The Advice Puzzle: An Experimental Study of Social Learning Where Words Speak Louder Than Actions*

Boğaçhan Çelen, Shachar Kariv and Andrew Schotter[†]
New York University

June 16, 2003

Abstract

This paper studies how individuals learn by observing the behavior of predecessors as well as from their advice. What we find is a truly puzzling result that we call the advice paradox. This paradox can be stated as follows: subjects in a laboratory social learning situation played with and without advice appear to be more willing to follow the advice given to them by their predecessor than to copy their action, despite the fact that both pieces of information are equally informative in equilibrium. The consequence of this advice paradox is that in experiments with advice subjects tend to herd more than they do in experiments where they can only view their predecessor's action. Remarkably, these herds tend to select the correct action and, hence, advice tends to be efficiency increasing when compared to experiments where subjects can only observe their predecessor's action.

Journal of Economic Literature Classification Numbers: C91,C92, D8.

Keywords: Social Learning, Informational Cascades, Herd Behavior, Advice, Experiment.

---

*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 acknowledge helpful discussions and suggestions of the participants of seminars at several universities.

[†]Department of Economics, New York University, 269 Mercer St., 7th Floor, New York, NY, 10003. Çelen: e-mail: bc319@nyu.edu, url: http://home.nyu.edu/~bc319. Kariv: sk510@nyu.edu, url: http://home.nyu.edu/~sk510. Schotter: as7@nyu.edu, url: http://home.nyu.edu/~as7.

1 Introduction

In recent years, a great deal of attention has been paid to the problem of *social learning*. In the literature associated with this problem[1], it is assumed that people learn by observing either all of or a subset of the actions of those who have gone before them. They use these actions to update their beliefs about the payoff-relevant state of the world and then take an action that is optimal given those beliefs. Using this approach a great deal has been learned about how and why people follow their predecessors, or herd, and how informational cascades develop.

The odd aspect of the social learning literature is that it is not very social. In the real world, while people learn by observing the actions of others, they also learn from their advice. For example, people choose restaurants not only by viewing which of them are popular, but also by being advised to do so. People choose doctors not by viewing how crowded their waiting rooms are, but by asking advice about whom to go to, and so on. Thus, social learning tends to be far more social than we, economists, depict it.

In this paper, we introduce advice giving into a standard information cascade problem of the type investigated theoretically by Çelen and Kariv (2002a) and experimentally by Çelen and Kariv (2002b, 2002c)[2]. We designed the experiment so that both pieces of information, action and advice, are equally informative in equilibrium. What we find is a truly puzzling result that we call the *advice puzzle*, which can be stated as follows: Subjects in a laboratory social learning situation played with and without advice appear to be more willing to follow the advice given to them by their predecessor than to copy their action, and the presence of advice increases subjects' welfare.

For example, consider a sequence of agents who must choose between two actions A and B under incomplete and asymmetric information. In the

[1] Banerjee (1992) and Bikhchandani, Hirshleifer and Welch (1992) introduced the basic concepts and stimulated further research in this area. For surveys see, Gale (1996), and Bikhchandani, Hirshleifer and Welch (1998). Among others, Lee (1993), Chamley and Gale (1994), Gul and Lundholm (1995), Smith and Sørensen (2000), and Çelen and Kariv (2002a) provide direct methodological extensions.

[2] Anderson and Holt (1997) investigate the observational learning model of Bikhchandani, Hirshleifer and Welch (1992) experimentally and Hung and Plott (2001) replicate and extend Anderson and Holt (1997) to investigate further possible explanations for cascade behavior. Schotter and Sopher (2001, 2002, 2003) and Chaudhuri, Schotter and Sopher (2002) use an intergenerational game set-up to introduce advice into a multi-period game experiment, but the games they study, while having advice, do not have the information cascade set-up.

standard social learning problem, each agent's payoff depends on her own action and on the state of nature so it does not depend directly on the actions of other agents. However, each agent's action reveals something about her private information, so an agent can generally improve her decision by observing what others do before choosing her own action. In social settings, where agents can observe one another's actions, it is rational for them to learn from each other.

Say that each subject can either observe an action, A or B, chosen by her immediate predecessor or receive a piece of advice, A or B, from her suggesting an action. In the model that underlies our experiment, we will see that in equilibrium receiving advice or observing predecessor action are equally informative (in fact identical). Despite this informational equivalence, in the laboratory, subjects tend to follow advice far more frequently than action. In other words, if told to take action A subjects tend to do so far more often than when they observe their predecessor actually taking action A. As we will say several times during the paper, this is an example of "words speaking louder than actions."

The consequence of this result is that in experiments with advice subjects tend to herd more than they do in experiments where they can only observe their predecessor's action. Remarkably, all herds turn out to be on the correct action and hence advice tends to be efficiency increasing. Even more surprising, however, is the fact that experiments run with only immediate predecessor advice appear to be as efficient as experiments run where subjects can observe all the decisions that have previously been made, i.e., they have *perfect information* about the entire history of actions that have been taken before them. Here, of course, the theory predicts that the perfect information environment would be more efficient since the environment is far richer informationally, yet this is not what happened in the laboratory.

This paper presents the results of a set of experiments run to investigate the role of advice in social learning situations. While we will elaborate on the results discussed above, we leave the advice puzzle as a homework problem for the reader – the type of problem that professors hope some smart student solves since they do not know the answer themselves.

The paper proceeds as follows: In Section 2 we formulate our questions about the role of advice in social learning into a set of research questions that we intend to investigate and answer in the remainder of the paper. Section 3 describes our experiments that are extensions of those run by Çelen and Kariv (2002b, 2002c) modified to include advice. Section 4 presents a short review of the relevant theory underlying the experiment. Section 5 presents our results and illustrates the advice puzzle in more detail. Section 6 offers some speculation as to why people are so disposed to follow advice and Section 7 concludes.

2 Research questions

There are several questions that we would like to address in our investigation of the impact of advice on behavior in social learning experiments. In this section, we quickly spell out what these questions are and then attempt to answer them in the remainder of the paper.

Our first question is at the core of the advice puzzle since it asks whether there is a bias on the part of subjects that leads them to be more predisposed to following the advice of their predecessors rather than imitating their actions. In other words, we would like to know if subjects tend to follow their predecessors advice more often than their actions when each is observed under identical circumstances? To illustrate, consider two subjects, one performing our Action-Only experiment (observes the preceding subject's action) and another performing our Advice-Only experiment (receives the preceding subject's advice). If the Action-Only subject observed her predecessor taking action A while the Advice-Only subject is told to choose action A by her predecessor, is the conditional probability of choosing A greater in the Advice-Only experiment? This leads to question 1.

Question 1: Do subjects tend to follow advice more often than action when each is observed under identical circumstances?

Whether following advice is beneficial or not depends on whether it increases our subjects' welfare. In our experiment, it can happen if receiving advice leads subjects to act more consistently with the Bayes rational solution of our Action-Only experiment, since only then can they be maximizing their expected payoffs. Our second question is, therefore, if receiving advice in our experiment leads subjects to act more consistently with the Bayes rational theory underlying our experiment. As we will see, our experiments test a continuous-signal social learning model and use a cutoff elicitation technique (instead of making a decision per se, subjects are asked to state a cutoff that then determines their action) to elicit subjects' beliefs. Hence, the experiments furnish a natural metric upon which to perform our test – the distance of our subject's actual cutoffs (conditional on the decision turn) from that predicted by the theory. These considerations lead to the following question.

Question 2: Do subjects in the Advice-Only or Action-Only experiments act in a manner that is more consistent with the Bayesian behavior for the game played? In other words, are their cutoffs in Advice-Only experiment closer to the cutoff prescribed by the theory than those of subjects in the Action-Only experiment.

In another experiment, subjects could both receive advice and observe the action taken by their predecessor (Action-Plus-Advice). A natural question is whether this changes their behavior away from what it was in the Advice-Only experiment. In fact, the Action-Plus-Advice experiment can give us some insight into whether subjects actually value advice more than action since in some cases subjects actually give advice that differs from the action they take. In such cases, the predecessor is possibly saying, "Do as I say not as I have done," and the question is which datum is more informative and why.

A related question is whether the conditional probability of following the advice to choose an action, say A, is higher if subjects have seen their predecessor chose A as well, or is it higher when the predecessor advised A but chose B?

Question 3: Does the behavior in the Action-Plus-Advice experiment change from what it was in the Advice-Only experiment? Which information, advice or action, is more valued by the subjects? And, under what circumstances do subjects offer advice, which differs from their action?

Our final question is in some sense the most important one since the focus of the social learning literature has so often been on the question of herding and informational cascades. An informational cascade is said to occur when after some time all individuals ignore their private information when making a decision, whereas herd behavior occurs when after some time all 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. Thus, an informational cascade implies a herd but a herd is not necessarily the result of an informational cascade. More important, since advice is a key ingredient into most decisions, it would be of interest to know whether it strengthens the urge to herd or diminishes it. In addition, from a welfare point of view it is important to know if herds, when they occur, turn out to be on the correct decision. If not, the ability of advice to strengthen herding behavior is welfare decreasing.

Question 4: Are there more or fewer herds and cascades in experiments with advice? Does the welfare of subjects increase when they have access to advice?

3 Experimental design

The data underlying our experiments are generated by the experiments run by the authors during the Fall of 2001 at the Center for Experimental Social Science (C.E.S.S.) at New York University and by two previously run experiments of Çelen and Kariv (2002b, 2002c), which are also discussed here for comparison purposes. We will designate the two newly run experiments as the Advice-Only and Action-Plus-Advice experiments, and the two previously run experiments as the Perfect-Information and the Action-Only experiments. All experiments utilized the same basic procedures but differed according to the information received by subjects. We will explain these informational regimes shortly.

The experiments were run using inexperienced undergraduate subjects at New York University who had no previous experience in social learning experiments. For each treatment, forty subjects were recruited from undergraduate economics classes at New York University. In any experiment eight subjects were recruited. After subjects read the instructions (see Appendix) they were also read aloud by an experimental administrator. The experiment lasted for about one and one-half hours. A $5 participation fee and subsequent earnings for correct decisions, were paid in private at the end of the session. Throughout the experiment, we assured anonymity and an effective isolation of subjects[3] in order to minimize any interpersonal factors that may cause a superfluous tendency towards uniform 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

[3] Participants' working-stations were isolated by cubicles making it impossible for participants to observe others' 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.

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.

In any of the experiments described below, upon being called to participate and before being informed of her private signal, a subject first received some information relevant to her decision-making i.e., either the action of the previous subject, her advice or both depending on the treatment. After receiving this information, 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 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. Everyone whose decision determines their action as A earned \$2 if the sum of the subjects private signals 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.

As mentioned above, the experiments differ by what type of information subjects are offered before they are allowed to state their cutoff values. In the Perfect-Information experiment (Çelen and Kariv (2002b)), each action was announced publicly, and thus was known to all successors. For example, the fifth subject to choose was informed what action the first, the second, the third and the fourth have taken. In the Action-Only experiment (Çelen and Kariv (2002c)), subjects were able to observe only the action taken by their immediate predecessor. For example, the fifth to choose was informed only what action the fourth participant has taken.

Two experiments containing advice were run. In one, what we have already called the Advice-Only experiment, when subjects were called upon to make their decision they were able to observe none of the actions taken by their predecessors. Rather, they received advice from their immediate predecessor as to what they felt was the correct action to take. In the other, the Action-Plus-Advice experiment, subjects were able not only to receive advice from their immediate predecessor, but also observed the action taken by her. In both cases, the predecessors gave the advice after the computer recorded the action for them according to their cutoff and after they observed their private signal. In addition, in the experiments containing advice, everyone earned $1 if her successor took the correct action. This was paid to insure that the advice subjects give is their best guess as to what the right action is. Table 1 summarizes our experimental design and procedures.

[Table 1 here]

4 Theory

In this section, we discuss the theoretical implications of the model tested in the laboratory[4]. One of the main goals of this section is to demonstrate that when compared with the Action-Only case, in the Advice-Only case advice cannot convey more information. This theoretical fact gives the advice puzzle its bite.

To formulate the Bayesian solution of the decision problem underlying our experimental design, suppose that the eight agents receive private signals $\theta_1, \theta_2, .., \theta_8$ that are independently and uniformly distributed on $[-1, 1]$[5]. Sequentially, each agent $n \in \{1, ..., 8\}$ has to make a binary irreversible decision $x_n \in \{A, B\}$ where action A is profitable if and only if $\sum_{i=1}^{8} \theta_i \geq 0$, and action B otherwise.

In what follows, we will first discuss the theory for the Action-Only experiment that constitutes backbone of all three experiments. After we do this, we comment on the Advice-Only and Action-Plus-Advice experiments in order to demonstrate their connection.

4.1 Action-Only

In the Action-Only case, except the first agent, everyone observes only her immediate predecessor's decisions. In such a situation, conditional on the information available to her, agent n's optimal decision rule is

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

and since individuals do not know any of their successors' actions,

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

[4] Çelen and Kariv (2002c) provides a detailed analysis of the Action-Only model and Çelen and Kariv (2001a) study a general version of the model.

[5] For expository ease, we normalize the signal space to $[-1, 1]$.

It readily follows that the optimal decision 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

$$\hat{\theta}_n = -\mathbb{E}\left[\sum_{i=1}^{n-1} \theta_i \mid x_{n-1}\right] \tag{2}$$

is the optimal cutoff which accumulates all the information revealed to individual n from her predecessor's action.

We proceed by illustrating the basic features of the decision problem. The first agent's decision is based solely on her private signal. Therefore, her optimal cutoff is $\hat{\theta}_1 = 0$ which determines that it is optimal for her to take action A if and only if $\theta_1 \geq 0$ and action B otherwise. Since the second agent observes the first's action, she conditions her decision on whether $x_1 = A$ or $x_1 = B$. Thus, according to (2) the second agent's cutoff rule is

$$\hat{\theta}_2 = \begin{cases} -\frac{1}{2} & \text{if} \quad x_1 = A, \\ \frac{1}{2} & \text{if} \quad x_1 = B. \end{cases} \tag{3}$$

By the time it is the third agent's turn to make a decision, the information inherent in the first agent's action is suppressed, but she can still draw a probabilistic conclusion about it by Bayes' rule. Thus, according to (2) the second agent's cutoff rule is

$$\hat{\theta}_3 = \begin{cases} -\frac{5}{8} & \text{if} \quad x_2 = A, \\ \frac{5}{8} & \text{if} \quad x_2 = B, \end{cases} \tag{4}$$

Proceeding with the example by adding agents who receive private signals and learn only from preceding agent's action, the cutoff rule, $\hat{\theta}_n$, of any agent n can take the two different values conditional on whether agent $(n-1)$ took action A or action B which we denote by

$$\begin{aligned} \overline{\theta}_n &= -\mathbb{E}\left[\sum_{i=1}^{n-1} \theta_i \mid x_{n-1} = A\right], \\ \underline{\theta}_n &= -\mathbb{E}\left[\sum_{i=1}^{n-1} \theta_i \mid x_{n-1} = B\right]. \end{aligned} \tag{5}$$

In Çelen and Kariv (2002a), we show that the dynamics of the cutoff rule $\hat{\theta}_n$ is described in a closed form solution recursively as follows

$$\hat{\theta}_n = \begin{cases} -\frac{1+\hat{\theta}_{n-1}^2}{2} & \text{if} \quad x_{n-1} = A, \\ \frac{1+\hat{\theta}_{n-1}^2}{2} & \text{if} \quad x_{n-1} = B, \end{cases} \tag{6}$$

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

It follows immediately from (6) that the cutoff rule partitions the signal space into three subsets: $[-1, \overline{\theta}_n), [\overline{\theta}_n, \underline{\theta}_n)$ and $[\underline{\theta}_n, 1]$. For high-value signals $\theta_n \in [\underline{\theta}_n, 1]$ and symmetric low-value signals $\theta_n \in [-1, \overline{\theta}_n)$ agent n follows her private signal and takes action A or B respectively. In the intermediate subset $[\overline{\theta}_n, \underline{\theta}_n)$, which we call an *imitation set*, private signals are ignored in making a decision and agents imitate their immediate predecessor's action. Furthermore, since $\{\overline{\theta}_n\}$ and $\{\underline{\theta}_n\}$ converge respectively to -1 and 1, imitation sets monotonically increase in n regardless of the actual history of actions, and thus, over time, it is more likely that imitation will arise.

4.2 Advice-Only

Next, we shall investigate the differences between the decision problem underlying our Action-Only and Advice-Only experiments. Our purpose is to demonstrate that advice cannot convey more information, i.e., the advice given is not more informative than the action observed. In fact, the two are generally informationally identical.

It readily follows that when agents believe that the advice given to them by their predecessor is identical to their action, the unique equilibrium in the Action-Only case prevails in the Advice-Only case. That is to say that with a consistent belief system, agent n's optimal decision takes the form of the cutoff strategy, (1) and (2), the optimal advice rule is to give advice equal to her chosen action. In other words, conditional on the piece of advice given to her, A or B, agent n takes an action by setting her cutoff at $\hat{\theta}_n$, which is the optimal cutoff that accumulates all the information revealed to individual n from her predecessor's advice, and advises her successor to take action A if and only if $\theta_n \geq \hat{\theta}_n$ and to take action B otherwise.

It is straightforward to verify however that this equilibrium is not the unique equilibrium in the Advice-Only case. In fact, there are two other equilibria. The first equilibrium is the mirror equilibrium to the one described above. That is, agents believe that the advice given to them by their predecessor is opposite to their chosen action, set their cutoffs optimally given their beliefs and advise their successor to take the opposite action as well[6]. Clearly, this equilibrium and the unique equilibrium in the Action-Only case are equally informative. The second is the babbling equi-

[6] To clarify, everyone who is advised by her predecessor to take action $A(B)$ believes that the action she actually took was $B(A)$ and thus sets her cutoff at $\underline{\theta}_n(\overline{\theta}_n)$ instead of $\overline{\theta}_n(\underline{\theta}_n)$. Then, everyone advises her successor to take action $A(B)$ if the action she herself took was $B(A)$.

librium in which the advice giving rule is noisy, in the sense that agents choose their advice randomly as A or B with equal chance, and agents ignore advice and make decisions solely on the basis of private information, by simply setting cutoffs at zero. In this equilibrium, no information is conveyed.

Throughout the paper, whenever we refer to the theoretical sequence of cutoffs we consider the unique equilibrium cutoffs in the Action-Only case as given by (6).

4.3 Action-Plus-Advice

In the Action-Plus-Advice game, agents are able not only to receive advice from their immediate predecessor, but also observe the action taken by her, which opens up signaling possibilities. This fact enables agents to engage in a more sophisticated and hence informationally richer strategy by combining all four available action-advice pairs (action A, advice B; action B, advice A, etc.) to partition their signal space into four regions and convey more information to their successors about their signals. Hence the informational pipeline is less constrained in this case and more informationally rich equilibria exist than in the Action-Only and Advice-Only games.

It is still the case, however, that the equilibrium in which the agents simply advise their successors to do as they do is an equilibrium. In other words, the unique equilibrium in the Action-Only case prevails also in the Advice-Plus-Action case. In particular, when a convention exists such that agents ignore conflicting advice and take decisions solely on the basis of the action observed, the resulting equilibrium is, of course, our familiar equilibrium to the Advice-Only game.

5 Results

We will organize the presentation of our results by answering the four questions described in Section 2.

5.1 Question 1

Do subjects tend to follow advice more often than action when each is observed under identical circumstances?

In short the answer to Question 1 is yes. Subjects tend to place more confidence in advice than action. This is our first instance of words speaking louder than actions. We will try to convince you of this answer in several ways.

First, we define decisions made by subjects as *concurring decisions* if the sign of their cutoff agrees with the action taken (advice observed). For instance, when a subject observes that her predecessor took action A (B) (gave advice A (B)) and adopts a negative (positive) cutoff, she demonstrates concurrence, since by selecting a negative (positive) cutoff she adopts a higher probability of taking action A (B). Similarly, if a subject observes action (advice) A (B) and selects a positive (negative) cutoff, then she disagrees with her predecessor. We say that such decisions are *contrary decisions.* Finally, *neutral decisions* are carried out by choosing a zero cutoff, which neither agrees nor disagrees with the predecessor's action (advice) but simply entreats choice based on private information.

Given this distinction we can easily see that subjects tend to follow the advice of their predecessor far more than they tend to copy their action. This is clearly seen in Table 2 which presents the percentages of choices made in our various experiments that were concurring, contrary and neutral.

Table 2: Agreement and Contrariness in Action-Only and Advice-Only Experiments

	Concurring	**Neutral**	**Contrary**
Action-Only	44.2%	16.6%	39.2%
Advice-Only	74.1%	9.1%	16.8%

As we can see, advice is followed far more than action. Over all decision turns, excluding the first, while subjects tend to set a cutoff consistent with the advice they receive 74.1 percent of the time in the Advice-Only experiment, they do so only 44.2 percent of the time in the Action-Only experiment. Added together with the neutral cutoffs, subjects tend to weakly agree, i.e., set an concurring or neutral cutoff, with advice 83.2 percent of the time in the Advice-Only experiment while they do so only 60.8 percent of the time in the Action-Only experiment. These two distributions are significantly different using a Kolmogorov-Smirnov test (p-value 0.000).

While Table 2 presents data on the number of decisions that were concurring, neutral or contrary, Figure 1 summarizes the percent of subjects who disagreed with the observed action in less than two rounds, three to five rounds and so on. What we see is that when subjects in the Advice-Only experiment disagreed with the advice they were given, 67.5 percent of the time they did so less than twice, while subjects in the Action-Only experiment, tended to disagree far more often – only 20.0 percent of the subjects disagreed two times or less and 40.0 percent of the subjects disagreed with the action they observed between six to eight times while that percentage was only 7.5 in the Advice-Only experiment. These two distributions are significantly different using a Kolmogorov-Smirnov test (p-value 0.000).

[Figure 1 here]

The signs of the cutoffs as indicating of agreement or disagreement tells only part of the story as it ignores the strength of this agreement or disagreement, which can be measured by the magnitude of the cutoff set. For example, if a subject observes action (advice) A and sets a cutoff close to -10, then not only does she agree with the action (advice) she observed, but she does so very strongly since she will almost surely take action A. In contrast, selection of a negative cutoff that is closer to zero clearly indicates a much weaker agreement.

Since the cutoff strategy is symmetric around zero, in the sense that the strength of agreement or disagreement is independent of the actual action observed[7], we proceed by transforming the data generated by our subjects in the following way: Take the absolute value of cutoffs in concurring decision points and minus the absolute value of cutoffs at contrary decision points. For instance, if a subject observes action (advice) A and selects a cutoff of -5, we take it as 5, since she acts in a concurring manner. On the other hand, if she places a cutoff of 5 we take it as -5, since she acts in a contrary manner. In the remainder of the paper we will refer to this as *mirror image transformation*. Figure 2 presents the theoretical cutoffs in the unique equilibrium in the Action-Only case and the mean cutoff (after mirror image transformation) of concurring decisions turn by turn for our Advice-Only and Action-Only experiments.

[Figure 2 here]

As you can see, there is little difference in the magnitude of the cutoffs set by subjects when they strictly agreed with either the advice offered or action observed by their predecessor. In other words, once a subject has decided to follow or imitate the advice offered or action taken, she does so with

[7] To make sure that there was no bias towards any of the actions A or B, we ran $y = \alpha D_A + \beta D_B + \varepsilon$, where y is the vector of reported cutoffs and D_x is the dummy variable which takes value of 1 when the action (advice) observed is x. We fail to reject the hypothesis $\alpha + \beta = 0$.

equal intensity[8]. Note also that there is a substantial degree of conformity with the theory in the magnitude of the cutoffs chosen by subjects when they agreed with the action (advice) observed. Figure 3 shows however that the situation reverses in the Action-Only experiment, particularly in late decision-turns, when we include neutral decisions in our sample[9].

[Figure 3 here]

So far, we focused on concurring decisions. There is, however, the complement subset of contrary decisions. Notice that once a subject decides not to follow her predecessor's action (advice), the intensity of her disagreement can be measured in several ways. Figure 4 presents the intensity of disagreement in two ways. First, we use the absolute value of the distance between the cutoff actually chosen and the one which would be selected if the subject acted according to the theoretical cutoff rule given by (6), and, second, by the absolute value of the distance of the chosen cutoff from zero. Note that the strength of disagreement is rather severe since when subjects disagree with their predecessor they tend to do so in quite an extreme way[10].

[Figure 4 here]

All of the results presented above condition our data on whether decisions are concurring or contrary. Figure 5 shows that if we do not condition the data on agreement and disagreement, it appears that overall there is a significant difference between the Action-Only and Advice-Only cutoffs. Specifically, we find that any difference in behavior is the result of the fact that far more subjects follow advice than imitate action. When advice is followed or action imitated, however, it is done so with the same intensity.

[Figure 5 here]

The regression analysis presented in Table 3 summarizes our discussion so far. We regress the mirror image transformation of the cutoff set by subjects on the decision turn at which this cutoff was set as well as a dummy

[8] A set of Two-sample Wilcoxon rank-sum (Mann-Whitney) tests run decision turn by decision turn detect no significant difference between the strongly agreeing cutoffs set in the Advice-Only and Action-Only experiments for any decision turn.

[9] A set of Wilcoxon tests run decision turn by turn detects a significant difference between the weakly agreeing cutoffs set in the Advice-Only and Action-Only experiments for decision turns 6 (p-value 0.009) and 7 (p-value 0.002).

[10] Wilcoxon tests detect a significant difference only for decision turns 4 (p-value 0.095) and 8 (p-value 0.046).

variable which takes a value of 1 if the experiment containing the observation was the Advice-Only experiment and 0 if the experiment was the Action-Only experiment. As we can see from Table 3, the experiment dummy is highly significant and positive indicating that for any decision turn the cutoffs set by a subject in the Advice-Only experiment can be expected to be 3.05 units higher than a cutoff set under identical circumstances in the Action-Only experiment indicating more confidence in advice than action. The significance of the dummy variable in the regression clearly indicates that the process of setting cutoffs is different in the Advice-Only and Action-Only experiments and that this difference is consistent with our observations above where we indicate that subjects are more persuaded by offered advice than observed action[11].

Table 3: Cutoff Behavior in Advice-Only and Action-Only Experiments

	Coefficient	**Std. Err.**	**t**	**p-value**
Constant	-0.51	0.475	-1.064	0.288
D(Advice-Only)	3.05	0.336	9.093	0.000
Turn 3	0.74	0.628	1.183	0.237
Turn 4	1.62	0.628	2.573	0.010
Turn 5	1.67	0.628	2.661	0.008
Turn 6	2.02	0.628	3.218	0.001
Turn 7	1.75	0.628	2.835	0.005
Turn 8	2.12	0.628	3.387	0.001

Obs.= 1050, $R^2 = .088$

5.2 Question 2

Do subjects in the Advice-Only or Action-Only experiments act in a manner that is more consistent with the Bayesian behavior for the game played? In other words, are their cutoffs in Advice-Only experiment closer to the cutoff prescribed by the theory than those of subjects in the Action-Only experiment.

It has been observed before in the work of Schotter and Sopher (2001, 2002, 2003) on inter-generational games, that advice appears to be rationality enhancing. That is, games played with advice are played in a manner

[11] GLS random-effects estimators and robust variance estimators for independent data and clustered data yield similar results.

which more closely adheres to the predictions of economic theory than do games in which subjects can only observe the history (or some subset of the history) of interactions before them, but unable to receive advice. The same appears to be true in the games studied here.

To demonstrate this conclusion, consider any of our subjects who have engaged in one of our experiments for 15 rounds. In each round, the subject was randomly assigned a decision turn so for any subject we have data on how she set her cutoff conditional on the decision turn and either the advice she was given or the action she observed, depending on the experiment she participated in. For each such situation, we also know the theoretical cutoff. Hence, our data provide us with a vector for each subject indicating the cutoff chosen at each of the decision turns while our theory provides us with an equilibrium cutoff for that situation.

More precisely, if we let $\tilde{\theta}_{i,r} = (\tilde{\theta}_{i,1}, ..., \tilde{\theta}_{i,15})$ be the vector of actual cutoffs chosen by subject i and $\hat{\theta}_{i,r} = (\hat{\theta}_{i,1}, ..., \hat{\theta}_{i,15})$ equilibrium cutoff for those situations faced by this subject, we can use the mean deviation (MD) for subject i,

$$MD_i = \frac{1}{15}\Sigma_{r=1}^{15}\left|\hat{\theta}_{i,r} - \tilde{\theta}_{i,r}\right|$$

as a goodness-of-fit measure between the cutoffs set by subject i and those prescribed by the theory. The smaller the mean MD for subjects in any experiment the closer is their behavior to that predicted by the theory. Figure 6 presents these MD calculations for each experiment in the form of histograms. The horizontal axis is the intervals of the MD scores and the vertical axis is the percentages of the subjects corresponding to this MD scores.

[Figure 6 here]

Figure 6 presents dramatic evidence that subject behavior in our Advice-Only experiment is, ironically, more consistent with the unique equilibrium of the Action-Only game than was the Action-Only data itself, i.e., the distribution of MD scores for the Advice-Only experiment was shifted to the left when compared to the Action-Only experiment. A Kolmogorov-Smirnov test confirms this observation by indicating that these two distributions of MD scores are significantly different from each other (p-value 0.000).

5.3 Question 3

Does the behavior in the Action-Plus-Advice experiment change from what it was in the Advice-Only experiment? Which infor-

mation, advice or action, is more valued by the subjects? And, under what circumstances do subjects offer advice, which differs from their action?

The Action-Plus-Advice experiment provides us with an extremely good opportunity to try to separate the impact of advice and action on behavior. The reason is that in a number of situations, subjects were faced with advice that was different from the action taken by the subject in the previous round. For example, in the Action-Plus-Advice experiment 83 out of the 525 decisions excluding the first decision turn (15.8 percent) were made under circumstances where the advice offered was different from the action observed in the previous period. If when these situations occurred, subjects chose to follow the advice of their predecessors rather than copying their action, we would interpret this as indicating that advice was more influential than action.

We investigate the differential impact of advice on behavior in several ways. First, consider the choice of a negative cutoff as indicating a preference for the A choice and the choice of a positive cutoff as a preference for the B choice. If the advice and action of a predecessor subject differ, then two cases can be observed. The predecessor chooses A and advises B or the predecessor chooses B and advises A. Based on either of these occurring, the successor subject could choose to set either a negative cutoff (a higher probability of taking action A) or a positive one (a higher probability of taking action B). This defines four contingencies as depicted in Table 4.

Table 4: Advice Taking in the Action-Plus-Advice Experiment

Successor / *Predecessor*	**Choose A Cutoff (-)**	**Choose B Cutoff (+)**	**Cutoff = 0**
Action A/Advice B	13 (15.66%)	33 (39.76%)	6 (7.23%)
Action B/Advice A	17 (20.48%)	7 (8.43%)	7 (8.43%)

As you can see, when the advice and action of one's predecessor differ, successors are far more likely to choose an action consistent with the received advice than the observed action. For example, in 60.2 percent of the cases where the advice offered differs from the action, subjects chose to follow the advice they received rather than imitate their predecessor's action while only 24.1 percent of the time they imitated the action taken and 15.7 percent of the time they were neutral and choose a cutoff zero.

Table 4 looks at behavior when the advice offered by a subject's predecessor differs from the action she took. But we might also ask whether getting advice that is consistent with the action taken by one's predecessor makes a subject more likely to follow it and if so more likely to set a more extreme cutoff indicating stronger agreement. A priori we would expect this to be the case since when advice agrees with a predecessors' action we should expect a subject to view it as more compelling.

What we find are mixed results. First, as illustrated in Table 5 it is true that subjects are more likely to follow advice (as indicated by the sign of their cutoff) when it is backed up by action. Note that if a subject is told to follow an action by a predecessor who took that action himself, such a recommendation is followed 84.2 percent of the time while such advice is followed only 74.1 percent of the time in the Advice-Only experiment. When just the action is observed, it is imitated only 44.2 percent of the time. So it should be clear that a predecessor who does as she says is seen as being more believable than one whose advice cannot be backed up by action.

Table 5: Decision Conformity With Advice and Action

	Action Taken		
	Concurring	**Neutral**	**Contrary**
Action-Only	44.2%	16.6%	39.2%
Advice-Only	74.1%	9.1%	16.8%
Action-Plus-Advice	84.2%	7.0%	8.8%

In addition, Figure 7 shows that when subjects are in an Action-Plus-Advice experiment and receive advice which is consistent with the action they observe, if they act in a contrary manner they tend to do so less often than they do in either the Action-Only or Advice-Only experiments. In fact, as we see in Figure 7, when they act in a contrary manner 80.0 percent of the time they make two or less contrary decisions indicating that most subjects follow advice backed up by action most of the time.

[Figure 7 here]

On the other hand, as Figure 8 indicates, when subjects set a cutoff conditional on receiving advice that is consistent with the action they observe, the magnitude of the cutoffs set does not differ very much from the cutoffs set by subjects in the Advice-Only experiment[12] (who were not able to see

[12] A set Wilcoxon tests detect no significant difference between the distributions of the cutoffs set in the Advice-Only and Action-Plus-Advice experiments for any turn.

if the advice offered was consistent with the actions taken by their predecessors). So again, we see that the impact of advice (this time backed up by action) is to increase the fraction of time recommendations are followed but, once they are followed, the strength of their commitment to the decision is practically identical.

[Figure 8 here]

Perhaps the best way to summarize our results here is to present the results of two simple regressions run to explain the cutoff behavior of subjects after they either view the action of the predecessor (Action-Only), receive advice (Advice-Only) or both receive advice and view the action of the predecessor (Action-Plus-Advice). It was our initial hypothesis that in any decision turn the cutoff set by a subject would be greatest (after mirror image transformation) in the Action-Plus-Advice experiment after receiving advice that was consistent with the action observed, i.e., after seeing a predecessor choose A (B) and then being told to choose A (B). Second highest after being advised to choose an action in the Advice-Only experiment. Third highest after observing an action in the Action-Only experiment, and smallest when one's predecessor chooses one action but advises another in the Action-Plus-Advice experiment.

To test this hypothesis we ran the following two regressions. We pooled our data and separated it into two sets. In one set we pooled the data from the Action-Only experiment as well as the data from the Action-Plus-Advice experiment, and in the other set we did the same thing for the Advice-Only and Action-Plus-Advice experiments. We then ran two regressions. In the first, we regressed the cutoff set by our subject on three dummy variables depicting whether the observation came from the Action-Only experiment, the Action-Plus-Advice experiment where advice was consistent with action or the Action-Plus-Advice experiment with inconsistent action and advice. Using the Action-Only dummy as the baseline, only two dummies were coded. We ran the same regression using the other data set and including the Advice-Only experiment as the baseline. The other right hand variables were the decision turn dummies.

The results of these regressions confirm our hypotheses. In brief, at any decision turn subjects tend to set the highest cutoff when they are in the Action-Plus-Advice experiment and receive advice which is consistent with the action they just observed. Their second highest cutoff is when they receive advice in the Advice-Only experiment. Third highest is when they view an action in the Action-Only experiment and finally they tend to set the lowest cutoff when they get conflicting advice from the action they observe. These results are presented in Tables 6a and 6b[13]. In both tables, dummy variable D_1 depicts an observation where the advice-action pair is either AA or BB while D_2 depicts an observation where the advice-action pair is either AB or BA. In Table 6a the constant term is associated with an observation coming from the Action-Only experiment while in Table 6b the observation comes from the Advice-Only experiment.

TABLE 6A: CUTOFF BEHAVIOR IN ACTION-ONLY AND ACTION-PLUS-ADVICE EXPERIMENTS

	Coefficient	**Std. Err.**	**t**	**p-value**
Constant (Action)	0.042	0.048	0.870	0.384
D$_1$(AA or BB)	0.347	0.042	8.157	0.000
D$_2$(AB or BA)	-0.370	0.065	-5.611	0.000
Turn 3	0.045	0.058	0.773	0.440
Turn 4	0.023	0.059	0.391	0.696
Turn 5	0.058	0.059	0.994	0.320
Turn 6	0.100	0.059	1.707	0.088
Turn 7	0.118	0.059	1.996	0.046
Turn 8	0.033	0.059	0.561	0.575

Obs. = 1050, R^2 = .151

TABLE 6B: CUTOFF BEHAVIOR IN ADVICE-ONLY AND ACTION-PLUS-ADVICE EXPERIMENTS

	Coefficient	**Std. Err.**	**t**	**p-value**
Constant (Advice)	0.257	0.056	4.526	0.000
D$_1$(AA or BB)	0.045	0.067	0.670	0.503
D$_2$(AB or BA)	-0.214	0.079	-2.707	0.007
turn 3	0.060	0.050	1.234	0.217
turn 4	0.097	0.049	1.965	0.049
turn 5	0.161	0.049	3.293	0.001
turn 6	0.218	0.049	4.481	0.000
turn 7	0.240	0.049	4.860	0.000
turn 8	0.228	0.049	4.653	0.000

Obs. = 1050, R^2 = .030

[13] GLS random-effects estimators and robust variance estimators for independent data and clustered data yield similar results.

Inspection of Tables 6a and 6b provides support for our hypothesis. For example, looking at Table 6a, first note that a subject's cutoff increases when she receives advice that is consistent with the action observed while it decreases if the action and advice disagree. In other words, mixed advice makes a subject less certain as to what to do and lowers the cutoff she employs. Note, however, that the cutoff does not increase as the decision turn increases since the coefficient on each decision turn coefficient is not significant. The results are slightly different in the regression presented in Table 6b that uses Advice-Only experiment as the baseline. Here, note that Advice-Only subjects are so confident about their cutoff that observing an action consistent with that advice in the Action-Plus-Advice experiment has no significant impact on their cutoff (D_1 is not significantly different from zero). Seeing mixed Advice-Only would lower their cutoff, however. In addition, the decision turn does have a significant impact of the cutoff level from decision turn 4 onward. Finally, looking at absolute levels we see that the regressions imply that cutoffs are highest in situations where advice and action are the same in Action-Plus-Advice, second highest in Advice-Only, third highest in Action-Only and finally lowest where advice and action are not the same in Action-Plus-Advice.

We next turn our attention the following question.

Under what circumstances do people offer advice which differs from their action?

It is obviously of interest to ask why a positive fraction of our subjects offer advice different from the action taken and under what circumstances this occurred. We conjectured that subjects would overturn their action (i.e., offer advice that differs from the action taken) when they set extreme cutoffs (i.e., cutoffs relatively close to -10 or 10) and the signal they observe is consistent with their beliefs but very close to their cutoff. Let us explain our reasoning here.

Say a subject is relatively certain that the sum of the signals is negative so that B is the correct decision. Under those assumptions, say she sets a cutoff of 8.5. This means that for any signal below 8.5, she would like to choose action B. Such a strong cutoff clearly indicates a strong belief that the sum of all signals is negative and that B is the profitable action. If the signal she receives is below 8.5 and very negative, say -5, she feels pretty safe in her belief and happy to have B chosen for her since her signal was such a strong confirmation of her prior belief. For the same reason, she would also be happy to offer B as advice. We will call a signal that is below a positive cutoff or above a negative cutoff a consistent signal since it confirms the subject's belief about the true state of the world in the sense that if gives evidence that the state she already believes is more likely to occur.

However, what if her signal is 8.49? Here, the signal is still consistent with her belief, i.e., she will still choose action B, but its magnitude shakes her faith in her prior. It is exactly under such circumstances that we expect to find that subjects offer advice which is opposite of the action they took. If their cutoff was not extreme, then we do not expect overturns since non-extreme cutoffs indicate lack of strength in a subject's belief about the state of nature. Hence, any realization of the signal is not likely to cause the subject to overturn it in her advice giving. To summarize, we expect that subject overturns occur when the signal observed is consistent but marginally so. If the signal were marginally inconsistent, for example 8.53 in our previous example, we expect subjects would tend not to overturn the A action chosen for them by the computer, at least according to our data.

Our data largely support these conjectures with one noticeable exception. To begin, overturns are relatively rare occurring only 17.5 percent of the times in the Advice-Only experiment and 15.8 percent of the times in the Action-Plus-Advice experiment. In addition, as Figure 9 illustrates it is also rare that any subject will overturn many outcomes even if she overturns some. For example, in the Advice-Only and Action-Plus-Advice experiments 65.0 and 67.5 percents respectively of our subjects, if they ever overturned, offered advice that overturned two or fewer of the actions they took out of 15 trials. Only 10.0 percent and 5.0 percent overturned six or more outcomes in the Advice-Only or Action-Plus-Advice experiments respectively. In short, overturns are rare and infrequently done by any given subject[14].

[Figure 9 here]

To investigate this conjecture we specified a Logit model in which the left hand variable is a binary variable which took a value of 1 when there was an overturn and 0 otherwise. The right hand variables consisted of the decision turn, **Turn**, the absolute value of the cutoff set by the subject, **Abs**, the distance between the absolute value of the cutoff and the mirror of the signal received, **Dst**. We also entered an interaction term, **Product**, consisting of

[14] In addition, the number of overturns is insensitive to the decision turn in that when a regression of the number of overturns on decision turn is run in both the Action-Plus-Advice and Advice-Only experiments, the slope coefficients are not significantly different from zero.

the distance and a dummy that take a value 1 if the signal and the cutoff have the same sign and the signal is below (above) a negative (positive) cutoff. We also entered a dummy, **AcAd**, for the experiment that takes the value 1 for the Action-Plus-Advice experiment and 0 for the Advice-Only experiment. The results are presented in Table 7.

Table 7: Overturning Behavior

	Odds Ratio	**Std. Err.**	**Z**	**p-value**
AcAd	0.973	0.169	-0.156	0.876
Turn	0.990	0.044	-0.216	0.829
Abs	1.093	0.034	2.912	0.004
Dst	0.834	0.019	-7.957	0.000
Product	1.070	0.049	1.487	0.137

Log likelihood = -429.34084, Pseudo R^2= 0.0925

We can see the model substantiates our conjecture. The probability of an overturn is increasing in how extreme the cutoff set by the subject (**Abs**), and decreasing in the distance between the cutoff set and the signal received (**Dst**). We see no experiment effect in the sense that the overturn behavior does not seem to be affected by whether we look at the Advice-Only or the Action-Plus-Advice experiment (**AcAd**). Finally, the interaction (**Product)** term is insignificant. Still, our original conjecture about when we expect overturns to occur does seem to be correct on this data set – overturns occur when the cutoff is high and the signal is close to it.

5.4 Question 4

Are there more or fewer herds and cascades in experiments with advice? Does the welfare of subjects increase when they have access to advice?

5.4.1 Payoffs and efficiency

Probably the most important question that we can ask about the impact of advice on social learning is whether the presence of advice increases the welfare of subjects over and above what it would be without it. In answering this question, we will have to examine the impact that advice has on herding and cascade behavior of subjects since one way that advice affects behavior is through its propensity to cause subjects to herd with greater frequency than they would in its absence. To begin, consider Table 8, which presents a summary our four experiments, it is clear that the mean payoffs of our subjects was highest in those experiments where advice was present.

Table 8: Summary of The Experiments

	Action-Only	Advice-Only	Action-Advice	Perfect-Information
Earnings	\$18.8	\$21.8	\$23.0	\$22.0
Herds*	8	25	36	27
% of Herds†	10.7	33.3	48.0	36.0
Incorrect Herds	0	0	0	1
Cascades	18	24	21	26
% of Cascades†	24.0	32.0	28.0	34.7
Overturns	234	167	142	173
% of Overturns§	44.6	31.8	27.0	32.9

* Herds of at least five subjects.
† Out of all 75 rounds.
§ Out of all 525 decision points excluding the first decision turn.

As we see, while earnings for taking the correct action in the Action-Only experiment averaged \$18.8 they average \$23.3 and \$21.8 for the Action-Plus-Advice and Advice-Only experiments. These increases represent increases of 24.3 percent and 16.4 percent respectively. In the Perfect-Information experiments of Çelen and Kariv (2002b) where subjects could see the entire history of actions before setting their cutoff values (but did not receive advice), earnings averaged \$22.0 indicating that advice with imperfect information is approximately as efficient as perfect information without advice. A set of binary Wilcoxon tests indicates that there is a significant difference between the sample of subject payoffs in the Action-Only experiment and all other experiments at the 5 percent level of significance. It also indicates that no difference exists between the payoffs of subjects in the Perfect-Information experiment and any of those with advice, substantiating our conclusions that the presence of advice seems to be a substitute for the extra information contained in the perfect information experiment.

5.4.2 Herd behavior and informational cascades

One of the main reasons why advice increases the payoffs and hence the welfare of our subjects is that it has a dramatic impact on our subjects' inclination to herd. We identify a subject who engages in cascade behavior as one who reports a cutoff of -10 or 10, and thus takes either action A or B, no matter what private signal she receives. In contrast, a subject who joins a herd but does engage in cascade behavior is one whose cutoff is in the open interval $(-10, 10)$, indicating that there are some signals that can lead her to choose action A, some that lead to B but when her private signal is realized she will act as her predecessors did. Finally, we say that a cascade occurs in the laboratory when beginning with some subject, all others thereafter follow cascade behavior, and herd behavior occurs when, beginning with some subject, all take the same action.

Herd behavior Table 9 shows the rounds in which herds of at least five subjects arise in Advice-Only and Action-Plus-Advice experiments. Note the dramatic impact that advice has on the propensity of our subjects to herd. While in our Action-Only experiments we observed herding of at least five subjects in only 8 of the 75 rounds (10.7 percent), in the Advice-Only and Action-Plus-Advice sessions herding occurred in 25 (33.3 percent) and 36 (48.0 percent) rounds respectively[15]. Moreover, in the Action-Plus-Advice experiment herd behavior developed even more frequently than in the Perfect Information experiments where we found that herding was the outcome in 27 of the 75 rounds (36.0 percent). Finally, the frequency in which herd behavior occurs in the Action-Plus-Advice experiment compares favorably to the 47 percent predicted by the theory.

[Table 9 here]

Obviously two conditions must be met if advice is going to be welfare increasing. First, the advice must be correct and second it must be followed. Miraculously, in these experiments, both conditions seemed to have been met. In the Advice-Only experiments, whenever herd behavior arises all of the advice given was consistent with the action herded up on. In the Action-Plus-Advice experiments, this was not the case in only 5 of the 36 herds (rounds 7.8, 7.11, 8.13, 8.15, and 9.8). In other words, when herds occurred those who herded tended to follow the advice given. More remarkably, in all experiments all herds turned out to be on the correct decision. This result

[15] As we show in Table 13, out of the 8 rounds in which a herd arises in the Action-Only sessions, in two rounds all eight subjects acted alike, in one the last six subjects and in five rounds the last five subjects acted alike. In the Advice-Only sessions, in six rounds all eight subjects acted alike, in seven rounds the last seven subjects, in seven rounds the last six subjects and in five rounds the last five subjects acted alike. In the Advice-Plus-Action sessions, in ten rounds all eight subjects acted alike, in five rounds the last seven subjects, in eleven rounds the last six subjects and in ten rounds the last five subjects acted alike.

is of a particular interest since one of the original concerns of the social learning literature was that herds and cascades might support or reinforce inefficient choices. Following Anderson and Holt (1997), these fears were supported by the results of many laboratory experiments.

Since herding was so prevalent when advice was present the converse, subjects not following the action taken by their predecessor, must have occurred less frequently. This, of course, was true. For example, in the Action-Only experiment if we exclude the first decision turn, such non-herding behavior occurred in 234 of the 525 decisions points (39.0 percent) compared to 167 (31.8 percent) and 142 (27.0 percent) in the Advice-Only and Action-Plus-Advice experiments respectively. Theory predicts that we would expect such behavior to occur only 19.0 percent of the times (given the distribution of signals and optimal cutoffs) so the frequency of not following one's predecessor was certainly greater than predicted. To sum up, our results on herd behavior indicate that advice is a strong force in the creation of uniform social behavior and welfare increasing.

Informational cascades While all cascades must be herds, the opposite is certainly not true. Our experiment is uniquely designed to distinguish between the occurrence of cascades and herd since we are able to observe subjects' cutoffs that are typically unobservable. Surprisingly, advice did not have a significant impact on the rate of occurrence of information cascades. In the Action-Only experiments, cascades, in the sense that from some subject on all acted irrespective of the content of their private signals by setting either -10 or 10 as their cutoffs, were observed in 18 rounds (24.0 percent), whereas in the Advice-Only and Action-Plus-Advice experiments cascades formed in 24 (32.0 percent) and 21 (28.0 percent) rounds respectively[16]. Table 10 summarizes the rounds in which the longest information cascades occurred in the Advice-Only and Action-Plus-Advice experiments.

[Table 10 here]

[16]In the Action-Only experiments, only in 2 rounds the last two subjects and in 16 rounds only the last subject followed a cascade behavior. In the Advice-Only experiments, in 1 round the last three subject, in 4 rounds the last two subjects and in 19 rounds the last subject followed a cascade behavior. In the Action-Plus-Advice experiments, in 2 rounds the last four subjects followed a cascade behavior, in 1 round the last three subject, in 5 rounds the last two subjects and in 13 rounds the last subject followed a cascade behavior.

5.4.3 The information content of advice

Our final comment deals with the question of how informative advice is. As we have seen in our discussion of herds and cascades, it appears as if the herding behavior and efficiency in the Advice-Only experiment replicates the Perfect-Information experiments of Çelen and Kariv (2002b). This is remarkable in that while in one experiment a subject gets to see all of the outcomes of all those who have chosen before her, in the Advice-Only experiment a subject only gets to see advice from someone who received advice herself with no subject ever being able to see the outcome of any decision turn. However, we would expect that if subjects treated advice as sufficient statistics encoding all of the information contained in the outcome histories we would expect to see very similar results across these experiments.

To give some insight into this question we performed a very simple exercise. We look at the mean cutoff of the decisions in the Perfect-Information experiment of Çelen and Kariv (2002b) where all predecessors acted alike before a given decision turn. Since the cutoff strategy is also symmetric around zero under perfect information, we use the mean of the absolute values of the cutoffs. For comparison purposes we then calculated the mean cutoffs of our subjects in both the Advice-Only and Action-Only experiments at the same decision turn. If one piece of advice (or one action) was as reliable as these unanimous histories, then we could expect that the cutoffs set in the Advice-Only experiments would, on average, be equal to those set by the Perfect-Information subjects. Figure 10 demonstrates that this is very much the case. As we see, while there is a great similarity between the cutoff set by subjects in the Perfect-Information experiment of Çelen and Kariv (2002b) and our Advice-Only experiment in the sense that both sets of cutoffs tend to be high, the cutoffs set in the Action-Only experiment are not only low but close to zero[17]. Subjects seem to treat one piece of advice as equivalent to pristine histories.

[Figure 10 here]

6 Concluding remarks

This paper has demonstrated the dramatic impact that advice has on the process of social learning. It has raised a behavioral puzzle that we call the

[17] A set of Wilcoxon tests run turn by turn detects no difference between the cutoffs of subjects in the Advice-Only and Perfect-Information experiments while there is a difference with cutoffs in the Action-Only experiment.

"Advice Puzzle" which is that people appear to be far more willing to follow the advice of people who go before them than to copy their action. While this may seem appropriate since one might think that advice embodies the wisdom of one's past, in fact, from an informational point of view, amongst a set of Bayes-rational subjects the two are equivalent.

This willingness to follow the advice of predecessors considerably increases the occurrences of herds when a comparison is made between the herding behavior of subjects in experiments with and without advice. Most importantly, advice is welfare increasing. Subjects do better when advice is allowed and advice tends, on average, to be correct so that it is worth following. To repeat the subtitle of this paper, this paper presents an experiment in social learning where words speak louder than actions. It presents an experimental design in which social learning is truly social in the sense that people cannot only copy the actions of the past but also receive advice from those who took them.

While we think we have raised an interesting puzzle, we still think that much more work needs to be done in answering it. For example, why are people so willing to follow advice? Why is advice, on average, more correct than actions? Is advice a prerequisite for herding in informationally limited environments? That is, where subjects can only see the action of their immediate predecessor instead of all the actions of subjects before them?

We feel that these and other questions like them will need to be answered before we can begin to fully sort out the factors that help us solve the advice puzzle.

References

[1] Allsopp L. and J. Hey, "Two Experiments to Test a Model of Herd Behavior." *Experimental Economics*, June 2000, *3*(2), pp. 121-36.

[2] Anderson L., "Payoff Effects in Information Cascade Experiments." *Economic Inquiry*, October 2001, *39*(4), pp. 609-15.

[3] Anderson L. and C. Holt, "Classroom Games: Information Cascades." *Journal of Economic Perspectives*, Summer 1996, *10*(4), pp. 187-93.

[4] Anderson L. and C. Holt, "Information Cascades in the Laboratory." *American Economic Review,* December 1997, *87*(5), pp. 847-62.

[5] Banerjee A., "A Simple Model of Herd Behavior." *Quarterly Journal of Economics,* August 1992, *107*(3), pp. 797-817.

[6] Bikhchandani S., D. Hirshleifer and I. Welch, "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascade." *Journal of Political Economy,* October 1992, *100*(5), pp. 992-1026.

[7] Bikhchandani S., D. Hirshleifer and I. Welch, "Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades." *Journal of Economic Perspective,* Summer 1998, *12*(3), pp. 151-70.

[8] Chaudhuri A., A. Schotter and B. Sopher, "Talking Ourselves to Efficiency: Coordination in Inter-Generational Minimum Games with Private, Almost Common and Common Knowledge of Advice." Mimeo, New York University, 2002.

[9] Çelen B. and S. Kariv, "Observational Learning Under Imperfect Information." *Games and Economic Behavior*, forthcoming, 2002a.

[10] Çelen B. and S. Kariv, "Distinguishing Informational Cascades from Herd Behavior in the Laboratory." Mimeo, New York University, 2002b.

[11] Çelen B. and S. Kariv, "An Experimental Test of Observational Learning under Imperfect Information." Mimeo, New York University, 2002c.

[12] Duffy J. and N. Feltovich, "Do Actions Speak Louder Than Words? An Experimental Comparison of Observation and Cheap Talk." *Games and Economic Behavior*, April 2002, *39*(1), pp. 1-27.

[13] Gale D., "What Have We Learned from Social Learning?", *European Economic Review,* April 1996, *40*(3-5), pp. 617-28.

[14] Hung A. and C. Plott, "Information Cascades: Replication and an Extension to Majority Rule and Conformity-Rewarding Institutions." *American Economic Review,* December 2001, *91*(5), pp. 1508-20.

[15] Schotter A. and B. Sopher, "Social Learning and Coordination Conventions in Inter-Generational Games: An Experimental Study." *Journal of Political Economy,* June 2003, *111*(3), pp. 498-529.

[16] Schotter A. and B. Sopher, "Advice and Behavior in Intergenerational Ultimatum Games: An Experimental Approach." Mimeo, New York University, 2002.

[17] Schotter A. and B. Sopher, "Trust and Trustworthiness in Intergenerational Games: An Experimental Study of Inter-generational Advice." Mimeo, New York University, 2001

Table 1

Action-Only	Observe predecessor's action (*A* or *B*)	Select a cutoff in [-10,10]	Receive signal Action recorded	
Advice-Only	Receive predecessor's advice (*A* or *B*)	Select a cutoff in [-10,10]	Receive signal Action recorded	Give advice (*A* or *B*) to successor
Advice-Plus-Action	Observe predecessor's action (*A* or *B*) Receive predecessor's advice (*A* or *B*)	Select a cutoff in [-10,10]	Receive signal Action recorded	Give advice (*A* or *B*) to successor
Perfect-Information	Observe the entire history of actions	Select a cutoff in [-10,10]	Receive signal Action recorded	

Table 9

Advice-Only

Session. Round*	Action herded	Cutoff and advice by turn							
		1	2	3	4	5	6	7	8
1.15	*A*	0 *A*	-5 *A*	-4 *A*	-3 *A*	-9 *A*	-10 *A*	-4.8 *A*	0
2.10	*A*	0 *A*	2 *A*	-5 *A*	2 *A*	0 *A*	-10 *A*	-9 *A*	-10
4.14	*B*	0 *B*	1.25 *B*	4.3 *B*	5 *B*	7 *B*	5 *B*	10 *B*	8
4.15	*A*	0 *A*	-7 *B* (shaded)	0 *A*	-2 *A*	-10 *A*	-7 *A*	2.3 *A*	-9.9
5.8	*B*	0 *B*	5 *B*	1.57 *B*	5 *B*	5 *B*	10 *B*	5 *B*	8
5.12	*A*	0 *A*	0 *B* (shaded)	-3.3 *A*	-2 *A*	-2.56 *A*	-3 *A*	-9 *A*	-7
Average**		0.0	3.4	3.0	3.2	5.6	7.5	6.7	7.2

Action-&-Advice

Session. Round*	Action herded	Cutoff and advice by turn							
		1	2	3	4	5	6	7	8
6.4	*A*	0 *A*	-5.4 *A*	-10 *A*	5 *A*	-2 *A*	-10 *A*	-2 *A*	-7
7.8	*A*	0 *A*	-4 *A*	0 *A*	-10 *A*	2 *A*	-9 *A*	-5 *B* (shaded)	-2
7.12	*B*	0 *B*	4 *B*	2 *B*	7 *B*	10 *B*	7.5 *B*	10 *B*	8
8.13	*A*	-3 *A*	-6 *A*	0 *A*	-9.9 *A*	-2 *A*	-6 *B* (shaded)	-4 *B* (shaded)	0
9.1	*A*	0 *A*	0 *A*	1 *A*	-1 *A*	1 *A*	-1.3 *A*	-7 *A*	-5
9.4	*B*	0 *B*	7 *B*	7 *B*	-4 *B*	2.5 *B*	-0.1 *B*	2 *B*	8
9.8	*A*	-7 *A*	-8 *A*	1 *A*	-1 *A*	-9 *A*	-10 *A*	-8 *B* (shaded)	1
9.14	*B*	0 *B*	0 *B*	10 *B*	2 *B*	5 *B*	8 *B*	7 *B*	-0.1
10.11	*A*	0 *A*	-8 *A*	-6 *A*	-4.5 *A*	-10 *A*	-10 *A*	-4 *A*	-3
10.15	*A*	-10 *B* (shaded)	0 *A*	-10 *A*	-9 *A*	-4.5 *A*	-9 *A*	-3 *A*	-7
Average**		2.0	4.2	4.7	5.3	4.8	7.1	5.2	4.1

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

** Average of the cutoffs absolute values.

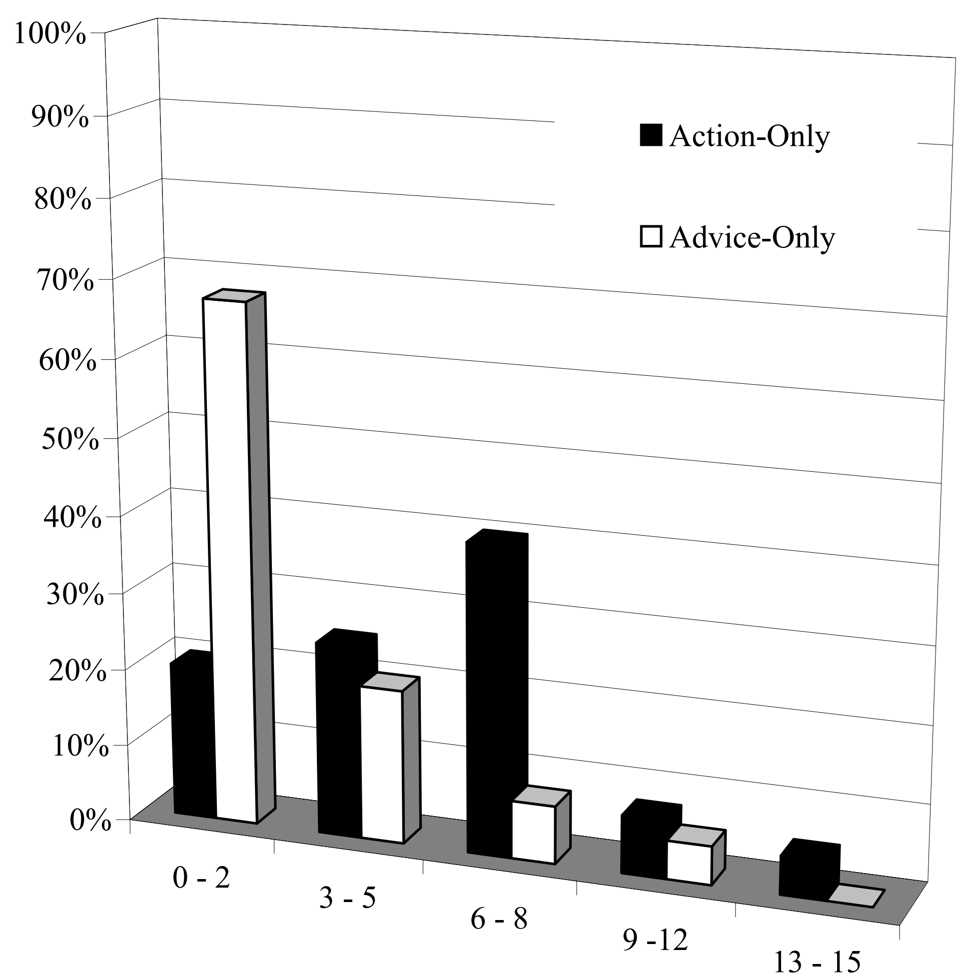
(shaded) Advice overturn.

Table 10

Data for the rounds with the longest informational cascades
Action:Advice
Cutoff
Private signal

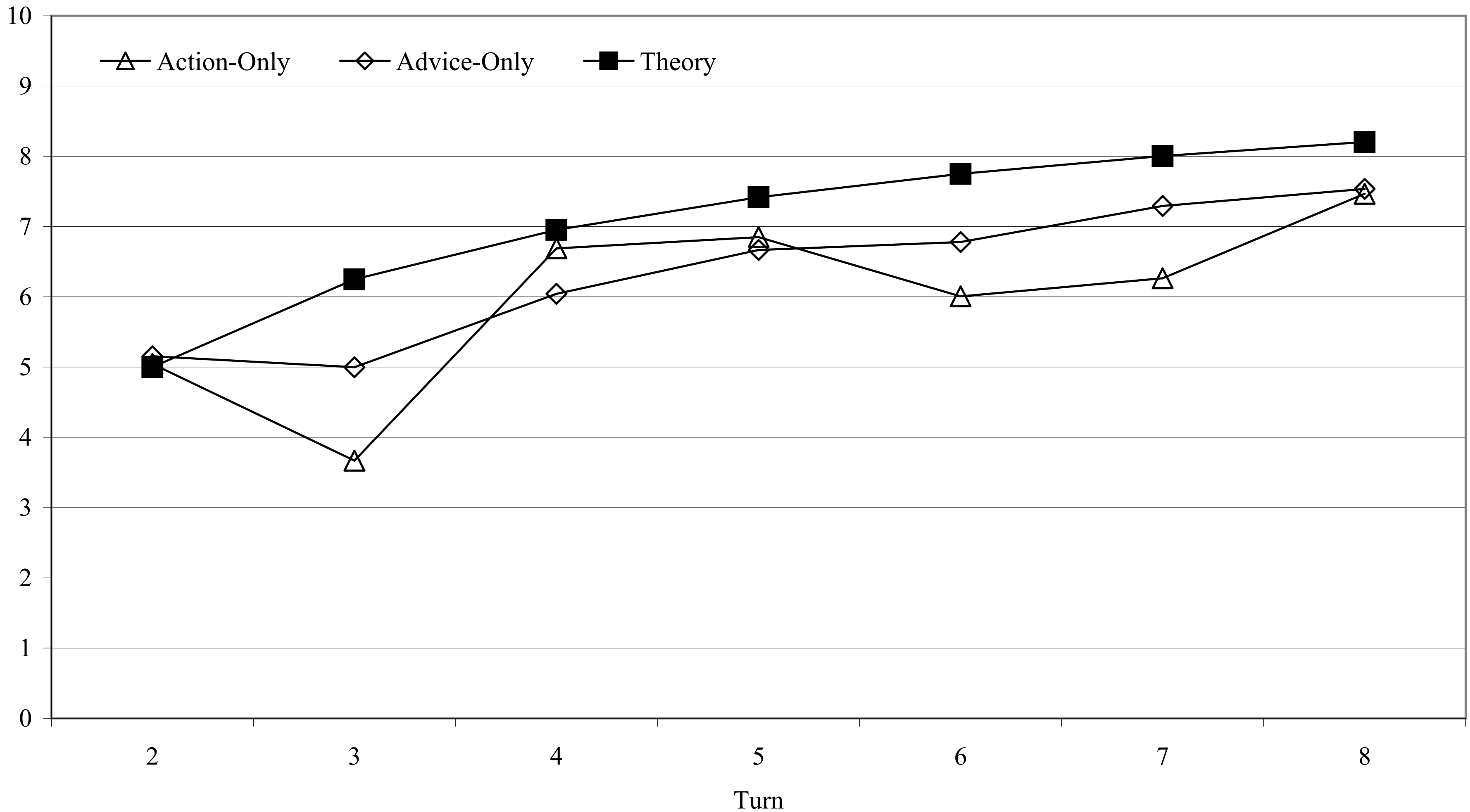
	Session. round	1	2	3	4	5	6	7	8	Sum of signals
Advice-Only	2.2	*A:B* -9 -8.01	*B:B* 0.6 -5.43	*B:B* -1 -2.11	*B:B* 0 -1.59	*B:B* 7 -7.67	*B:B* 10 0.21	*B:B* 10 0.55	*B:* 10 -8.19	-32.2
Action-Plus-Advice	6.14	*A:A* 0 3.71	*A:A* -10 0.87	*B:A* -4 -4.39	*A:A* -5.5 8.76	*A:A* -10 0.67	*A:A* -10 8.36	*A:A* -10 9.06	*A:* -10 5.94	33.0
	6.15	*B:B* 9.8 -6.71	*A:A* 5 8.29	*A:A* -10 -9.78	*A:A* -2 4.68	*A:A* -10 9.12	*A:A* -10 0.30	*A:A* -10 6.51	*A:* -10 4.76	17.2
	7.15	*B:B* 0 -9.43	*B:A* 4 3.55	*A:A* 2 6.91	*A:A* -6 2.03	*A:A* -7.5 4.16	*A:A* -10 0.35	*A:A* -10 9.08	*A:* -10 2.52	19.2

Figure 1: The distribution of contrary subjects



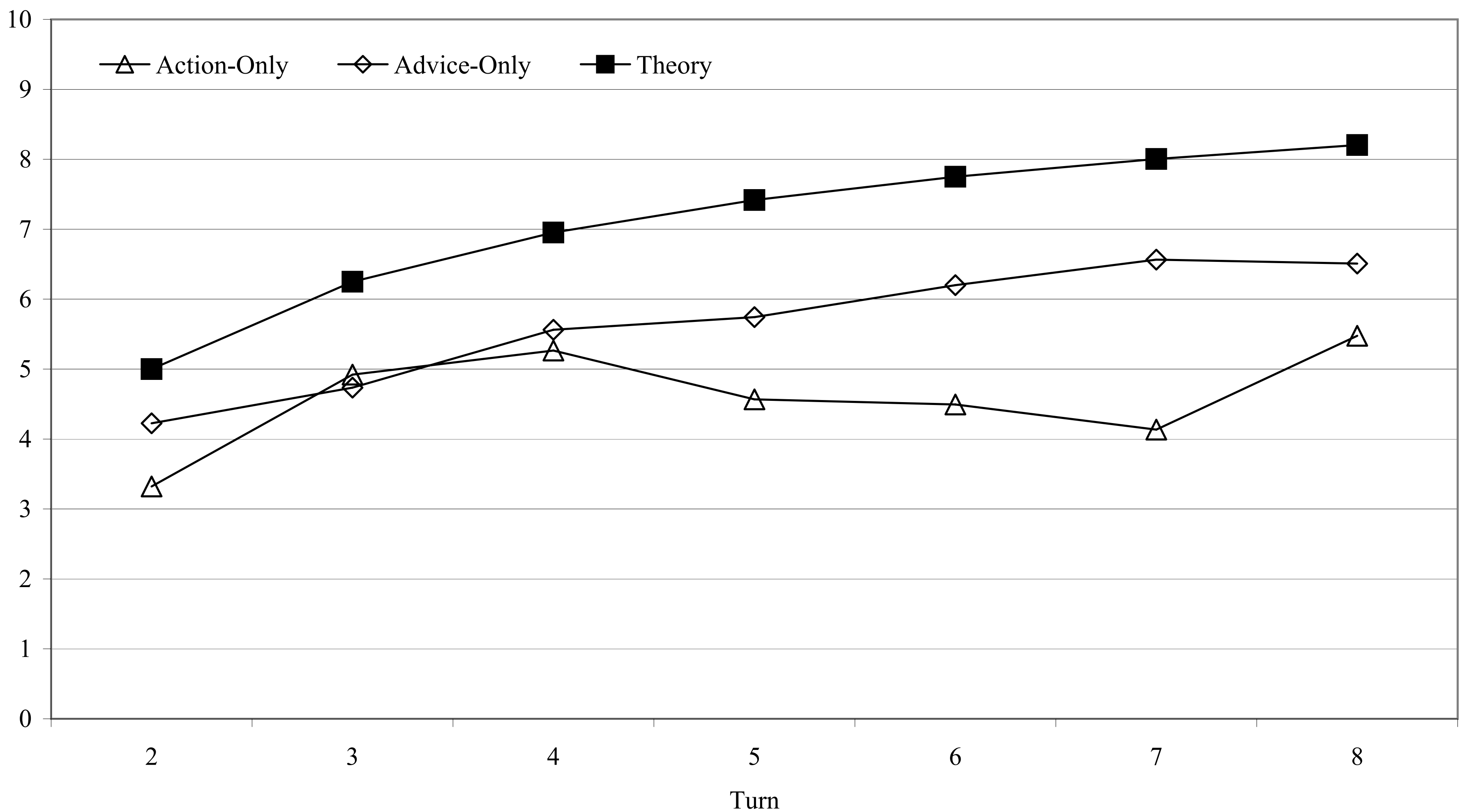
The percent of subjects who disagreed with the observed action (advice) in less than two rounds, three to five rounds and so on.

Figure 2: Mean cutoffs by decision turn in concurring concurring decisions



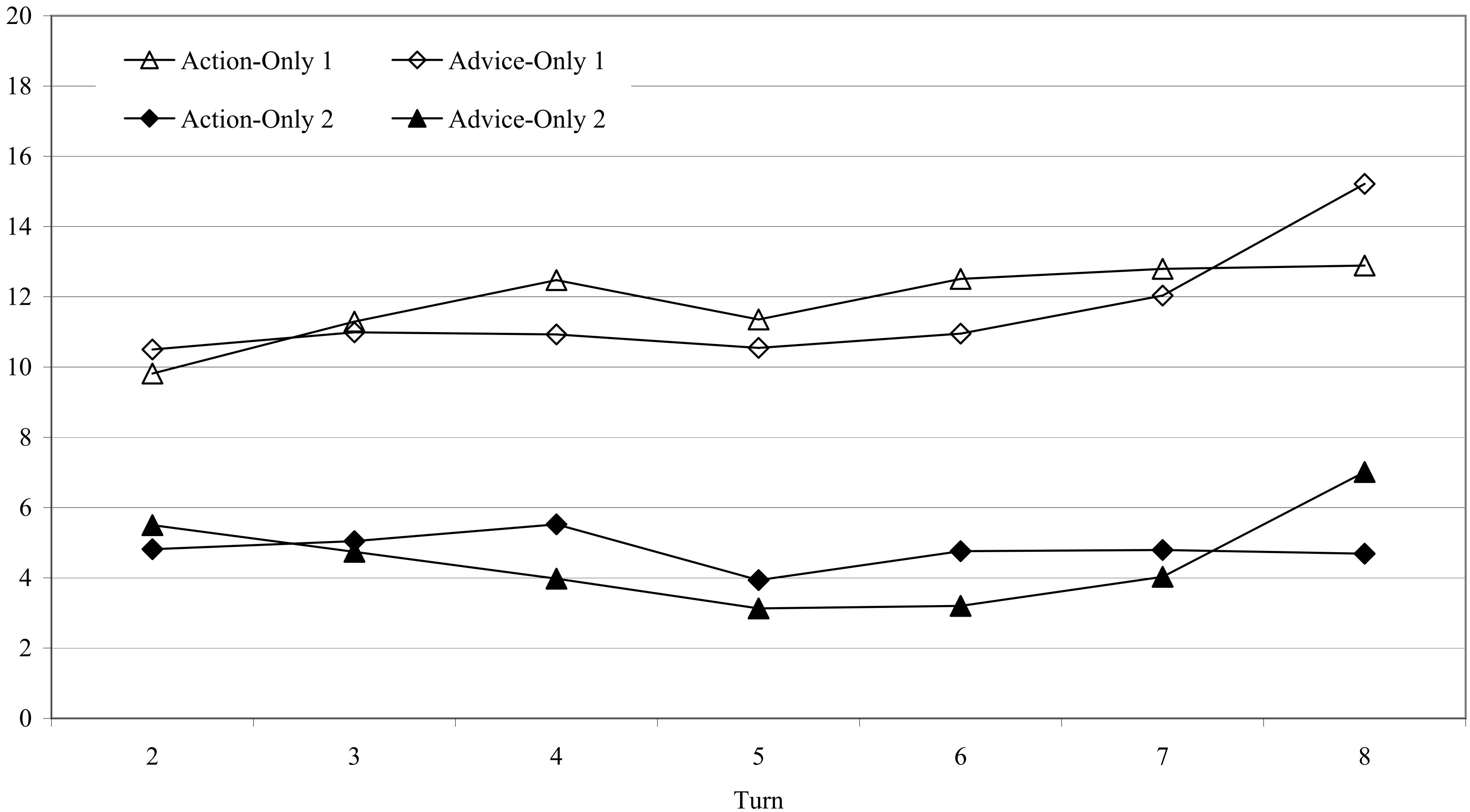
Conditional means where the conditioning is done on whether the subject's decision was a cooccurring decision.

Figure 3: Mean cutoffs by decision turn in weakly concurring decisions



Conditional means where the conditioning is done on whether the subject's decision was weakly concurring (we include the neutral decisions in our means).

Figure 4: The strength of disagreement



The absolute value of the distance between the cutoff chosen and that which would be set if the subject acted according to the theoretical cutoff rule (1), and between the cutoff chosen and zero (2).

Figure 5: Unconditional mean cutoffs by decision turn

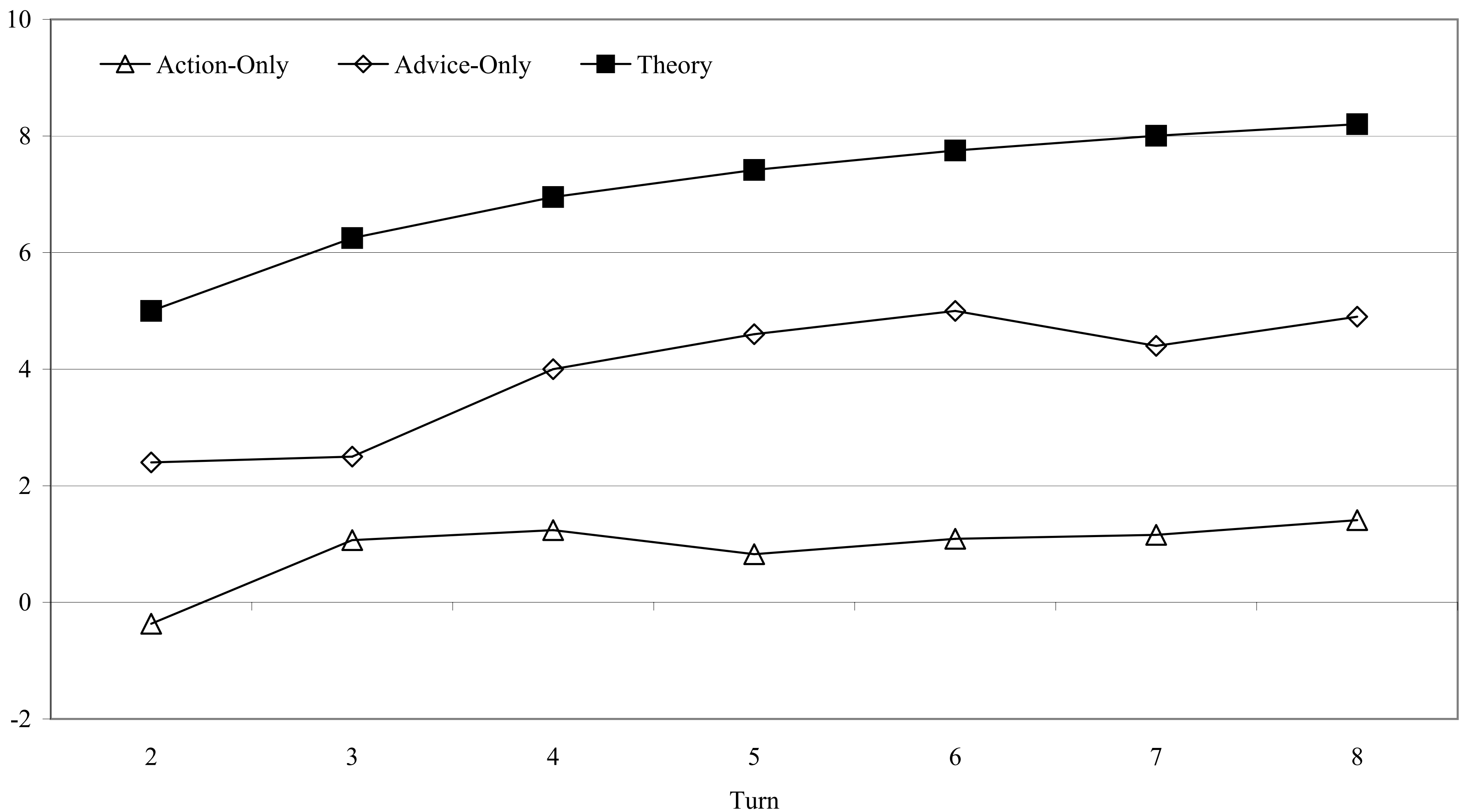
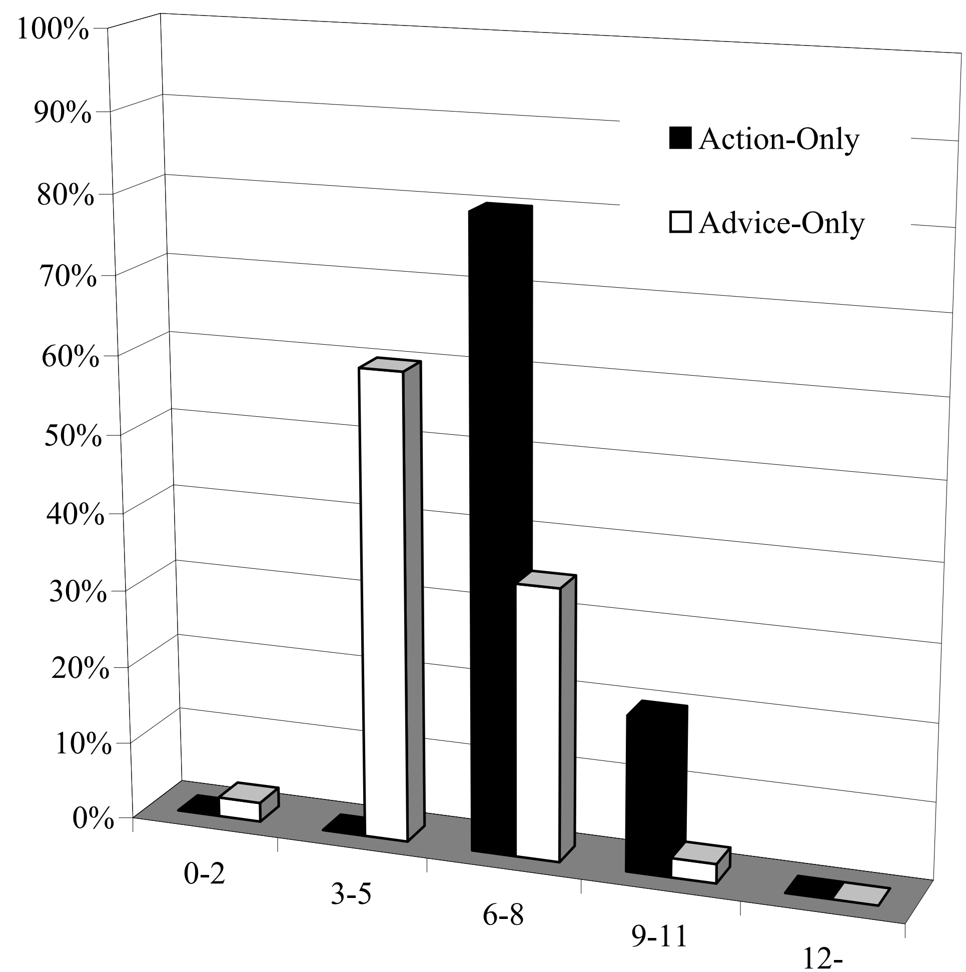
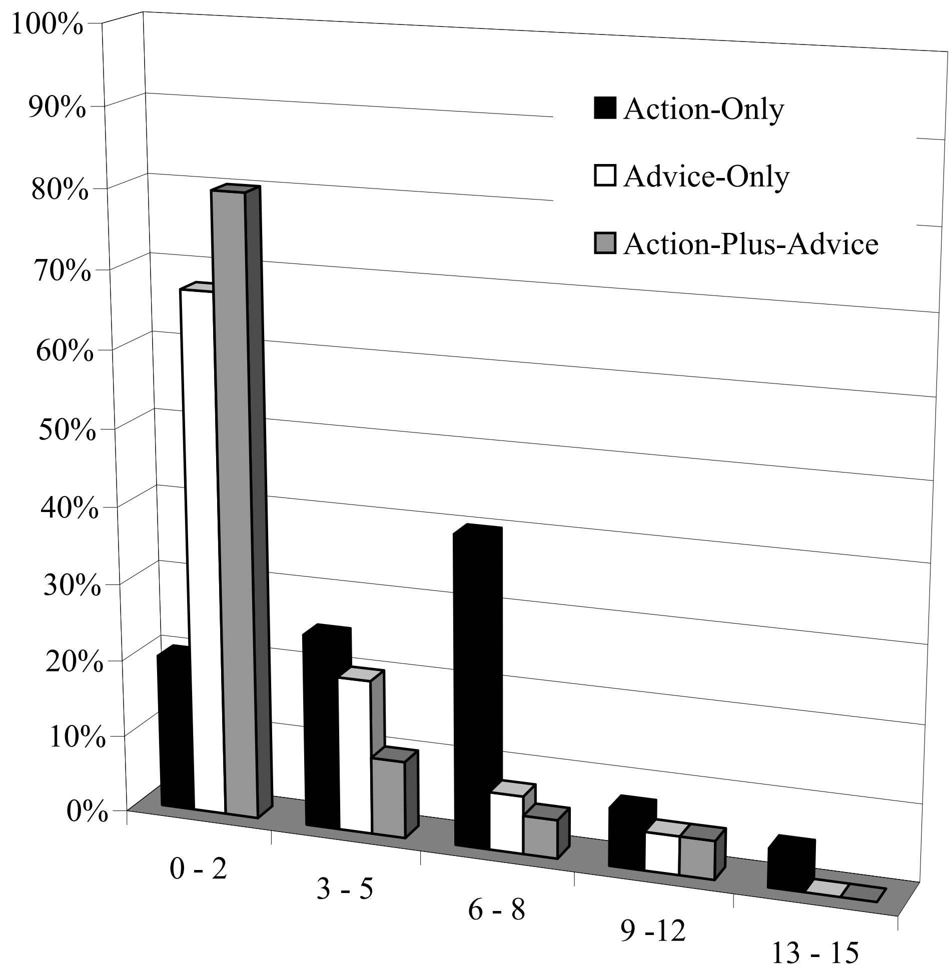


Figure 6: The distribution of subjects' MD scores



The histograms show that subject behavior is more consistent with the theory in the Advice-Only experiment as the distribution of MD scores shifts to the left when calculated using the Advice-Only data.

Figure 7: The distribution of contrary subjects



The percent of subjects who disagreed with the observed action (advice) in less than two rounds, three to five rounds and so on.

Figure 8: Unconditional mean cutoffs by decision turn

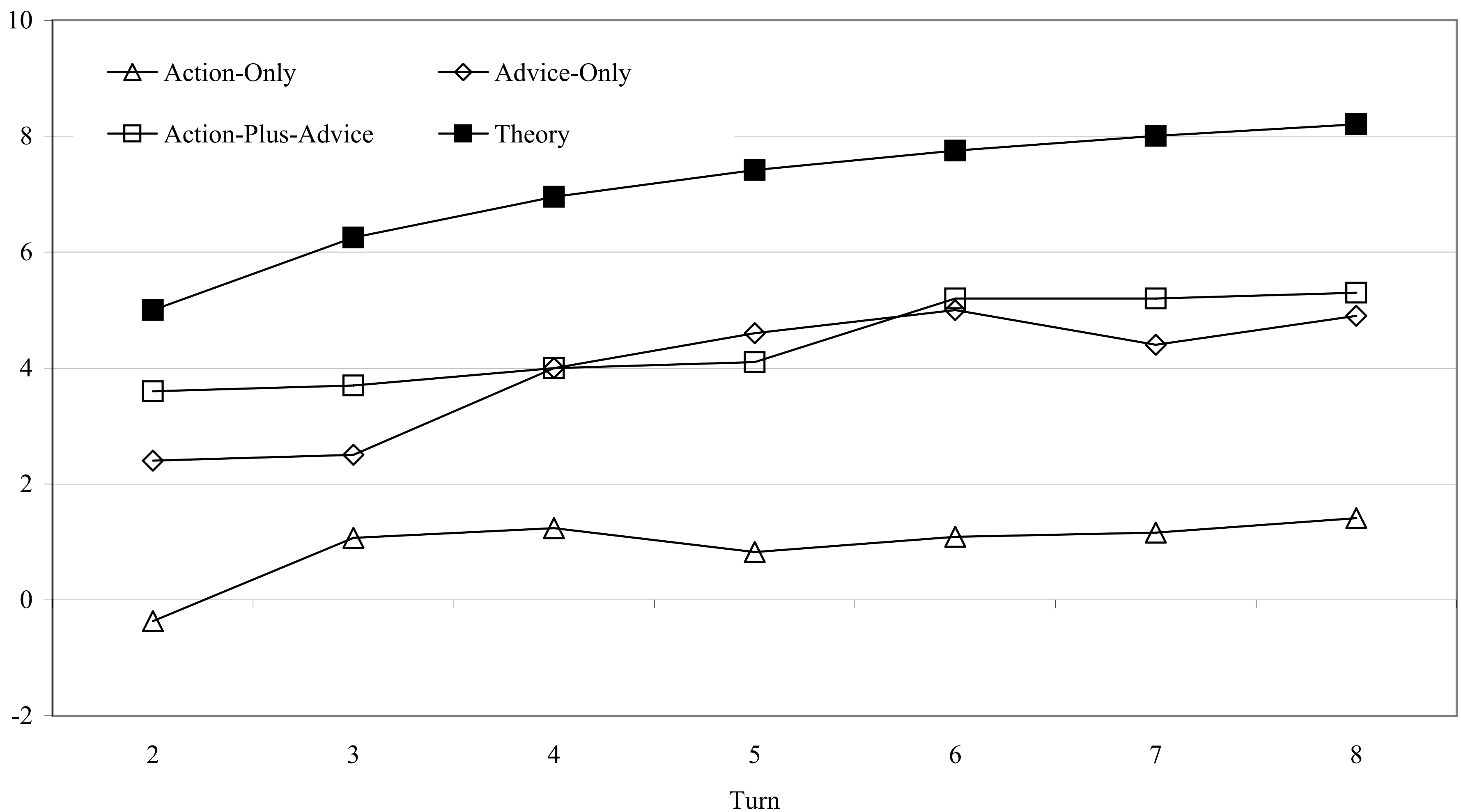


Figure 9: The distribution of advice overturning subjects

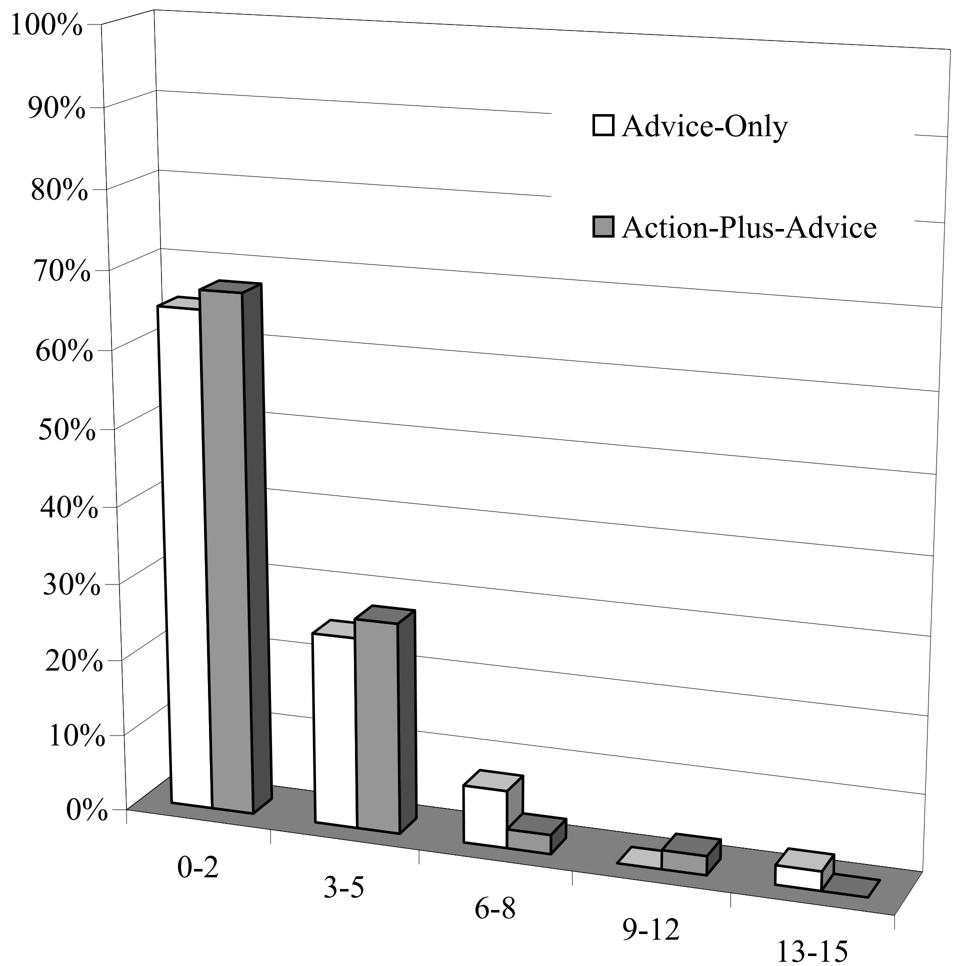
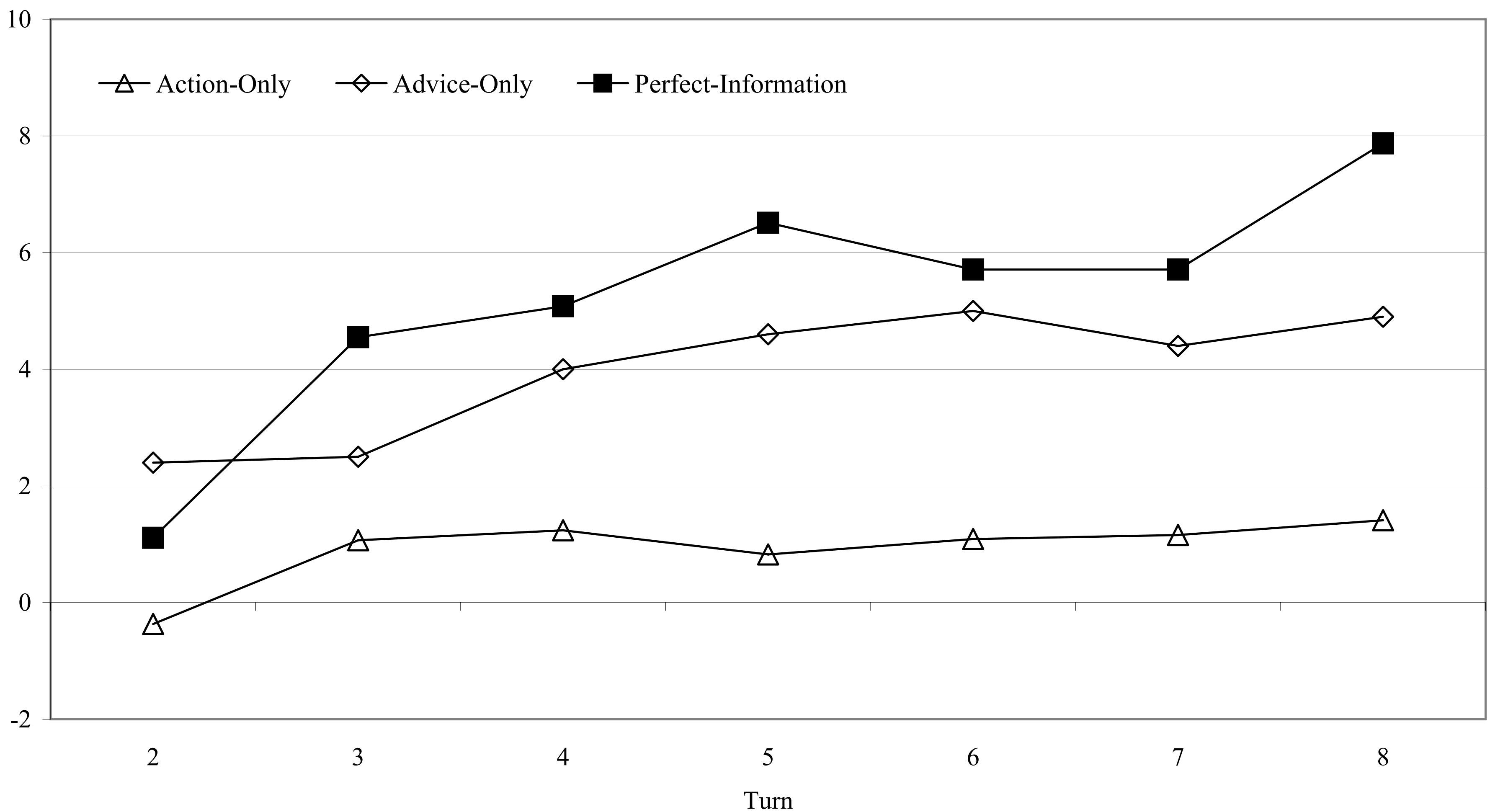


Figure 10: Mean cutoffs by decision turn in the Action-Only, Advice-Only and Perfect-Information



The mean cutoffs and the mean cutoff reported in Çelen and Kariv (2002b) experiment under perfect information in decision points in which all predecessors acted alike.