Information Cascades:
Evidence from a Field Experiment with Financial Market Professionals

Abstract:
Previous empirical studies of information cascades have used either naturally occurring data or laboratory experiments. In this paper, we combine attractive elements from each of these lines of empirical research by observing market professionals from the Chicago Board of Trade (CBOT) in a controlled environment. Our analysis of over 1500 individual decisions suggests that CBOT market professionals behave quite differently than a control group of students. For instance, market professionals are better able to discern the *quality* of public signals and their decisions are not affected by the domain of earnings. These results have interesting implications for market efficiency and are important in both a positive and normative sense.

Keywords: Herd Behavior, Futures Traders, Experiments

Acknowledgements: Thanks to the Editor, Robert Stambaugh, an Associate Editor, and an anonymous reviewer for sharp comments that improved the study considerably. John Di Clemente, former Managing Director of Research at the Chicago Board of Trade, authorized this study. Also special thanks to CBOT staff Dorothy Ackerman Anderson, Frederick Sturm, and Keith Schap for their incredible support on site. Valuable comments and suggestions were provided by seminar participants at Harvard University, University of Chicago, University of Maryland, George Mason University, American University, the NCR-134 Conference on Applied Commodity Price Analysis, Forecasting and Market Risk Management, in St. Louis, Missouri, April 2003, the Economic Science Association International Conference, in Pittsburgh, Pennsylvania, July 2003, and the Financial Management Association Annual Meeting in New Orleans, October, 2004. Liesl Koch, Vernon Smith, and Georg Weizsacker provided comments on an earlier version of this manuscript. The views expressed in this paper are those of the authors and do not, in any way, reflect the views or opinions of the U.S. CFTC.

Correspondence regarding this paper should be addressed to:
Associate Chief Economist
Room 8010, U.S. Commodity Futures Trading Commission
Three Lafayette Centre
1155 21st Street, N.W.
Washington, DC, 20581; Tel (202) 418-5063 (W); Fax: (202) 418-5660; E-mail: mhaigh@cftc.gov

In economic and financial environments where decision makers have imperfect information about the true state of the world, it can be rational to ignore one's own private information and make decisions based on what are believed to be more informative public signals. In particular, if decisions are made sequentially and earlier decisions are public information, "information cascades" that give rise to herding behavior can be a general result. Information cascades arise when individuals rationally choose identical actions despite having different private information.[1] Cascades may arise in myriad settings, including technology adoption, medical treatment choices, and responses to environmental hazards. Arguably the most well known herds or cascades occur in financial markets, where bubbles and crashes may be examples of such behavior.[2]

Information cascades can be suboptimal, since the private information of cascade followers is not revealed. As a result, the small amount of information revealed early in a sequence has a large impact on social welfare. One consequence is that cascades can be fragile as well, with abrupt shifts or reversals in direction when new information becomes available (Bikhchandani et al. (1992, hereafter BHW; 1998); Gale (1996); Goeree et al. (2004b)). Indeed, some argue that volatility induced by herding behavior can increase the fragility of financial markets and destabilize the broader market system (Eichengreen et al. (1998); Bikhchandani and Sharma (2000); Chari and Kehoe (2004)).

Previous empirical approaches that examine cascade behavior can be divided into two classes: regression-based tests using naturally occurring data and laboratory experiments using data gathered from student subjects. Bikhchandani and Sharma (2000) review the extant regression-based results for herding in financial markets but note the difficulty of controlling for underlying fundamentals. A result of this difficulty, they argue, is that there is often "a lack of a direct link between the theoretical discussion of herding behavior and the empirical specifications used to test for herding."[3] The laboratory environment allows for the control of public and private information so explicit tests of theoretical predictions are made more easily. Yet an important debate exists about the relevance of experimental findings from student subjects for understanding phenomena in the field. For example, there are numerous reasons to suspect that professional behavior in the field might differ from student behavior in laboratory experiments, including training, regulation, developed rules of thumb, and the overall experimental environment (see, e.g., Harrison and List (2004)). Locke and Mann (2005) argue that financial market research that ignores the effect of professional expertise is likely to be received passively because "ordinary" individuals, as opposed to professional traders, are too far removed from the price discovery process. Bikhchandani and Sharma (2000, p. 13) argue similarly that "To examine herd behavior, one needs to find a group of participants that trade actively and act similarly."

We find these arguments compelling and therefore combine the most attractive aspects of the two classes of empirical research—observation of professionals in a controlled environment—to extend the literature in several new directions. First, we compare the behavior of market professionals from the floor of the Chicago Board of Trade (CBOT) with that of college students in an experimental setting in which the underlying rationality of herd behavior can be identified. Given the vast normative implications of work that has established the importance of the domain of earnings for decision making under risk (Kahneman and Tversky (1979); Shefrin and Statman (1985);

Odean (1998)), we also examine the behavior of each group in the gain and loss domain. We further examine whether, and to what extent, cascade formation is influenced by both private signal strength and the *quality* of previous public signals, as well as decision heuristics that differ from Bayesian rationality. Finally, within the group of market professionals, we examine the extent to which differences in cascade formation are associated with individual characteristics, such as whether the participant is a day trader.

Empirical findings, gained from an examination of more than 1500 individual decisions, lend some interesting insights into cascade behavior. A key finding is that market professionals tend to make use of their private signal to a greater degree than their student counterparts. Furthermore, professionals base their decisions on the *quality* of the public signal to a *greater* extent than do students. As a result, the professionals are involved in weakly fewer overall cascades, and significantly fewer *reverse cascades* (cascades that lead to inferior outcomes). This result is novel to the literature and has important implications for financial markets.[4] Further, the behavior of students is consistent with the notion that losses loom larger than gains, while market professionals are unaffected by the domain of earnings. This finding is consonant with Locke and Mann (2005), Genesove and Mayer (2001), and List (2003; 2004), who find, in much different environments, that market experience is associated with a decline in deviations from classical assumptions.

While we observe behavioral differences across subject pools, within the market professional group we find some significant differences in behavior as well. For example, Bayesian play is correlated with market experience and day traders are much more likely to join an informational cascade than non-day traders. Finally, we present data on the prevalence of non-Bayesian decision heuristics, an area in which the two subject pools demonstrate similarities.

The remainder of the study is crafted as follows. Section II outlines the basic theory and experimental design. Section III presents our empirical results. Section IV considers the implication of our results for financial markets and briefly discusses the use of professionals in experimental practice more broadly. Section V concludes.

II. Theory and Experimental Design

Imitative behavior associated with herding has often been viewed as the product of irrational decision-making (Keynes (1936); Shleifer and Summers (1990); Hirshleifer (2001)). Alternatively, models due to Banerjee (1992), BHW (1992), and Welch (1992) discuss the conditions under which it is rational to join a cascade. The model we present below, and the experimental environment we implement, is consistent with the work of these authors in that it is predicated on Bayesian updating of beliefs, given private signals and a history of observable actions.[5] The empirical investigation of the cascade phenomenon raises interesting questions beyond whether agents update information in a manner that is consistent with Bayes' rule.[6] Since the formation of informational cascades is a social phenomenon, individual behavior may depend on how agents view the rationality of others. In particular, we examine how our two subject pools respond to uncertainty about the quality of information that arises due to potential deviations from Bayesian rationality by others. We adopt two approaches. First, we use a model in which the null hypothesis is that Bayesian rationality is universally applied and is common knowledge. Second, we estimate a quantal response equilibrium (QRE) model that assumes decision error (McKelvey and Palfrey (1998); Goeree et al. (2004a).

II.A. Theoretical Model and Predictions

Consider an environment in which there are two possible underlying states of nature $\Omega = \{A, B\}$, with the true state denoted by $\omega \in \Omega$. Each of a set of $I = \{1,2,..,n\}$ agents receives an independent private signal, $s_i \in \{a,b\}$ that is informative in the sense that $\Pr(A \mid a) > \Pr(B \mid a)$ and $\Pr(A \mid b) < \Pr(B \mid b)$. Signal precision, given by $\Pr(s = \omega \mid \omega)$, is identical for all agents. After receiving their signal, each agent chooses either *A* or *B* with their choice, c_i. If $c_i = \omega$, a reward, normalized to 1, is received by individual *i*. Those choosing the incorrect state receive 0. Each individual receives their signal in an exogenously determined choice order. Along with their private signal s_i, each agent observes the history of choices, $H_i = \{c_1,...,c_{i-1}\}$. The prior probability of an underlying state, given by $\Pr(\omega = A) = p$ and $\Pr(\omega = B) = 1 - p$, is common knowledge. If all individuals update beliefs according to Bayes' rule and this updating is common knowledge, the posterior probability $\Pr(\omega \mid H_i, s_i)$ is easily derived. The formation of an information cascade in this setting is demonstrated via a simple example, parameterized with the probabilities from one of our experimental treatments.

We choose $\Pr(\omega = A) = \Pr(\omega = B) = p = 1/2$ as the prior probability, and the precision of the symmetric signal is given by $\Pr(a \mid A) = \Pr(b \mid B) = 2/3$, and so $\Pr(b \mid A) = \Pr(a \mid B) = 1/3$. Suppose that $s_1 = a$. Bayes' rule implies that

$$\Pr(\omega = A \mid s_1 = a) = \frac{\Pr(a \mid A)\Pr(A)}{\Pr(a \mid A)\Pr(A) + \Pr(a \mid B)\Pr(B)} = \frac{2}{3}.$$

An expected utility maximizer would, therefore, predict *A* as the state of nature since expected profits for announcing *A*, π_A, exceed those for announcing *B*, π_B.[7] If the second subject also receives an *a* signal, updating according to Bayes' rule yields

$$\Pr(\omega = A \mid H_2 = A, s_2 = a) = \frac{\Pr(a \mid A)^2}{\Pr(a \mid A)^2 + \Pr(a \mid B)^2} = \frac{4}{5}.$$

Thus, two consecutive identical announcements yield a posterior probability of 0.80 in favor of the indicated urn.[8] As a result, the third decision maker should "follow the herd" and choose $\omega = A$ regardless of her signal, as can be seen by examining the posterior where an opposing *b* signal is the private draw of the third player after two consecutive *A* announcements:

$$\Pr(\omega = A \mid H_3 = A, A, s_3 = b) = \frac{\Pr(a \mid A)^2 \Pr(b \mid A)}{\Pr(a \mid A)^2 \Pr(b \mid A) + \Pr(a \mid B)^2 \Pr(b \mid B)} = \frac{2}{3}.$$

We denote a decision of this type—consistent with Bayesian rationality, but in which one's own private signal is ignored—a *cascade decision*. In this example, the decision maker in the third position reveals nothing about their private information and thus imposes an informational externality on future players. The analysis implies that, with this parameterization, public announcements are uninformative whenever the number of public signals of one type exceeds the other by two or more. As a result, if a cascade has not started, two consecutive low probability draws can result in a *reverse cascade* in which everyone rationally herds on the incorrect state.

II.B. Experimental Design

Anderson and Holt (1997) present a seminal experimental investigation of cascade formation, using a subject pool of undergraduates. To ensure comparability of our results to the extant literature, we use experimental protocols that are closely related to their work.[9] The parameterization in the example above is consistent with their *symmetric* treatment, with its name derived from the fact that $\Pr(a \mid A) = \Pr(b \mid B)$. The experimental sessions we conduct include 15 *rounds* of the basic game for a group of either five or six players whose choice order in each round — either first, second, third, …., sixth — is determined by a random draw.

A round begins with the experimental monitor selecting the state of nature with a roll of a die that is unobserved by the subjects. Subjects gain information about the state by drawing a single ball out of an unmarked bag into which the contents of the selected urn have been transferred. The draw is made while the subject is isolated from other players. The monitor is informed of the choice of the state, and announces it publicly. After all subjects have made their choices, the true state is revealed.

To provide exogenous variation in the informational content of the private signal across treatments, we use two urn types. In the *symmetric* treatment, Urn A contains two type *a* balls and one type *b* ball, while Urn B contains two of type *b* and one of type *a*. To create the *asymmetric* treatment, we add four *a* balls to both urns, yielding 6 (5) *a* signals and 1 (2) *b* signal in the *A* (*B*) state. This change results in a significant dilution of the strength of an *a* signal; the relative weakness of which can be observed in Table I, which provides posterior probabilities for all possible signal histories for both the symmetric and asymmetric urn types. As an example, the two-thirds probability arising after a single *a* draw in the symmetric treatment arises after four consecutive *a* draws in the asymmetric setting. One consequence of the change in signal strength is that in the asymmetric treatment a cascade on the *B* state should take place after one *b* signal even with either one or two *a* signals in the game's history.

The difference in signal strength across urn types allows us to investigate the relationship between Bayesian updating and a choice heuristic based on a counting rule. In the symmetric treatment, the optimal decision is always consistent with choosing the state with the most informative signals. In the asymmetric case, a number of sequences violate this counting rule in that it is optimal to choose B even when there are fewer *b* signals. These four *non-counting rule* sequences, $(a,b) \in \{(2,1),(3,1),(3,2),(4,2)\}$, are underlined in Table I. The asymmetric treatment, therefore, allows us to gain insights into the extent to which decisions are better represented as counting and perhaps only superficially Bayesian.

To provide exogenous variation in the earnings domain, we randomly place subjects in either a *gain* or *loss* treatment for all 15 rounds. The treatment was implemented so that in gain (loss) space a correct (incorrect) inference about the underlying state resulted in positive (negative) earnings of $1 for students and $4 for the market professionals.[10] An incorrect (correct) choice in gain (loss) space resulted in no earnings. To provide similar monetary outcomes across treatments, in the loss treatments, students and market professionals were, respectively, endowed with $6.25 and $25.00.[11] We believe that this is the first study to vary the gain/loss domain in cascade games.

Experimental subjects in a particular session consisted entirely of one of the two subject types: either students or market professionals. The experimental sessions with market professionals were conducted at the Chicago Board of Trade (CBOT) and the student data were gathered from undergraduates at the University of Maryland in College Park. The CBOT (student) subject pool included 55 (54) subjects recruited from the floor of CBOT (the university). The resulting experimental design is a 2x2x2 factorial across urn type (symmetric (S) or asymmetric (A)); across domain (gains (G) or losses (L)); and across subject type (college undergraduates (C) or market professionals (M)). Each experimental session consists of a group of either five or six participants making decisions within the same treatment type over 15 rounds. Table II summarizes our experimental sessions.

III. Experimental Results

Table III.a presents descriptive statistics from the experiment. We include the rate of Bayesian decision making and the rate of cascade formation for the pooled data and by treatment, with a Bayesian decision defined assuming common knowledge of Bayesian rationality (no decision error). The 20 experimental sessions yielded a total of 1,647 decisions, 1,284 (78 percent) of which were consistent with a perfect Bayesian equilibrium.[12] Cascade decisions (i.e., Bayesian decisions in contrast to the private signal) occurred in 15 percent of the choices. Of these, just under one quarter (55 out of 245) were "reverse" cascades, resulting in the wrong inference about the underlying state.

Perhaps more revealing than the aggregate number of cascades is the proportion of cascade decisions made when the opportunity arises. Recall that a cascade decision is possible only when the private draw is inconsistent with the probability weight derived from the choice history and one's own private signal. In our data, cascade formation was possible in 441 of the decisions, representing 27 percent of the total. Cascades were realized in 245 (56 percent) of these cases. These results are presented in the *potential* and *realized* cascades columns of Table III.a.

Table III.a also reports statistics disaggregated by subject and treatment type. In aggregate, 81 percent (75 percent) of the students' (market professionals') decisions were consistent with Bayesian Nash equilibrium. Decisions of individual subjects ranged from 38 percent to 100 percent Bayesian (these results are not shown to conserve space), and of the fourteen subjects perfectly consistent with Bayesian rationality, ten were students.[13] In situations where Bayesian behavior required one to ignore private information, fewer agents were Bayesian: the final column of Table III.b shows that 61 percent (49 percent) of students (market professionals) ignored their signal when doing so lead to a cascade. Interestingly, rates of cascade formation and Bayesian decision making are lower in the asymmetric treatments for both subject pools.

The final set of descriptive statistics is presented in Table III.b, which displays results from the asymmetric treatments pooled, and parsed by subject pool and sequence type, where the type is either a counting rule or non-counting rule sequence.[14] Table III.b demonstrates that both Bayesian decision-making and cascade formation decline when the rules are not reinforcing: the proportion of Bayesian decisions by students (market professionals) declines from 76 percent (76 percent) to 49 percent (42 percent), and the rate at which cascades are realized declines from 58 percent (55 percent) to 39 percent (26 percent). These results suggest that the non-counting rule sequences pose a challenge for both subject pools.[15]

To permit more formal inference, we apply a variety of parametric and non-parametric statistical techniques and group our results into five categories. Three of the categories compare students and market professionals to consider differences in (1) Bayesian decision making, (2) cascade formation, and (3) behavior across the gain/loss domain. A fourth category concentrates on data from market professionals by making use of additional demographic data collected during the experiment. The fifth category considers the exogenous alteration of signal strength through the use of the symmetric and asymmetric urns. Our analysis leads to a first insight:

> Result 1. *Market professionals are less Bayesian than students. Despite this behavioral discrepancy, earnings are not significantly different across subject pools.*

To provide evidence of this result we employ both unconditional and conditional statistical tests. When using unconditional tests, we recognize the data dependencies within an experimental session by using session level aggregates to yield the most conservative statistical tests. Our unconditional test used to support Result 1 is a non-parametric Mann-Whitney U test, which indicates that the rate of Bayesian decision making differs across subject pools at a level of significance of $p = .052$.[16]

To complement this analysis, we use conditional testing procedures that make use of the panel nature of our data; in particular, we use a random effects probit specification of the following form:

$$Baye_{it} = \beta` X_{it} + e_{it}, \qquad e_{it} \sim N[0,1], \qquad (1)$$

where $Baye_{it}$ equals unity if agent i was a Bayesian in round t under the assumption of no decision error by preceding players, and equals zero otherwise; X_{it} includes treatment effects (*gain*, *sym*, and *trader*) and other variables predicted to influence play (*order*_x, *diff*, and *heurist*). The treatment variables are as defined above: *gain* equals one (zero) for sessions over gains (losses); *sym* equals one (zero) for the symmetric (asymmetric) sessions; and *trader* equals one (zero) for the market professionals (students).

The remaining variables are defined as follows: *order_x (where x=2,..,6)* is a categorical variable indicating the positional order in which the individual choice was made. The posterior probability is incorporated in the variable *diff* which is calculated as $\left|\Pr\left(\omega = A \mid H, s\right) - .5\right|$, and thus measures the accrued public and private information at the disposal of each decisionmaker. Thus, the *diff* variable varies from zero to one-half, increasing with evidence of the underlying state.[17] The variable *heurist* is equal to one (zero) for non-counting rule (counting rule) sequences. In a perfect Bayesian equilibrium, the coefficients of these latter two variables should not differ from zero.

We specify $e_{it} = u_{it} + \alpha_i$, where the two components are independent and normally distributed with mean zero: $\mathrm{Var}(e_{it}) = \sigma_u^2 + \sigma_\alpha^2$. We estimate equation (1) using the maximum likelihood approach derived in Butler and Moffitt (1982). Estimation of this model is amenable to Hermite integration. To estimate the model, we use a twelve-point quadrature and the method of Berndt et al. (1974) to compute the covariance matrix.

Empirical results are reported in Table IV, which presents marginal effects associated with a change in each of the regressors computed at the overall sample means.[18] Concerning subject pool effects, results from both a likelihood-ratio test and the *trader* dummy variable in the pooled regression model (panel 4a) support the non-parametric findings: market professionals are less Bayesian than students.[19] The estimated marginal effect in the pooled model suggests that traders are 6 percent less likely to be Bayesian, and this effect is significant at the $p < .05$ level.

Despite the noisier environment (fewer professionals are Bayesian), market professionals and students choose the correct underlying state at similar rates. Indeed, using a Mann-Whitney U-test, we find that we cannot reject the homogenous null that success rates are similar at conventional levels (p= 0.29), leading to the result that earnings are similar across the subject pools. To dig a level deeper into this finding, we estimate a model similar in spirit to equation (1), but make the dependent variable *win* dichotomous and equal to unity (zero) if the individual chose correctly (incorrectly). An additional independent variable, *round,* is included to identify learning during the course of the session. *Round* is a time trend and increases from 1 to 15 within a session.[20]

Empirical results summarized in Table V support the non-parametric finding concerning earnings and provides more formal evidence of the second half of Result 1: the *trader* variable in panel 5a of Table V is not significantly different from zero at conventional levels (p = .27). This result suggests that traders and students choose the correct urn at similar rates. The two groups differ, however, in their temporal play, as evidenced by the significant (insignificant) and positive marginal effect of *round* for the traders (students), consistent with learning effects among traders.

Besides providing empirical support for Result 1, the models in Tables IV and V reveal some of the important effects of the other independent variables. For example, the *diff* and *heurist* coefficient estimates in the pooled model of Table IV show that a marginal change in the posterior probability has a large positive effect (66 percent), while decisions in the *counting rule* sequences are 23 percent less likely to be Bayesian than those in which counting and Bayesian posterior imply the same result. Similar insights are found when splitting the sample by subject type, as summarized in panels 4b and 4c of Table IV. In addition, the effect of *diff* is found to be statistically significant for both subject pools in the Table V *win* models.

Interestingly, urn symmetry, as captured by the *sym* dummy variable, is not significant for the market professionals in either model, implying that the difference across urn types, for the traders, is captured by the counting rule distinction. Yet, in Table IV we find that the urn difference has a significant influence on students, who are much more likely to be Bayesians in the symmetric treatment (see panel 4b). A final important difference is that the *order_x* variables indicate a decline in Bayesian behavior among market professionals who chose in the third through fifth position. The magnitude of the effect is rather large, having from one-third to two-thirds of the effect of the counting rule sequences as represented in the *heurist* variable (Table IV, panel 4c). In contrast, the students show no such effect. The behavior reflected in this finding is consistent with the idea that the market professionals recognize that no new additional information is added by choices once a herd has been formed.

Given the significance of the *diff* and *heurist* variables, we further explore the individual data in a QRE model, which examines the degree to which incentives affect error rates in decisionmaking. Following Anderson and Holt (1997), we focus on data from our symmetric sessions and make use of the QRE model developed by McKelvey and Palfrey (1995; 1998); (see also Goeree et al. (2004a, 2004b)). The QRE model assumes that the probability of choosing an urn is increasing in its expected value. And, given the positive and significant coefficient on the *diff* variable in Table IV, the usefulness of such a model for both the students and market professionals appears evident. For parsimony, we point the reader to the Appendix and only briefly describe the results here.

Table A1 in the Appendix reports estimates of the lambda parameter in the QRE model. The lambda parameter indicates the extent to which noise affects decision outcomes. As $\lambda \to \infty$, the choice converges to the Bayesian outcome. As $\lambda \to 0$, the decisions become purely random. Significant differences in lambda across the subject pools are observed at choice orders one, two, and five as shown by the p values in column "p" of Table A1. Particularly noticeable is the difference at choice order two where the students exhibit few errors. The differences in noise in the first two choice orders leads to quite different behaviors in choice order three, despite the fact that estimates of the lambda parameter are indistinguishable.

The lambda estimates imply that the two subject pools have similar deviations from Bayesian rationality at choice order three. Thus, the market professionals' tendency to rely on their own signal due to errors in earlier rounds is as rational as the students' decision to ignore theirs and join the cascade. Table A2 clarifies the meaning of this result by examining in detail the impact of the noisy decision process on revealed public information and choice probabilities for the first three rounds of play. For comparison, the posteriors and choice probabilities assuming a perfect Bayesian equilibrium are presented, as are the actual individual decisions.

Consider the posterior probability for choice order 3 in Table A2, the first choice where a cascade may form in the symmetric treatment, when the signal history is AAb (or BBa). In this case, the posterior probability of urn A has dropped from 0.67 for the most likely urn to 0.51 (0.59) for the market professionals (students). Thus, while ignoring one's private information is optimal for both groups, the noise in prior decisions dilutes the strength of the signals, with the market professionals facing essentially a random choice. And, the probability that urn A is chosen is .54 (.83) for the market professionals (students). The differences across the sequences in choice order three highlights the fact that noise in the decision making process dilutes the value of the public signal.

Together, both the probit and QRE estimations provide insights into the differences across the subject pools. An interesting finding is that the two groups have similar earnings, despite the noisier environment of the market professionals. Further exploration into this observation leads to the following two insights:

> Result 2a. *In aggregate, the rate of cascade formation is not significantly different for the students and market professionals, but market professionals enter into fewer reverse cascades in the asymmetric treatments.*
>
> Result 2b. *Market professionals are better able to discern the quality of the signal associated with other players' announcements than students.*

Evidence in favor of Results 2a and 2b follow from both non-parametric and parametric statistical tests. Even though realized cascades are roughly 60 percent among students and only 50 percent among professionals (see Table III.a), using a Mann-Whitney test the homogeneous null cannot be rejected at conventional levels (Mann-Whitney p=.33).

While this result indicates that there is only weak evidence that students enter into a greater number of cascades than professionals, there are significant differences across subject pools in the rate of cascade formation in the asymmetric urn treatment. Table III.a shows that in the asymmetric treatment only 12 percent (8 of 66) of the cascades entered by market professionals are reverse cascades. This is roughly half of the rate observed for students (25 of 99), a difference that is statistically significant at the $p < 0.05$ level using a Mann-Whitney test.

To complement these non-parametric insights, we estimate models similar to equation (1), but make the dependent variable equal to one when a cascade is formed and zero otherwise. To conserve space, we do not formally table these results since they reinforce the non-parametric insights gained above, but in the cascade formation model that pools the symmetric and asymmetric data, we find that cascade formation is similar across the students and market professionals. Alternatively, when we focus on reverse cascades and use only the data from the asymmetric urn treatments, we find that students enter significantly more reverse cascades than professionals.

These results cannot be explained by our model of decisionmaking based on posterior probabilities derived from signals and actions. We therefore investigate the hypothesis that market professionals use auxiliary information that the students ignore in order to avoid reverse cascades. To do so, we augment the cascade formation model discussed above by considering whether subjects use information specific to individuals selecting prior to them in the current round. To this end, we construct two variables that each provide an indication of the Bayesian decision making of subjects who preceded each player in a particular round.

These variables, denoted *othb_max* and *othb_min*, provide an indication of previous play of the most Bayesian and least Bayesian players. For example, for *othb_min* for player i whose choice order is x in round t we calculate

$$othb\_min_{it}^{x} = \min\left[\frac{\sum_{t=1}^{t-1} baye_{jt}}{t-1} \quad \forall\ j : x_{jt} < x_{it}\right] \qquad (2)$$

The proportion of Bayesian decisions for the individual with the lowest proportion among all j preceding the current decision maker is used as the independent variable in this case, although the empirical results are robust to other specifications including replacing the min operator with the mean or max. In the case of *othb_max*, we simply replace "min" with "max" in equation (2). Note that these variables are calculated for each *t* (round of the game) so that it includes only those decisions that have already occurred. The variables *diff*, *heurist*, and *gain* are also included and are defined as in the previous models.

Empirical results are presented in Table VI. Since the results across models yield similar insights concerning the nature of interpreting signals, we focus on the *othb_min* results. Although *othb_min* is insignificant in the pooled specification in panel 6a, this result masks a difference in how the two subject pools respond to the announcements of others. Results in panel 6c of Table VI suggest that cascade formation for the market professionals is significantly and substantially associated with the *quality* of the others' signals. The marginal effect of a higher minimum in the preceding players' share of Bayesian decisions is 47 percent, the largest of the variables that are statistically significant, and an indication of the impact of the inferred signal quality on the willingness to make a decision that relies on others. This variable is significant and negative in the student sample (panel 6b).

Using *othb_max* in the regression yields an insignificant effect for the students, while the market professionals again respond positively, with a marginal effect of 57 percent (detailed results omitted). We therefore conclude that the market professionals make better use of the available public information, incorporating evidence on others' rationality in their decision making in a way that is payoff relevant.[21] Note also that, in contrast with what was found for all decisions (Table IV), the *diff* variable is not significant for either group when restricting attention to the subset on cascade formation.

One may wonder whether the result on signal quality is due to market professionals having a greater level of previous interaction with one another than students have had, or if, alternatively, there is evidence of learning in the experiment. To explore this issue, we again examine changes in behavioral patterns during an experimental session. The evidence is consistent with the view that market professionals learn over these 15 rounds. Comparing behavior from the first and last three rounds of a session, we find that market professionals: a) significantly reduce the rate at which they join reverse cascades (from 13 percent to 2 percent) and b) increase the rate at which they join cascades with good outcomes (from 24 to 46 percent). Both results are statistically significant in probit specifications that include the cascade type as the dependent variable, and the temporal variable along with the control variables as independent variables (full results omitted to conserve space). By contrast, there are no significant changes in the rate of cascade formation for either type of cascade for the student subjects.

Our final insight concerning the comparison between students and professionals concerns the domain of the game:

Result 3. *Bayesian behavior of the student population is affected by the gain/loss domain, while market professionals are unaffected by the domain of earnings.*

Summary evidence in favor of this result can be found in Table III.a, where we observe that professionals exhibit a similar degree of Bayesian decisionmaking across the domains (roughly 75 percent). Yet, for the student sample we observe that Bayesian play increases in the loss domain. For example, considering the asymmetric treatments, we find that a Mann-Whitney test indicates that college students are less Bayesian in the gain treatment than in the loss treatment, while market professionals are unaffected by the domain of earnings (students: $p < 0.08$; traders: $p = 0.61$).[22]

Empirical estimates in Table IV provide additional evidence of this result. In the pooled data (panel 4a), the dummy variable *gain* is not significant at conventional levels, and it remains insignificant for the market professionals' specification (panel 4c). For the students, however, the parameter estimate is both significant ($p = 0.028$) and negative, indicating a 6 percent increase in Bayesian behavior in the loss domain. This result is consistent with the notion that, for the student population, losses loom larger than gains and is consonant with results in List (2003, 2004), who explored loss aversion in a much different environment. Nevertheless, consistent with the notion that repetition might attenuate such anomalies (see, e.g., Knez et al. (1985); Coursey et al. (1987)), exploring the data from the student sessions provides some evidence that the effect of the domain is mitigated via repetition.

While Results 1-3 highlight differences between the professional and student subjects, we also find important differences within the group of market professionals that are relevant for understanding their decision processes. We supplement our data with a survey implemented at the end of the experimental session and upon exploring these data more closely, we find:

Result 4. *Behavioral differences exist within the professional subject pool.*

Evidence of this result can be obtained by augmenting equation (1) using the additional demographic data collected from the CBOT floor personnel after the experiment. In this respect, we focus on collected data from a group of 28 of the 55 traders who reported information on *intensity*: the average number of contracts traded per day, *gender* – one for female, zero otherwise, *yrs* – years of experience, *income,* and *overnight*, which is dichotomous variable that equals one if the trader takes overnight positions and zero otherwise. Panel a. of Table VII reports on the Bayesian decision making and panel b. the cascade formation for these traders.

Concerning Bayesian decisionmaking, we find that the *diff* variable is not significantly different from zero. Indifference to the magnitude of the posterior, for the Bayesian models, has not been observed elsewhere in our study, and as discussed previously is consistent with Bayesian rationality, and inconsistent with theories of decision error. Variables that are significant include *heurist, intensity*, and *overnight*. As with the previous results reported in Table IV, *heurist* has a strong negative effect (-39.1 percent). Trading intensity increases Bayesian behavior slightly (0.4 percent) and overnight trade has a significantly negative impact on the rate of Bayesian decision making (-17.8 percent). The probit estimates in panel b. reveal that day traders are much more likely to join an informational cascade, as are those with higher trading intensity, with the marginal effects -80 percent on *overnight*, and 2.9 percent on *intensity*.

For those making consequential trading decisions, the link between trading intensity and Bayesian rationality is consonant with the empirical results of Locke and Mann (2005), Genovese and Mayer (2001), and List (2003, 2004), who find similar results in diverse settings that include financial, housing, and memorabilia markets. The result on trading style we believe is novel, and we offer some thoughts on its implications in the discussion section, below.

Results 1-4 highlight differences in cascade formation and Bayesian decision making across subject types, and have included the exogenous alteration of signal strength due to urn type through the *heurist* variable. Our final result looks more closely at the impact of signal strength:

Result 5. *Deviations from Bayesian norms are greatest when the counting rule and Bayesian updating make different predictions.*

Our probit specifications suggest that in those cases where counting and Bayesian rationality make different predictions, both market professionals and students are less Bayesian. Table VIII presents all of the observed signal patterns for the asymmetric treatment. Those in which the counting rule and Bayesian posteriors yield different predictions are in bold type. Statistical tests confirm what a visual scan of the data suggests: Bayesian behavior is significantly reduced in the non-counting rule sequences.[23] In fact, the four non-counting rule sequences have lower rates of Bayesian decision making than any of the other sequences, despite the fact that others have smaller *diff* values.

Figure 1 illustrates this insight by presenting the proportion of Bayesian decisions for all observed histories of play as a function of the posterior probability. The non-counting rule sequences (square entries) are uniformly lower than the other choice histories, represented as black diamonds. Compiling the results from Figure 1, we find that Bayesian behavior occurs at a rate of 44 percent in the non-counting rule choice histories and in 81 percent of the remaining choice histories in the asymmetric treatments. There is an important difference in the rate of Bayesian behavior in non-counting rule sequences that depends on whether one's decision involves choosing to join a cascade. The difference is best explained by considering whether individuals rely on their private signal. Restricting attention to non-counting rule sequences, we find that individuals are Bayesian in 31 percent of the cases when the decision involves choosing to enter a cascade. Sixty-nine percent therefore follow their own signal. By contrast 74 percent of decisions are Bayesian when there is no potential cascade and the decision is consistent with one's private information (see Table IX). Thus, when the signal history requires that Bayesian agents ignore their own signal, agents generally fail to do so. As a result, the failure of cascade decisions implies 69 percent rely on their own information – a result statistically indistinguishable from the 74 percent who do so when it is optimal. We conclude that, for the non-counting rule sequences, a Bayesian perspective provides a less accurate description of decision making than the simple rule of using private information.

IV. Discussion

Our cascade game data yields interesting evidence of heterogeneity both across the two subject pools and within the market professional group. Simple measures of performance indicate that the students outperform the market professionals. Controlling for learning about signal quality, however, makes clear that the market professionals use a more sophisticated decision process, more finely parsing the quality of public information and relying on their own signal more frequently. Within the market professional group, trading style has a strong effect on behavior, with those taking overnight positions entering cascades much less frequently.

We view these results as having potentially interesting implications for financial markets, although care must be taken with the interpretation, in part because of the fixed payoff that subjects received in our experiment.[24] Fixed prices, however, are not irrelevant in financial markets as variability in order size means that prices need not change with each transaction. Thus it is reasonable to study cascade decisions occurring at a constant price as well as those that lead to a change in price.

We believe it is plausible that the heterogeneity among traders regarding cascade formation may be related to differences in their trading practices, including those around fixed prices. Local floor traders who do not take overnight positions typically specialize as market makers and are more likely to be faced with situations in which herding, including herding at a constant price, is part of their trading practice. This type of herding may occur, for example, when several floor traders each take a portion of a large institutional order. Manaster and Mann (1999) provide evidence that market makers are willing to give up their advantage in executions, narrowing or eliminating the bid-ask spread, when they have an informational advantage over the outside order. If information is dispersed among traders heterogeneously, the situation is similar to the cascade environment we have studied. A crucial difference is that timing and transaction size in the market is endogenous, and ultimately, of course, prices do change.[25]

Avery and Zemsky (1998) introduce flexible pricing into the BHW model, and find that for cascades to form the *value uncertainty*, which was implemented in our experimental protocol, needs to be accompanied by *event uncertainty* – the possibility of a change in asset value, and *composition uncertainty* – which implies that the distribution of trader types is not common knowledge. Our results on the discernment of the quality of public announcements suggest that experienced professionals are better able to estimate the composition of the distribution of trader types, and so may act to mitigate price bubbles and crashes.[26] Clearly, additional research regarding the impact of trader specialization is warranted, but our findings highlight the benefits of controlled experimentation with non-student subject pools.

We believe that our findings may also shed light on other types of cascade behavior. Consider Welch's (1992) interesting model of initial public offerings (IPOs), for example, which addresses cascade formation at a fixed price due to regulatory requirements for IPOs. Welch finds that issuing firms have an interest in pricing to generate an informational cascade in order to increase the probability of a successful offering. Our results emphasizing the potential for cascade fragility arising from variation in the ability to interpret signal quality may be important in this context. One possible implication is that when underpricing of offerings is optimal in the Welch model, heterogeneity in signal strength and interpretation might play an instrumental role since reverse cascades in which no investment occurs will be fragile. The welfare implications, however, are not immediately obvious given that the resulting cascades are of shorter duration. Further, the importance of the effect may differ across firms or industries depending on the economies of scale of the investment and thus their need to have full or only partial subscription (Welch 1992, p. 709).

Both the differences due to specialization and the heterogeneity in signal quality and processing abilities suggest fruitful research opportunities. A more complete understanding of how the specialized skills of market participants interact in price discovery could also be explored in experiments that move towards a full market setting, but in which liquidity and informational conditions are varied in a controlled manner. A natural part of this research program would be to extend the current environment to study the impact of heterogeneity on the IPO model of Welch (1992). There is a recent study that provides evidence from asset market experiments with student subjects that mixed experience levels can reduce the incidence of bubbles and crashes (Dufwenberg et al.

2005). Heterogeneous subject pools that include professionals would shed crucial light on this issue, and help to identify the mechanisms underlying cascade formation and fragility in settings that mix fixed and variable prices.

V. Concluding Comments

In this study, we introduce market professionals from the CBOT floor to a controlled experimental environment. Making use of information cascades games, we report several insights. While student subjects more closely follow Bayes' rule, they do not perform significantly better than the market professionals along the important dimension of earnings. This puzzle is explained by the fact that professionals are more sophisticated in their use of public information. This result manifests itself over the course of the decision process: market professionals are less Bayesian when making decisions later in the choice order in a cascade game, behavior that is consistent with recognizing that previous announcements are not informative once a cascade has formed.

Market professionals also learn over the course of an experimental session to account for the quality of others' decisions, information that the student subjects fail to use. A further insight is that market professionals are also consistent in behavior over the gain and loss domains, while students, in aggregate, show evidence consistent with the notion of loss aversion. Perhaps most provocatively for the operation of markets, we find an important heterogeneity among the market professionals who make salient trading decisions that depends on their trading style. In summary, our data reveal that the decisions of market professionals are consistent with behaviors that may mitigate informational externalities in market settings, and thus reduce the severity of price bubbles due to informational cascades.

Besides revealing both positive and normative insights, our work also has methodological content. For example, it highlights the potential for experiments with students and professionals to be complementary inputs to research when field data is suggestive but inconclusive. Indeed, in making the transference of insights gained in the laboratory with student subjects to the field, a necessary first step is to explore how market professionals behave in strategically similar situations. In this spirit, we have focused on the representativeness of the sampled population to lend insights into what empirical results are similar across subject pools. A related issue concerns the representativeness of the environment, which also merits serious consideration. For example, before we can begin to make reasonable arguments that behavior observed in the lab is a good indicator of behavior in the field, we must explore whether the other dimensions of the laboratory environment might cause differences in behavior, including the abstract task, the stakes, the good, and the institution. While our research represents a necessary first step in the discovery process, we hope that future efforts will explore more fully other potentially important dimensions of the controlled laboratory experiment.

References

Anderson, L. R. 1994. Information Cascades. Ph.D. dissertation, University of Virginia.

Anderson, L. R. and C. A. Holt. 1997. Information Cascades in the Laboratory. *The American Economic Review* 87, no. 5: 847-62.

Avery, C., and P. Zemsky. 1998. Multidimensional Uncertainty and Herd Behavior in Financial Markets. *American Economic Review* 88, no. 4: 724-48.

Banerjee, A. V. 1992. A Simple Model of Herd Behavior. *Quarterly Journal of Economics* 107: 797-818.

Barberis, N., A. Shleifer, and R. Vishny. 1998. A Model of Investor Sentiment. *Journal of Financial Economics* 49: 307-43.

Berkman, H. 1996. Large Option Trades, Market Makers, and Limit Orders. *Review of Financial Studies* 9, no. 3: 977-1002.

Berndt, E. et al. 1974. Estimation and Inference in Nonlinear Structural Models. 653-65.

Bikhchandani, S., D. Hirshleifer, and I. Welch. 1998. Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades. *Journal of Economic Perspectives* 12, no. 3: 151-70.

———. 1992. A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *Journal of Political Economy* 100, no. 5: 992-1026.

Bikhchandani, S., and S. Sharma. 2000. Herd Behavior in Financial Markets: A Review. *IMF Working Paper 48*.

Butler, J. S. R. Moffitt. 1982. A Computationally Efficient Quadrature Procedure for the One-Factor Multinomial Probit Model. *Econometrica* 50: 761-64.

Celen, B., and S. Kariv. 2005. An Experimental Test of Observational Learning under Imperfect Information. *Economic Theory* 26, no. 3: 677-99.

———. 2004. Distinguishing Informational Cascades From Herding Behavior in the Laboratory. *American Economic Review* 94: 484-97.

Chamley, C. 2004. *Rational Herds: Economic Models of Social Learning*. Cambridge , U.K.: Cambridge University Press.

Chamley, C., and D. Gale. 1994. Information Revelation and Strategic Delay in a Model of Investment. *Econometrica* 62, no. 5: 1065-85.

Chari, V. V., and Patrick J. Kehoe. 2004. Financial Crises as Herds: Overturning the Critiques. *Journal of Economic Theory* 119: 128-50.

Cipriani, M., and A. Guarino. 2005a. Herd Behavior and Contagion in Financial Markets. Forthcoming*: Economic Theory* .

———. 2005b. Herding and Price Convergence in a Laboratory Financial Market. Forthcoming*: American Economic Review* .

Coursey, Don, Hovis, John, and Schulze, William. "The Disparity Between Willingness to Accept and Willingness to Pay Measures of Value." *Quarterly Journal of Economics*, 1987, 102(3), pp. 679-90.

Daniel, K., D. Hirshleifer, and A. Subrhamanyam. 2001. Mispricing, Covariance Risk, and the Cross-Section of Security Returns. *Journal of Finance* 56: 921-65.

Devenow, A., and I. Welch. 1996. Rational Herding in Financial Economics. *European Economic Review* 40: 603-15.

Drehmann, M., J. Oechssler, and A. Roider. 2005. Herding and Contrarian Behavior in Financial Markets - An Internet Experiment. Forthcoming: *American Economic Review*.

Dufwenberg, M. T. Lindqvist E. Moore. 2005. Bubbles and Experience: An Experiment. Forthcoming *American Economic Review* .

Eichengreen, B., D. Mathieson, B. Chandha, A. Jansen, L. Kodres, and S. Sharma. 1998. Hedge Funds and Financial Market Dynamics. *IMF Occasional Paper* 172.

Fama, E. 1998. Market Efficiency, Long-term Returns, and Behavioral Finance. *Journal of Financial Economics* 49: 283-306.

Gale, D. 1996. What Have We Learned From Social Learning? *European Economic Review* 40: 617-28.

Garber, Peter. 2000. *Famous First Bubbles: The Fundamentals of Early Manias*. Cambridge, Mass.: MIT Press.

Genesove, D., and C. Mayer. 2001. Loss Aversion and Seller Behavior: Evidence from the Housing Market. *Quarterly Journal of Economics* 116: 1233-60.

Gigerenzer, G., and D.J. Murray. *Cognition as Intuitive Statistics*. Hillsdale, NJ: Erlbaum.

Goeree, J. K, C.A. Holt, and T.R. Palfrey. 2004a. Regular Quantal Response Equilibrium. *Working Paper: California Institute of Technology*.

Goeree, J. K., T. R. Palfrey , B. W. Rogers, and R. D. McKelvey. 2004b. Self-correcting Information Cascades. *Working Paper: California Institute of Technology*.

Harrison, G.W. and J.A. List. 2004. Field Experiments. *Journal of Economic Literature* December 42(4): 1009-1055.

Hirshleifer, D. 2001. Investor Psychology and Asset Pricing. *Journal of Finance* LVI, no. 4: 1533-97.

Hirshleifer, David, and Siew Hong Teoh. 2003. Herd Behavior and Cascading in Capital Markets: a Review and Synthesis. *European Financial Management* 9, no. 1: 25-66.

Huck, S., and J. Oechssler. 2000. Informational Cascades in the Laboratory: Do They Occur for the Right Reasons? *Journal of Economic Psychology* 21: 661-71.

Hung. A.A., and C. Plott. 2001. Information Cascades: Replication and an Extension to Majority Rule and Coformity-Rewarding Institutions. *American Economic Review* 91, no. 5: 1508-20.

Kahneman, D., and A. Tversky. 1979. Prospect Theory: An Analysis of Decision Under Risk. *Econometrica* 47, no. 2: 263-91.

Keynes, John M. 1936. *The General Theory of Employment, Interest and Money*. London : Macmillan.

Knez, P. V. L. Smith and A. Willliams. 1985. Individual Rationality, Market Rationality, and Value Estimation. *American Economic Review* 75, no. 2: 397-402.

Koessler, F., and A. Ziegelmeyer. 2000. Tie-breaking Rules and Informational Cascades: A Note. *University Louis Pasteur Working Paper* .

Kubler, D. and G. Weizsacker. 2004a. Limited Depth of Reasoning and Failure of Cascade Formation in the Laboratory. *Review of Economic Studies* 71: 425-41.

———. 2004b. Are Longer Cascades More Stable? *LSE Working Paper* .

Lee, In Ho. 1998. Market Crashes and Informational Avalanches. *Review of Economic Studies* 65, no. 4: 741-60.

List, J. A. 2003. Does Market Experience Eliminate Market Anomalies? *Quarterly Journal of Economics* 118, no. 1: 41-71.

———. 2004. Neoclassical Theory versus Prosepect Theory: Evidence From the Marketplace. *Econometrica* 72, no. 2: 615-25.

Locke, P. R., and S.C. Mann. 2005. Professional Trader Discipline and Trade Disposition. *Journal of Fincancial Economics* 76, no. 2: 401-44.

MacKay, Charles. 1980. *Extraordinary Popular Delusions and the Madness of Crowds*. New York: Three Rivers Press.

Manaster, S, and S. C. Mann. 1999. Sources of Market Making Profits: Man Does not Live by Spread Alone. *Working Paper: Texas Christian University*.

McKelvey, R. D., and T.R. Palfrey . 1998. Quantal Response Equilibria for Extensive Form Games. *Experimental Economics* 1, no. 1: 9-41.

———. 1995. Quantal Response Equilibria for Normal Form Games. *Games and Economic Behavior* 10: 6-38.

Noth, M., and M. Weber. 2003. Information Aggregation with Random Ordering: Cascades and Overconfidence. *Economic Journal* 113: 166-89.

Odean, T. 1998. Are Investors Reluctant to Realize their Losses? *Journal of Finance* 53, no. 5: 1775-98.

Sgroi, D. 2003. The Right Choice at the Right Time: A Herding Experiment in Endogenous Time. *Experimental Economics* 6, no. 2: 159-80.

Shefrin, H. M., and M. Statman. 1985. The Disposition to Sell Winners too Early and Ride Losers too Long. *Journal of Finance* 40: 777-90.

Shleifer, A., and L.H. Summers. 1990. The Noise Trader Approach to Finance. *Journal of Economic Perspectives* 4, no. 2: 19-33.

Smith, L., and P. Sorenson. 2000. Pathological Outcomes of Observational Learning. *Econometrica* 68, no. 2: 371-98..

Tversky, A., and D. Kahneman. 1974. Judgment Under Uncertainty: Heuristics and Biases. *Science* 185: 1124-31.

Vives, X. 1996. Social Learning and Rational Expectations. *European Economic Review* 40, no. 3-5: 589-601.

Welch, I. 1992. Sequential Sales, Learning, and Cascades. *Journal of Finance* 47, no. 2: 695-732.

Willinger, M, and A. Ziegelemeyer. 1998. Are More Informed Agents Able to Shatter Information Cascades in the Lab? In, *The Economics of Networks: Interaction and Behaviours.* Eds. Cohendet et al.Springer.

Appendix: QRE Estimation Results

Results in Table IV lead us to investigate a modification of the quantal response equilibrium (QRE) developed by McKelvey and Palfrey (1995, 1998). By accounting for the noise associated with the probabilistic choice rule, the QRE yields alternative measures of the posterior public belief. Our approach invokes a rational expectations assumption for the error distribution. There is mixed evidence for this assumption (see Goeree et al., (2004b) and Kubler and Weizsacker (2004a; 2004b)), but several alternative specifications yield similar insights (these results are available upon request).

Consider that the probability of choosing urn A is given by:

$$pr(c_i = A \mid H_i, s_i) = pr\left(\pi_i^A + \varepsilon_i^A > \pi_i^B + \varepsilon_i^B\right) = pr\left(\varepsilon_i > 1 - 2\pi_i^A\right)$$

where $\pi_i^A = pr(A \mid H_i, s_i) * \$W = \$W - \pi_i^B$ and for the stochastic component, $\varepsilon_i = \varepsilon_i^A - \varepsilon_i^B$. For comparability across subject pools we normalize so that *$W* = 1 for both subject pools. If the errors have an extreme value distribution, then the conditional probability of the urn choice is given by the logistic distribution yielding

$$pr(c_i = A \mid H_i, s_i) = \frac{1}{1 + \exp\left(\lambda_i\left(1 - 2\pi_i^A\right)\right)}.$$

The lambda parameter indicates the extent to which noise affects decision outcomes. As $\lambda \to \infty$, the choice converges to the Bayesian outcome. As $\lambda \to 0$, the decisions become purely random. Note that the history under the QRE assumption, and thus π_i^A, is affected by the noise parameter, lambda, from previous choice orders.

In our estimation, we follow Anderson and Holt (1997) and focus on the symmetric data. The QRE results using these data are displayed in Tables A1 and A2. Our results emphasize the fact that not only the numbers of each signal, but the order in which they are revealed have an important impact on behavior. For example, note the posterior probability for order choice 3 in Table A2, the first choice where a cascade may form in the symmetric treatment, when the signal history is AAb (or BBa) (since the AAb and BBa are symmetric, Table A2 reports the results from these sequences as one choice history (AAb); all other symmetric choice sequences are treated similarly). In this case the posterior probability of urn A has dropped from 0.67 for the most likely urn to 0.51 (0.59) for the market professionals (students). Thus, while ignoring one's private information is optimal for both groups, the noise in prior decisions dilutes the strength of the signals, with the market professionals facing essentially a random choice and the probability that urn A is chosen is .539 (.833) for the market professionals (students). In comparison the ABa sequence, which has an identical posterior probability when there is no noise, the posterior 0.64 (0.65) for market professionals (students), and the optimal decision is made uniformly by both subject pools. This difference across the sequences in choice order three highlights the fact that noise in the decision making process dilutes the value of the public signal.

Table A1 Lambda estimate for Quantal Response Equilibrium, Symmetric Gain treatment.

Choice Order	M	C	p
1	4.59	7.12	0.094
2	4.56	27.75	0.012
3	8.67	8.62	0.505
4	3.90	4.99	0.258
5	2.48	6.34	0.026

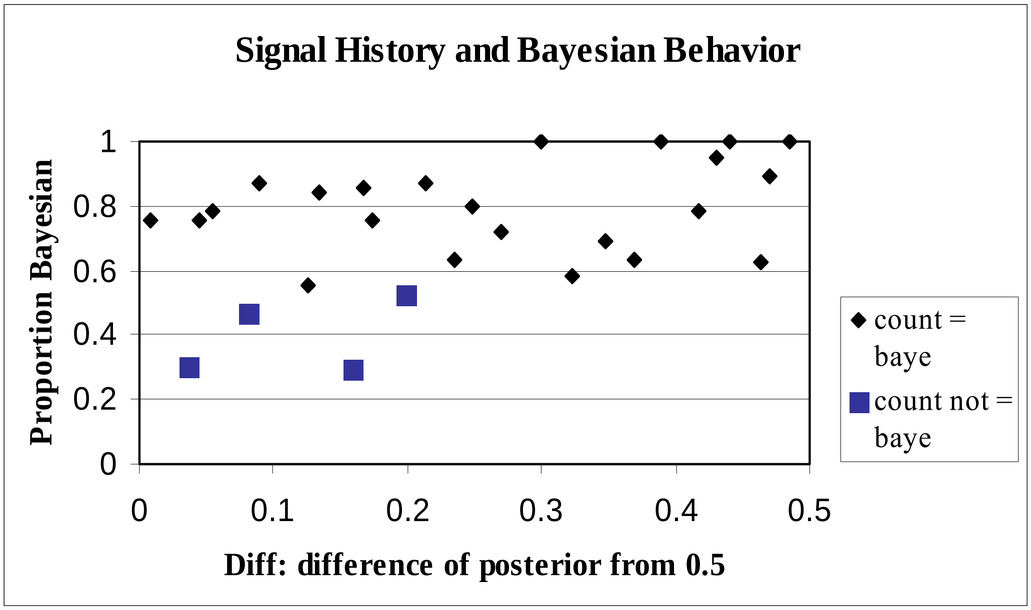
Columns M and C report the lambda parameter for market professionals and college students. Column p reports the one-tailed p-value for the null hypothesis that the lambda parameter does not differ across the two groups.

Table A2 Posterior Probabilities and Choice Probabilities with QRE Decision Error for Both Market Professionals (M) and College Students (C)

		Choice Probability $pr(c = A \mid H, s, \lambda)$			Posterior Probability $pr(\omega = A \mid H, s, \lambda)$			Decisions					
			M	C		M	C	M			C		
Choice Order	History & Signal	Bayes	QRE		Bayes	QRE		A	B	Proportion A	A	B	Proportion A
1	A	1.00	0.82	0.92	0.67	0.67	0.67	37	8	0.822	43	4	0.915
2	Aa	1.00	0.91	0.99	0.80	0.76	0.78	23	3	0.885	27	0	1.000
2	Ab	0.50	0.36	0.15	0.50	0.44	0.47	4	15	0.211	3	17	0.150
3	AAa	1.00	0.99	0.99	0.89	0.81	0.85	14	0	1.000	17	1	0.944
3	AAb	1.00	0.54	0.83	0.67	0.51	0.59	7	6	0.538	12	0	1.000
3	ABa	1.00	0.939	0.92	0.67	0.65	0.64	10	0	1.000	8	0	1.000
3	ABb	0.00	0.049	0.03	0.33	0.32	0.31	1	7	0.125	0	9	0.000

Calculations are for the first three choices of the symmetric gain treatment for market professionals (M) and college students (C), with the choice probability and the posterior probability adjusted for decision error. For comparison, the probabilities assuming a perfect Bayesian equilibrium (Bayes) and the actual decisions are also presented. Due to the symmetry of the treatment, the history and signal combination also represents its complement. For example the row reporting history and signal "ABa" also includes the "BAb" sequences.

Figure 1: Counting Rule Heuristic



The proportion of Bayesian decisions for every realized posterior probability is presented as a data point. The choice histories in which the counting rule and Bayesian posterior yield different predictions are presented as dark squares. All other sequences are presented as black diamonds. Note that the sequences in which Bayesian behavior and the counting rule heuristic make different predictions have a uniformly lower proportion of Bayesian decisions than the others.

Table I. Posterior Probabilities: Symmetric (upper) and Asymmetric (lower, and *italic*) Urns

a \ b	0	1	2	3	4	5	6
0	0.500 *0.500*	0.330 *0.333*	0.200 *0.200*	0.110 *0.111*	0.060 *0.059*	0.030 *0.030*	0.020 *0.015*
1	0.670 *0.545*	0.500 *0.375*	0.330 *0.231*	0.200 *0.130*	0.110 *0.070*	0.060 *0.036*	
2	0.800 *0.590*	0.670 <u>*0.419*</u>	0.500 *0.265*	0.330 *0.153*	0.200 *0.083*		
3	0.890 *0.633*	0.800 <u>*0.464*</u>	0.670 <u>*0.302*</u>	0.500 *0.178*			
4	0.940 *0.675*	0.890 *0.509*	0.800 <u>*0.341*</u>				
5	0.970 *0.713*	0.940 *0.554*					
6	0.980 *0.749*						

Entries represent the posterior probabilities for all possible sequences of draws for both symmetric (upper) and asymmetric (lower, and *italic*) treatments based on choice histories (a, b). The prior probability of an urn is 0.5 in (0,0). Underlined entries in the asymmetric urn are those consistent with counting heuristic sequences that will yield different predictions from the posterior prior probability.

Table II. Experimental Design

Panel A: Ten Market Professional Sessions

	Symmetric Urn		Asymmetric Urn	
	Gains	Losses	Gains	Losses
Number of Sessions	3	1	3	3
Participants in Session	5	5	One with 5, two with 6	6
Total number of Decisions	225	75	255	270
Average Earnings of Participants	$43.20	-$20.80	$39.06	-$22.89

Panel B: Ten Student Sessions

	Symmetric Urn		Asymmetric Urn	
	Gains	Losses	Gains	Losses
Number of Sessions	3	1	3	3
Participants in Session	One with 5, two with 6	5	One with 5, two with 6	5
Total number of Decisions	267	75	255	225
Average Earnings of Participants	$11.61	-$2.80	$11.00	-$6.40

Panel A (B) shows that Market Professionals (Students) were exposed to either the Symmetric or Asymmetric urn and played the game in either the gain or loss domain. The symmetric urn consisted of 3 balls — two *a* and one *b* in Urn A, and one *b* and two *a* in Urn B. The Asymmetric urn consisted of 7 balls — six *a* and one *b* in Urn A, and five *a* and two *b* in Urn B. The number of decisions is a function of the number of players, the number of games, and the number of rounds in each game.

Table III.a. Disaggregated Decision Making across Treatments

Treatment		Bayesian	Cascades (total)	Reverse Cascades	Potential Cascades	Realized Cascades
1. Pooled Data						
C & M n=1647	Proportion Number	0.780 1284	0.149 245	0.033 55	0.268 441	0.556 245/441
2. College Student Treatments (C)						
C n = 822	Proportion Number	.814 669	.178 146	.045 37	.292 240	.608 146/240
SGC n = 267	Proportion Number	.940 251	.157 42	.041 11	.172 46	.913 42/46
SLC n = 75	Proportion Number	.960 72	.067 5	.013 1	.080 6	.833 5/6
AGC n = 255	Proportion Number	.682 174	.251 64	.051 13	.451 115	.557 64/115
ALC n = 225	Proportion Number	.764 172	.155 35	.053 12	.324 73	.480 35/73
3. Market Professional Treatments (M)						
M n = 825	Proportion Number	.745 615	.120 99	.021 18	.244 201	.493 99/201
SGM n = 225	Proportion Number	.818 184	.098 22	.022 5	.142 32	.688 22/32
SLM n = 75	Proportion Number	.867 65	.147 11	.067 5	.213 16	.688 11/16
AGM n = 255	Proportion Number	.714 182	.133 34	.008 2	.275 70	.486 34/70
ALM n = 270	Proportion Number	.681 184	.133 32	.022 6	.307 83	.385 32/83

The *Bayesian* column represents the total number of decisions (and proportion) that were consistent with a perfect Bayesian equilibrium. *Cascade* decisions (those which are Bayesian but private information ignored) and *reverse* cascades (same as cascades but the wrong inference of the underlying state takes place) occupy the next two columns. The *potential* cascades category represents the proportion (and number) of cascades that could have occurred when it was possible to make one, and the *realized* cascades category represents the proportion of those potential cascades that were actually realized. "n" = number of decisions. Treatment codes are S = symmetric, A = asymmetric, G = gain, L = loss, C = college student, M = market professional.

Table III.b Decision Making by Counting Rule Predictions (Asymmetric Treatments)

		Bayesian	Cascades (total)	Reverse Cascades	Potential Cascades	Realized Cascades
1. Pooled Data						
C & M n = 1005	Proportion Number	.709 712	.164 165	.033 33	.339 341	.477 165/341
Count = Baye n = 843	Proportion Number	.759 640	.152 128	.024 20	.267 225	.565 128/225
Count ≠ Baye n = 162	Proportion Number	.444 72	.228 37	.080 13	.716 116	.313 37/116
2.College Student Treatments (C)						
C n = 480	Proportion Number	.721 346	.206 99	.052 25	.392 188	.527 99/188
Count = Baye n = 412	Proportion Number	.760 313	.189 78	.036 15	.325 134	.582 78/134
Count ≠Baye n = 68	Proportion Number	.485 33	.309 21	.147 10	.794 54	.389 21/54
3. Market Professional Treatments (M)						
M n = 525	Proportion Number	.697 366	.126 66	.015 8	.291 153	.431 66/153
Count = Baye n = 431	Proportion Number	.759 327	.116 50	.011 5	.211 91	.550 50/91
Count ≠Baye n = 94	Proportion Number	.415 39	.170 16	.032 3	.660 62	.258 16/62

The *Bayesian* column represents the total number of decisions (and proportion) that were consistent with Bayesian updating. *Cascade* decisions (those which are Bayesian but private information ignored) and *reverse* cascades (same as cascades but the wrong inference of the underlying state takes place) occupy the next two columns. The *potential* cascades category represents the proportion (and number) of cascades that could have occurred when it was possible to make one, and the *realized* cascades category represents the proportion of those potential cascades that were actually realized. "n" = number of decisions. Treatment codes are S = symmetric, A = asymmetric, G = gain, L = loss, C = college student, M = market professional.

Table IV. Bayesian Decisions: Probit Model

Dependent variable: baye	4a: Pooled Model (combining market professionals and students) n = 1647 Pr(Baye=1)=.818			4b. Student Model n = 822 Pr(Baye=1)=.868			4c. Market Professionals Model n = 825 Pr(Baye=1)=.772		
Ind. Variables:	Marginal Effect	z stat	P>|z|	Marginal Effect	z stat	P>|z|	Marginal Effect	z stat	P>|z|
Diff	0.655	5.53	0.000	0.769	5.06	0.000	0.546	3.11	0.002
Heurist	-0.232	-4.95	0.000	-0.161	-2.51	0.012	-0.284	-4.47	0.000
Gain	-0.030	-1.13	0.259	-0.060	-2.19	0.028	0.015	0.34	0.737
Sym	0.102	3.64	0.000	0.145	4.88	0.000	0.037	0.79	0.430
Trader	-0.060	-2.32	0.020	-	-	-	-	-	-
order_2	-0.023	-0.66	0.507	0.019	0.54	0.590	-0.084	-1.42	0.157
order_3	-0.041	-1.06	0.291	0.017	0.42	0.673	-0.120	-1.86	0.063
order_4	-0.120	-2.91	0.004	-0.052	-1.13	0.261	-0.205	-3.12	0.002
order_5	-0.035	-0.95	0.343	0.017	0.46	0.649	-0.107	-1.71	0.087
order_6	-0.040	-0.83	0.408	-0.035	-0.56	0.577	-0.080	-1.05	0.294
	Log Likelihood: -766.487, Wald $\chi^2_{(10)}$ = 141.03, Prob > $\chi^2_{(10)}$ = 0.000			Log Likelihood: -328.95, Wald $\chi^2_{(10)}$ = 87.77, Prob > $\chi^2_{(9)}$ = 0.000			Log Likelihood: -427.283, Wald $\chi^2_{(10)}$ = 68.91, Prob > $\chi^2_{(9)}$ = 0.000		

The dichotomous dependent variable in all three probit models (pooled, student, and market professional) is coded one for a decision consistent with the Bayesian posterior and zero otherwise. Independent variables include *diff*, which is $\left|prob\left(urn = A\right) - .5\right|$, where the $prob\left(urn = A\right)$ is the posterior probability arising from the combination of public and private information at the disposal of each decision maker. The variables *gain*, *sym*, and *trader* (in the case of the pooled model) are dichotomous and distinguish the treatments. *Heurist* is a dummy variable equal to one for the non-counting rule sequences and zero for all others. *order_x (where x=2,..6)* is a categorical variable indicating where in the round of play the decision was made. The Wald statistic tests the null hypothesis that all coefficients are zero.

Table V. Winning Decisions: Probit Model

Dependent variable: *win*	5a: Pooled Model (combining market professionals and students) n = 1647 Pr(Win=1)=.702			5b. Student Model n = 822 Pr(Win=1)=.729			5c. Market Professionals Model n = 825 Pr(Win=1)=.670		
Ind. Variables:	Marginal Effect	z stat	P>|z|	Marginal Effect	z stat	P>|z|	Marginal Effect	z stat	P>|z|
Diff	1.070	7.85	0.000	1.039	5.51	0.000	1.053	5.55	0.000
Heurist	-0.008	-0.19	0.846	-0.009	-0.14	0.886	-0.050	-0.31	0.754
Gain	0.083	2.65	0.008	0.113	2.52	0.012	0.158	1.26	0.207
Sym	-0.028	-0.81	0.418	-0.013	-0.31	0.759	-0.122	-0.86	0.388
Trader	-0.033	-1.10	0.272	-	-	-	-	-	-
Round	0.005	1.95	0.051	0.003	0.68	0.499	0.022	2.06	0.040
order_2	0.008	0.22	0.824	0.014	0.29	0.774	0.008	0.05	0.960
order_3	0.019	0.49	0.627	0.016	0.30	0.767	0.073	0.45	0.656
order_4	0.048	1.26	0.206	-0.001	-0.02	0.988	0.293	1.79	0.074
order_5	0.057	1.53	0.126	0.081	1.64	0.102	0.010	0.62	0.534
order_6	0.063	1.33	0.184	0.046	0.63	0.528	0.192	0.96	0.337
	Log Likelihood: -964.06, Wald $\chi^2_{(11)}$ = 99.52, Prob > $\chi^2_{(11)}$ = 0.000			Log Likelihood: -462.25, Wald $\chi^2_{(10)}$ = 54.39, Prob > $\chi^2_{(10)}$ = 0.000			Log Likelihood: -498.55, Wald $\chi^2_{(10)}$ = 48.30, Prob > $\chi^2_{(10)}$ = 0.000		

The dichotomous dependent variable in all three probit models (pooled, student, and market professional) is coded one for a decision that correctly predicts the underlying state and zero otherwise. Independent variables include *diff*, which is $\left|prob\left(urn = A\right) - .5\right|$, where the $prob\left(urn = A\right)$ is the posterior probability arising from the combination of public and private information at the disposal of each decision maker. The variables *gain*, *sym*, and *trader* (in the case of the pooled model) are dichotomous and distinguish the treatments. *Heurist* is a dummy variable equal to one for the non-counting rule sequences and zero for all others. *Round* represents a time trend that increases from 1 to 16 with each completed play of the cascade game. *Order_x (where x=2,..6)* is a categorical variable indicating where in the round of play the decision was made. The Wald statistic tests the null hypothesis that all coefficients are zero.

Table VI. Cascade Formation: Probit Model

Dependent variable: cascade	6a: Pooled Model (combining market professionals and students) n = 416 Pr(Cascade=1)=.588			6b. Student Model n = 226 Pr(Cascade=1)=.676			6c. Market Professionals Model n = 190 Pr(Cascade=1)=.493		
Ind. Variables:	Marginal Effect	z stat	P>|z|	Marginal Effect	z stat	P>|z|	Marginal Effect	z stat	P>|z|
Diff	0.861	0.75	0.453	-0.136	-0.08	0.939	1.689	1.13	0.259
Othb_min	-0.014	-0.09	0.924	-0.572	-2.54	0.011	0.469	2.11	0.035
Heurist	-0.354	-4.35	0.000	-0.331	-2.49	0.013	-0.394	-3.95	0.000
Gain	0.133	1.67	0.095	0.021	0.19	0.846	0.120	1.12	0.261
Sym	0.126	1.03	0.303	0.389	4.09	0.000	-0.200	-1.16	0.246
Trader	-0.111	-1.44	0.151	-	-	-	-	-	-
order_2	-0.146	-1.13	0.260	0.064	0.39	0.696	-0.332	-2.33	0.020
order_3	0.078	0.72	0.469	0.167	1.24	0.214	0.039	0.24	0.810
order_4	0.042	0.39	0.696	0.048	0.33	0.744	0.066	0.44	0.658
order_5	0.233	2.35	0.019	0.169	1.21	0.225	0.308	2.13	0.033
	Log Likelihood: -245.22, Wald $\chi^2_{(10)}$ = 46.24, Prob > $\chi^2_{(10)}$ = 0.000			Log Likelihood: -125.20, Wald $\chi^2_{(9)}$ = 27.78, Prob > $\chi^2_{(9)}$ = 0.001			Log Likelihood: -111.06, Wald $\chi^2_{(9)}$ = 27.42, Prob > $\chi^2_{(9)}$ = 0.0012		

The dichotomous dependent variable in all three probit models (pooled, student, and market professional) is coded one for a cascade decision and zero otherwise. Independent variables include *diff*, which is $\left|prob\left(urn=A\right)-.5\right|$, where the $prob\left(urn=A\right)$ is the posterior probability arising from the combination of public and private information at the disposal of each decision maker. The variables *gain* and *trader* (in the case of the pooled model) are dichotomous and distinguish the treatment/subject type. *Othb_min* is the proportion of Bayesian decisions by the individual with the lowest proportion among all preceding the decision maker. The othb_min is calculated in each round of the game to include only those decisions that have already occurred. *Heurist* is a dummy variable equal to one for the non-counting rule sequences and zero for all others. *Order_x* is a categorical variable indicating where in the round of play the decision was made. Note: Because the othbys variable is not applicable for those in the first round or first in choice order in subsequent rounds (they do not observe others' decisions in the current round), these observations are excluded. This results in the exclusion of 25 of the 441 potential cascades. The order_2 dummy variable is also excluded and choice order two serves as the baseline to which others are compared. The Wald statistic tests the null hypothesis that all coefficients are zero.

Table VII. Bayesian and Cascade Behavior of Traders

	7a: Trader subset of CBOT Market Professionals n = 227 Dependent Variable: Baye Pr(Baye=1)=.745			7b.Trader subset of CBOT Market Professionals n = 66 Dependent Variable: Casc Pr(Casc=1)=.388		
Ind. Variables:	Marginal Effect	z stat	P>|z|	Marginal Effect	z stat	P>|z|
Diff	0.467	1.34	0.181	3.710	0.65	0.517
Heurist	-0.391	-3.05	0.002	0.001	0.05	0.997
Gain	-0.023	-0.24	0.81	0.163	0.59	0.552
Sym	0.009	0.08	0.938	0.498	0.79	0.432
order_2	0.061	0.64	0.523	n/a	n/a	n/a
order_3	-0.029	-0.28	0.778	-0.185	-0.48	0.629
order_4	-0.095	-0.86	0.392	-0.219	-0.68	0.498
order_5	-0.024	-0.22	0.828	0.474	1.22	0.221
order_6	0.089	0.67	0.504	0.229	0.36	0.721
Intensity	0.004	2.44	0.015	-0.029	-1.91	0.056
Gender	0.069	0.58	0.561	-0.955	-0.2	0.838
Experience (yrs)	-0.001	-0.18	0.859	0.011	0.34	0.735
Income	0.013	0.55	0.582	0.394	1.58	0.115
Overnight	-0.173	-2.03	0.042	-0.804	-2.24	0.025
	Log Likelihood: -93.93, Wald $\chi^2_{(14)}$ = 52.74, Prob > $\chi^2_{(14)}$ = 0.0000, Pseudo R-squared = 0.22			Log Likelihood: -15.62, Wald $\chi^2_{(13)}$ = 34.20, Prob > $\chi^2_{(13)}$ = 0.0011, Pseudo R-squred = .52		

The dichotomous dependent variable in panel a. is coded one for a decision consistent with the Bayesian posterior and zero otherwise. For panel b. cascade formation is indicated by a one and cascade failure by a zero. Independent variables include *diff*, which is $\left| prob\left(urn = A\right) - .5 \right|$, where the $prob\left(urn = A\right)$ is the posterior probability arising from the combination of public and private information at the disposal of each decision maker. The variables *gain* and *sym* are dichotomous and distinguish the treatments. *Heurist* is a dummy variable equal to one for the non-counting rule sequences and zero for all others. *Order_x (where x=2,..6)* is a categorical variable indicating where in the round of play the decision was made. *Intensity* reflects the level of trading intensity among participants, measured as the number of contracts traded per day. *Gender* is 1 for female and zero for male. *Experience* (years) and *income* (dollars) and *overnight* (one for holding overnight positions, zero for daytrader) are additional control variables.

Table VIII. Posterior Probability Urn is *A* and Proportion of Bayesian Decisions (Counting Rule Sequences in Bold)

a b	0	1	2	3	4	5	6
0		0.33 *0.85*	0.20 *1.00*	0.11 *1.00*	0.06 *1.00*	0.03 *0.89*	0.02 *1.00*
1	0.55 *0.76*	0.38 *0.56*	0.23 *0.72*	0.13 *0.64*	0.07 *0.95*	0.04 *0.63*	
2	0.59 *0.87*	**0.42** ***0.46***	0.26 *0.63*	0.15 *0.69*	0.08 *0.79*		
3	0.63 *0.84*	**0.46** ***0.30***	**0.30** ***0.52***	0.18 *0.58*			
4	0.67 *0.76*	0.51 *0.76*	**0.34** ***0.29***				
5	0.71 *0.87*	0.55 *0.78*					
6	0.75 *0.80*						

The number of A and B signals are given in the first row and first column, respectively. The pairs of numbers within an (a,b) pair represent the Bayesian posterior (upper number) and the proportion of Bayesian decisions (lower number and in *italics*). Those in bold type are the *counting heuristic* sequences. Thus (2,1) has a posterior probability of 42 percent that the urn is *A* (diff=0.08). Forty-six percent made the Bayesian decision in this case. By contrast the (2,0) sequence (in which diff=0.09) has a posterior probability of 0.59, and 87 percent of those decisions were Bayesian.

Table IX: Bayesian Behavior and Potential Cascade: By Counting Rule Sequences

		No Potential Cascade	Potential Cascades	Total
All n = 1005	Proportion Number	.82 541/657	.48 166/348	.70 707/1005
Non-Counting Rule n = 843	Proportion Number	.83 512/618	.57 128/225	.76 640/843
Counting Rule n = 162	Proportion Number	.74 29/39	.31 38/123	.41 67/1624

The proportion of Bayesian decisions both when a cascade is possible and when one is not for both counting rule and non-counting rule sequences in the asymmetric treatments are provided in the table. When there is no potential cascade the proportion of Bayesian decisions (.74) is the proportion in which one follows the private signal. When there is a potential cascade (1-.31=.69) is the proportion of decisions that follow the private signal.

Endnotes

[1] Herding is a more general phenomenon than an informational cascade though both result in behavioral conformity. The homogeneity of a herd may arise through other than informational means such as payoff externalities, preferences for conformity, or sanctions. A comprehensive taxonomy of herd behavior is developed by Hirshleifer and Teoh (2003); Smith and Sorenson (2000). Devenow and Welch (1996), and Bikhchandani and Sharma (2000) also discuss alternative sources of herd behavior and review the extant literature.

[2] It has been argued, also, that information cascades can explain a large variety of social behaviors such as fashion, customs, and rapid changes in political organization. Anderson (1994), Banerjee (1992), Bikhchandani et al. (1992; 1998), and Welch (1992) discuss a variety of interesting examples. A number of historical anecdotes can be found in MacKay (1980) and Garber (2000).

[3] Fama (1998) discusses the interpretation of empirical results as evidence of irrational behavior.

[4] Combined with the insights gained from the models of Barberis et al. (1998), Daniel et al. (2001), and from Hirshleifer (2001), our results indicate that the ability of the strength of evidence and weight of evidence to have a differential impact on asset pricing is a potentially powerful phenomenon.

[5] As discussed below, our experimental environment makes use of a binary signal and binary state and a fixed payoff regardless of the history of announcements. Avery and Zemsky (1998), Lee (1998), Chari and Kehoe (2004), and Cipriani and Guarino (2005a) explore more general settings in which variable pricing reduces but does not eliminate the potential for information cascades. Chamley (2004) provides a comprehensive review of rational herding models.

[6] The ability of humans to reason in a Bayesian manner seems to depend on how information is presented. Studies that present base rates as percentages often show that we are poor "intuitive statisticians" (Tversky and Kahneman (1974)). Decisions tend to be more consistent with Bayesian rationality when individuals experience probability distributions through repeated exposure (see, e.g., Gigerenzer and Murray (1987)). Our experiment is consistent with protocols that have been shown to give Bayesian decision making its best chance.

[7] $\pi_A - \pi_B = \frac{\$W}{3}$ in the gain treatments, after an initial *a* signal, where \$W is the win amount. Treatments over gains and losses yield identical predictions (i.e., expected losses are minimized by picking the most probable urn).

[8] A second A announcement could arise in this setting if the second subject received a *b* signal. We consider that announcing an A given the history Ab to be inconsistent with Bayesian

rationality, although alternative interpretations are possible. Since the posterior probability is .5 in this case a tie-breaking rule must be invoked. We follow Anderson and Holt (1997) in assuming individuals who are indifferent announce their own signal. This is sensible if individuals recognize the possibility of decision error in previous announcements. Alternative tie-breaking rules include random choice as in BHW (1992) and a "non-confident" rule in which one ignores one's own information (Koessler and Zieglemeyer 2000). In our treatments the Anderson and Holt rule is followed 81 percent of the time, with most of the deviations occurring in the early rounds of play.

[9] Our experimental instructions are available upon request. Note that Anderson and Holt (1997) found that cascades formed in roughly 70 percent of the rounds in which they were possible. Deviations from Bayesian cascade formation occurred most often when a simple counting rule gave a different indication of the underlying state. Extensions to the experimental literature have introduced relevant complications to the cascade process that include costly information, endogenous sequencing of choice order, collective decision making, expanded signal spaces, and payoff externalities (Celen and Khariv (2004, 2005); Cipriani and Guarino (2005b); Drehmann et al. (2005); Huck and Oechssler (2000); Hung and Plott (2001); Kubler and Weizsacker (2004a); Noth and Weber (2003); Sgroi (2003); Willinger and Ziegelmeyer (1998)).

[10] CBOT officials suggested that designing a 30-minute game with an expected average payout of approximately $30 was more than a reasonable approximation of an average trader's earnings for an equivalent amount of time on the floor. In our experiments the median earnings for the market professionals were slightly in excess of this amount and therefore likely to be salient.

[11] To ensure that subjects departed with positive money balances we had both subject pools participate in other unrelated games during the experimental session.

[12] In the discussion that follows we use the term "Bayesian decision" to mean that the decision is consistent with the predictions of perfect Bayesian equilibrium.

[13] Thirteen of the fourteen who were perfectly consistent with Bayesian rationality were in the symmetric urn treatment. One market professional was perfectly Bayesian in the asymmetric setting.

[14] We will see below that there are differences between the symmetric and asymmetric treatments even after controlling for the counting rule sequences, as a result we do not pool the symmetric results in this table.

[15] Anderson and Holt (1997) found that the rate of Bayesian behavior in the non-counting rule sequences was 50 percent, comparable to our student population rate of 49 percent and close to the pooled rate of 44 percent.

[16] There are ten session level observations for each subject pool as summarized in Table II.

[17] The posterior, and thus the *diff* variable, remains constant once a cascade has formed, unless a decision breaking the cascade is observed.

[18] Results are robust to inclusion of a time trend for round or time dummies (categorical time dummy variables for each round of play). Below we discuss further our evidence of learning.

[19] A Chow test rejects the null hypothesis of no differences across the subject pools at the $p < 0.01$ level.

[20] Several specifications of the model of Bayesian decision making were tested for learning and no such effect was found.

[21] Support for the significant differences between subject pools found in the parametric results is also found in non-parametric (Mann-Whitney) tests.

[22] Due to the small number of sessions at the individual treatment level, p-values for the Mann-Whitney test are reported for observations aggregated at the individual participant level.

[23] We use a Wilcoxon matched pairs test with the proportion of Bayesian decisions the variable of interest, aggregated at the session level. The diff variable for the *counting rule* sequences are in the range from 0.0 to 0.2, and all other sequences with diff variables in this range are included for the paired comparison. Using data from the twelve asymmetric sessions we find that the *counting rule* sequences reflect less Bayesian decision making despite roughly equivalent diff scores at $p<.01$.

[24] There has been a long and important debate on the relevance of cascade models for financial markets (Vives (1996)). Avery and Zemsky (1998) showed that the introduction of variable prices to the BHW model can eliminate informational cascades (herding in their terminology) under certain conditions. Lee (1998), Chari and Kehoe (2004) and Cipriani and Guarino (2005a) demonstrated the potential for informational cascades in the variable price setting by introducing transaction costs, endogenous timing, and preference heterogeneity.

[25] One mechanism through which cascades might arise is, in the jargon of the trading floor, when local traders "lean on" large orders by trying to enter the market on the same side and at the same price. Locals who trade alongside an institutional order accumulate a position knowing that they can transact with the institution and avoid a loss. The decision process associated with deciding to trade with the institution has the character of a fixed price cascade. Berkman (1996) discusses in the context of option markets how market makers supply liquidity in the presence of large fixed price orders. Chamley and Gale (1994) introduce endogenous timing in a cascade model that predicts the least informed would trade later, and potentially face adverse prices.

[26] Drehmann et al. (2005) test experimentally a version of the Avery and Zemsky (1998) model that omits event and composition uncertainty and find behavior fairly consistent with its predictions, though subject to decision error and contrarian behavior.