7:B 10 -8.19	8:B 10 1.29	-19.5

Key: (shaded) - Cascade behavior.
* - Decisions inconsistent with the observed history.

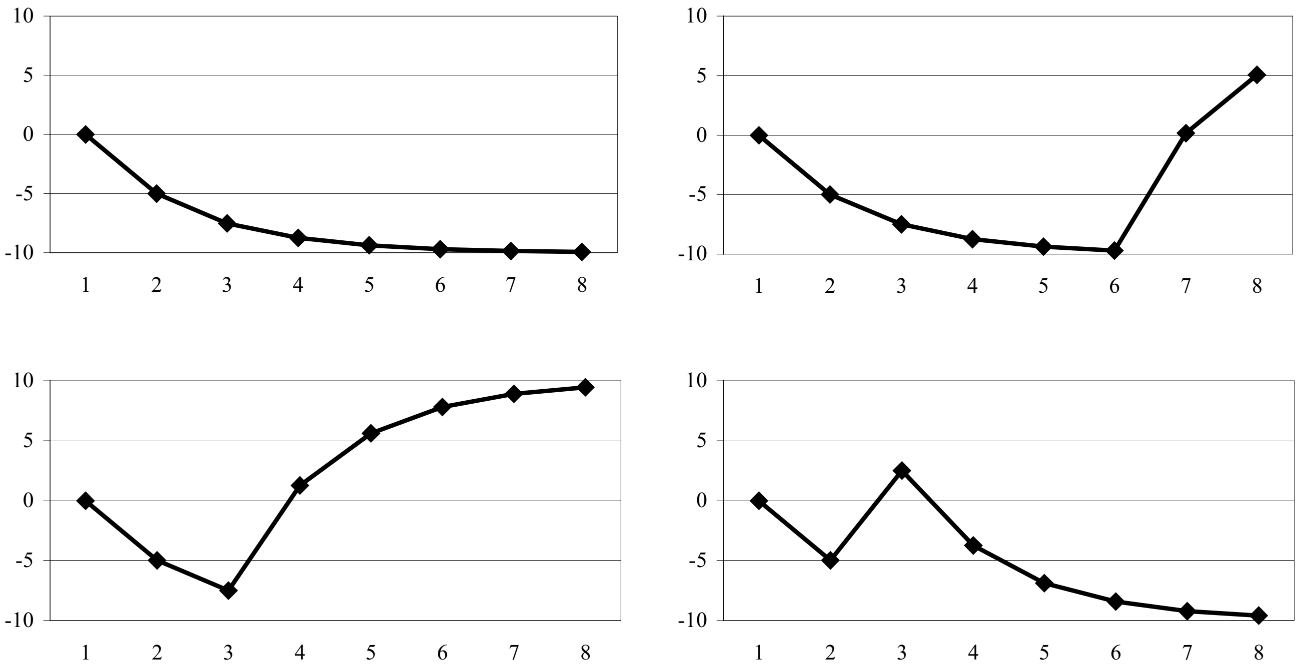
<u>Table 3: The econometric results by turn</u>

Turn	2	3	4	5	6	7	8
# of obs.	75	75	75	75	75	75	75
$\hat{\alpha}$	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-squared	0.07	0.31	0.39	0.51	0.47	0.45	0.45

Key: (standard errors)

Appendix B: Figures

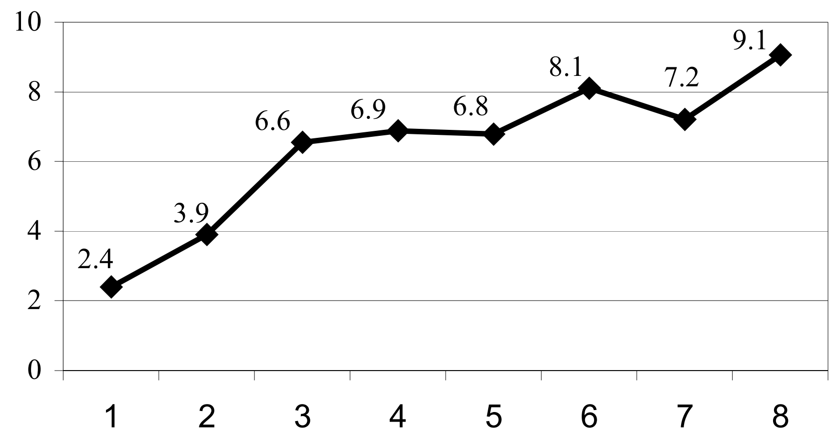
Figure 1. The process of cutoffs for various histories of eight *dm*s.



Top left: a history in which all *dm* choose A; Bottom left: the third *dm* overturns the behavior; Top right: the sixth *dm* overturns the behavior; Bottom right: both the second and the third *dm*s overturn the behavior.

<u>Figure 2. The average* cutoffs for all histories in which all subjects acted alike</u>

(session.rounds: 1.11, 2.7, 2.10, 2.11, 2.12, 2.14, 3.1, 4.3, 4.9, 4.12, 4.15, 5.3, 5.5, 5.9)



* Average of the cutoffs absolute values. Note that the cutoff strategy is symmetric around zero.

<u>Figure 3. The econometric results by turn</u>

$\hat{\beta}_n$

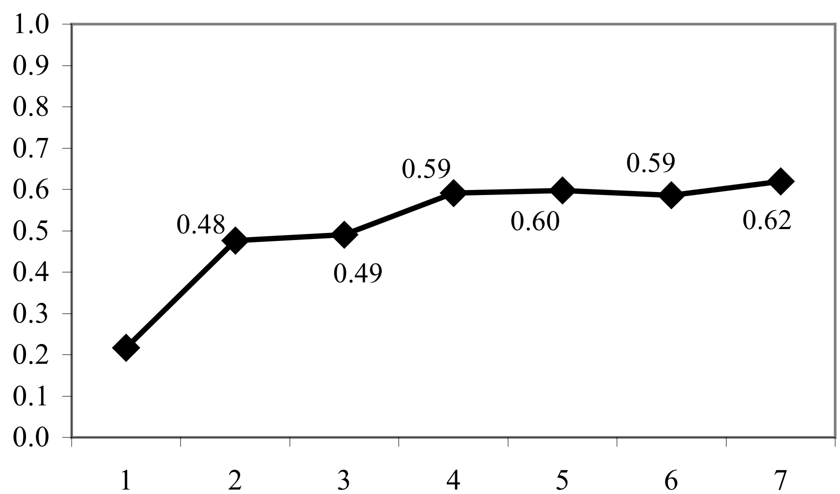


Figure 4. The process of cutoffs where all decision-makers choose action A
Bayes and Estimated

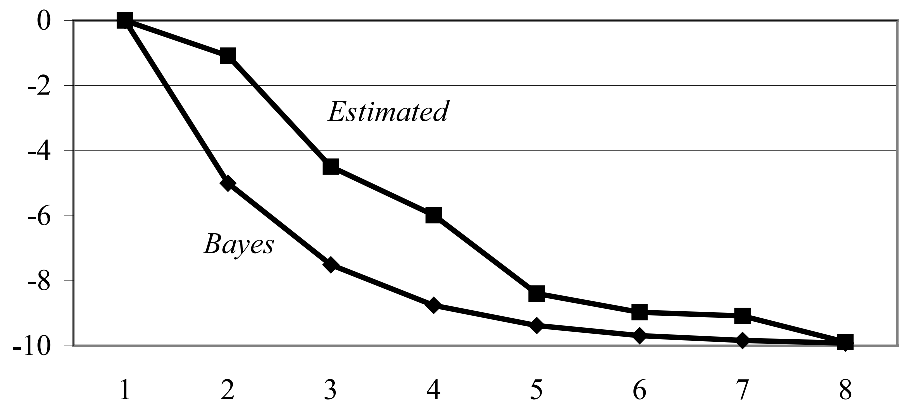
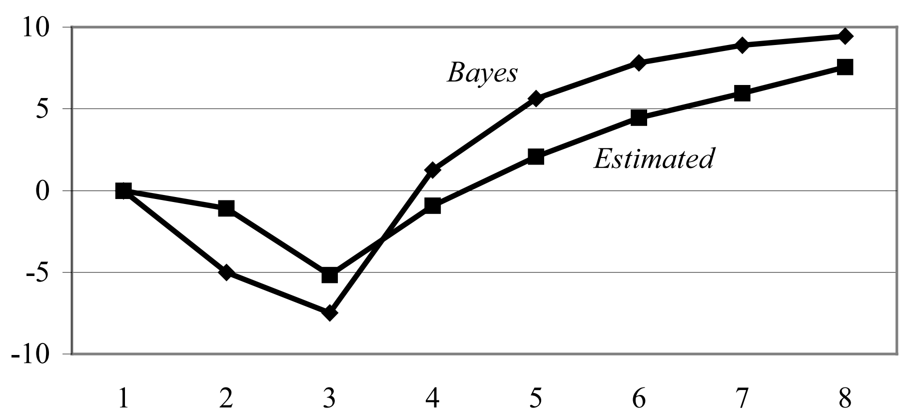


Figure 5. The process of cutoffs incase of an 'overturn' in behavior
Bayes and Estimated



Appendix C: Instructions

<u>Introduction</u>

This is an experiment in the economics of decision-making. A research foundation has provided funds for conducting this research.

Your earnings will depend partly on your decisions and partly on chance. If you follow the instructions and make careful decisions, you may earn a considerable amount of money.

At this time, you will receive $5 as participation fee (simply for showing up on time). Details of how you will make decisions and gain subsequent earning will be provided below.

In this experiment, you will participate in 15 independent and identical (of the same form) decision problems.

In all decision problems, all participants take decisions one after the other (hereafter, line of participants). In each decision problem, you will be randomly assigned the role of one of the participants, labeled 1,2,…,8 according to their position along the line of participants. When you are being called to make your first decision you will be informed of your position along the line of participants for the first decision problem. After all have participated in the first decision problem, you will be then randomly placed in the line of participants for the second decision problem. Again, upon being called to make your second decision you will be informed of your position along the new line for the second decision problem, and so on.

Your position along the line of participants of each decision problem depends solely upon chance and is independent of your position along the line of participants of any other decision problem.

Next, we will describe in details the process that will be repeated in all 15 decisions problems.

<u>A decision problem</u>

Each decision problem starts by having the computer selecting randomly eight numbers $x_1, x_2, ..., x_8$ from the set of real (including decimals) numbers [-10,10]. That is, x_1 is equally likely to be any real number between –10 and 10, x_2 is equally likely to be any real number between –10 and 10, and so on.

The numbers selected in each decision problems are independent of each other and of the numbers selected for any of the other decision problems.

You will be informed only of the number corresponding to your position along the line of participants. For example, if you are the third participant you will be told only what is the value of x_3. The value of this number is your private information and should not be shared with any of the other participants. Each of the other participants will also be informed only on the number that corresponds to her position along the line of participants.

In each decision problem, you will be asked to choose one of the actions – ***A*** or ***B***. Action ***A*** will yield you a payment of $2 when the sum of the eight numbers $x_1 + x_2 + ... + x_8$ is positive (including zero) and nothing when this sum is negative. Conversely, action ***B*** will yield you a payment of $2 when the sum of the eight numbers $x_1 + x_2 + ... + x_8$ is negative and nothing when this sum is positive.

That is, **action *A* is profitable if and only if $x_1 + x_2 + ... + x_8 \geq 0$, and action *B* is profitable if and only if $x_1 + x_2 + ... + x_8 < 0$.**

Since others may have participated in the decision problem before you, you will observe a history of actions. That is, you will be able to observe the actions that have previously taken. For example, if you are the fifth participant to choose, you will be informed what action the first, the second, the third and the fourth have taken.

Note that unless you are the first participant to do the decision problem, when it will be your turn, you will see some history of actions.

After observing the actions that have been taken by the participants before you, and **before** being informed what is the value of the number corresponding to your position along the line of participants, say x_i, you will make your decision.

Instead of choosing action *A* or action *B* directly, you will have to state the minimum number, x_i, between –10 and 10 for which you wish to choose action *A*.

Only after submitting your decision, the computer will inform you what is the value of the number corresponding to your position along the line of participants. The computer will then record your action as ***A*** if this number is equal to or higher than the minimum value you stated. Otherwise, the computer will record your action as ***B***.

For example, suppose you are the forth participant to make a decision. Then, you will observe a history of actions of the form:

Participant	Action
1	***A***
2	***B***
3	***B***

That is, x_1 was equal to or higher than the minimum value given by the first participant for choosing action ***A***. And, on the other hand, x_2 and x_3 were lower than the minimum values given by the second and third participants for choosing action ***A*** respectively.

Next, you will have to state a minimum number, say *y*, between –10 and 10 for which you wish to choose action ***A***.

Only then, you will be informed by the computer what is the value of x_4. The computer will record your action as ***A*** if $x_4 \geq y$ and will record your action as ***B*** if $x_4 < y$.

After all participants have made decisions, the computer will inform everyone what the sum of the eight numbers, $x_1 + x_2 + ... + x_8$, actually was.

Everyone whose decision determines her action as ***A*** earns \$2 if $x_1 + x_2 + ... + x_8 \geq 0$ and nothing if $x_1 + x_2 + ... + x_8 < 0$. On the other hand, everyone whose decision determines her action as ***B*** earns \$2 if $x_1 + x_2 + ... + x_8 < 0$ and nothing if $x_1 + x_2 + ... + x_8 \geq 0$.

<u>Rules</u>

Please do not talk with anyone during the experiment. We ask everyone to remain silent until the end of the last decision problem.

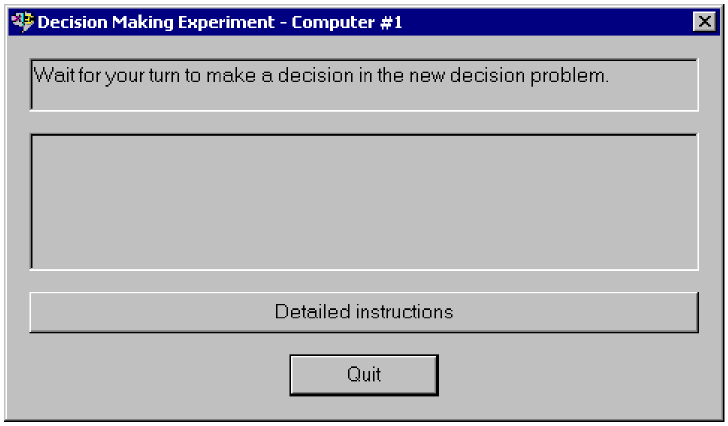
Your participation in the experiment and any information about your earnings will be kept strictly confidential. Your payment's receipt and participant form are the only places in which your name and social security number are recorded.

If there are no further questions, at this time we will move to describe the computer program that you will use to make your decisions.

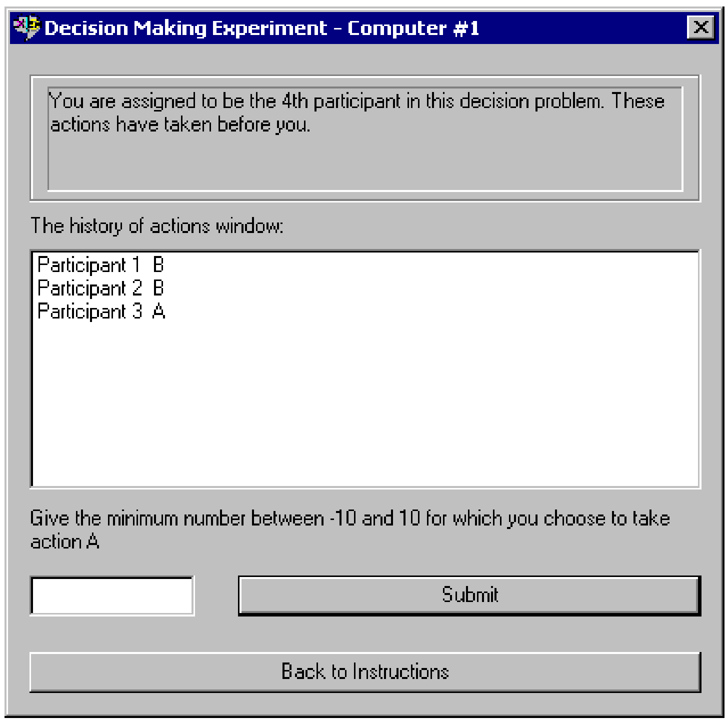
<u>The Computer Program Description</u>

To make your decisions you will use the computer. To begin with, one of the instructors will activate your program. At the beginning of the first decision problem, the **waiting window** (shown below) will pop up and will appear on your screen until it is your turn to participate.

At this point, take a minute to write down the number of the computer you are using as appears on the top line of the **waiting window.** At the end of the experiment, you should use your computer number to claim your cumulative earnings.



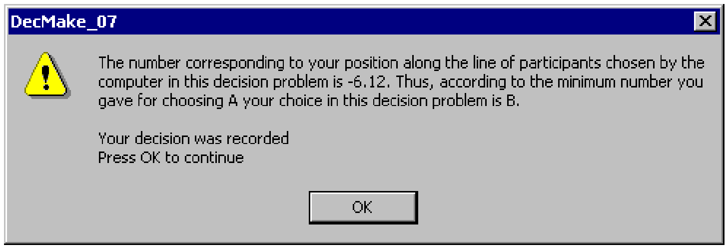
After everyone that was assigned the role of one of the participants before you have made her decision, you will be called to make your decision. At this time, the **program dialog window** will pop up and replace the **waiting window**.



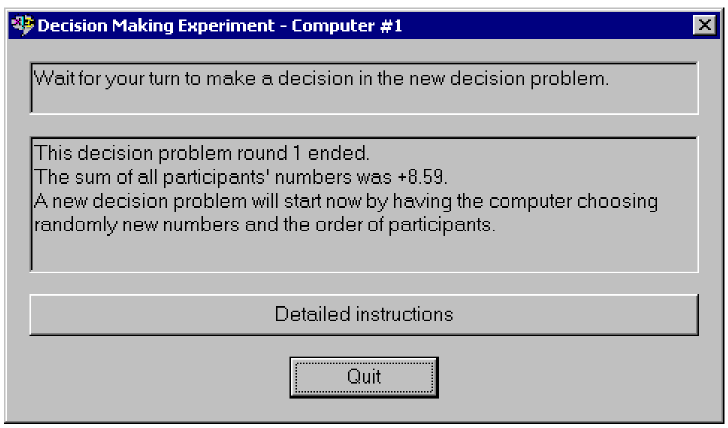
On the top of the **program dialog window,** you will be informed of your position along the line of participants for this decision problem. The history of actions is given to you in the large window that appears below.

When you are ready to make your decision, all you need to do is just use the mouse to click on the your **Give the minimum number for which you choose to take action *A* window** on the bottom of the **program dialog window**. Then, use the keyboard to fill a real (including decimals) or an integer as the minimum number between –10 and 10 for which you wish to choose action ***A***. After that, confirm your decision by clicking on the **Submit** button.

After you will submit your decision, the computer will inform you what is the value of the number corresponding to your position along the line of participants and what action has been recorded according to your decision. To end your participation in this decision problem press the **OK.**



At this point, At that point, your **waiting window** will pop up again and the next participant will be called to make her decision. The first decision problem will end after all participants have made their decisions. Then, all will be informed in the **waiting window** what the sum of the eight numbers, $x_1 + x_2 + ... + x_8$, in the first decision problem actually was.



After letting you observe the results of the first decision problem, the second decision problem will start by having the commuter selecting randomly eight new numbers $x_1, x_2, ..., x_8$ and a new participations' line. Again, you will observe the **waiting window** until it will be your turn to participate in the second decision problem.

This process will be repeated until all 15 independent and identical decision problems are completed. At the end of the last decision problem, you will be informed about the end of the experiment.

If there are no further questions, you are ready to start. An instructor will approach your desk and activate your program.