MPhil Finance Thesis

***Investigating Herding Behavior in an Experimental Setup*[1]**

Vinayak Dalmia*
*Judge Business School, University of Cambridge, CB2 1AG, UK.

June 01, 2009

Abstract

An "information cascade" occurs when initial decisions coincide in a way that it is optimal for each of the subsequent individuals to ignore her or his private signals and follow the established pattern. Moscarini et. al (1996) developed a social learning model hitherto unexplored in an experimental setting. They considered the consequences on learning if the state of the world is changing stochastically over time. Such a changing world scenario has many parallels in the functioning of financial markets where agents with incomplete information pass on signals and there is always a chance that the underlying world may change between time periods. Using the experimental methodology developed by Anderson and Holt (1997), the theoretical propositions were put to test. The broad results from this experiment were not inconsistent with the theory. Further extensions for new research are also discussed.

[1] This research was supported by a funding grant from the SOCIALNETS project. I wish to thank Jon Crowcroft for his constant guidance, patience and support. I also thank Nishanth Sastry and Eiko Yoneki for creating the experimental software; without which this endeavour would not be possible. I am also grateful to Marco Ottaviani for helpful discussions and advice. Last, but not least I thank the Economics Faculty at the University of Cambridge for allowing me to use their computer laboratory. Raw Data is available from the researcher upon request. All mistakes are my own.

I. Introduction

There are uncountable social and economic situations in which our decision making is influenced by what others around us are doing. Think everyday life: we often decide on which restaurants, jobs, and schools (sometimes even degrees) to choose based on how popular they seem to be. Moreover one of the most striking regularities of human society is localized conformity. Japanese act Japanese, Indians act Indian and British act British.

Economists have proposed several different theories to explain the existence of social conformity. These include benefits from conformity for its own sake[2], sanctions imposed on deviants[3], externalities[4], and *social learning*, which refers to situations in which individuals learn by way of observing the behavior of others.

Among these many theories, social learning alone explains not only why a society settles on a single pattern of behavior but also why mass behavior may be idiosyncratic, error-prone and fragile, in the sense that small shocks may cause behavior to shift suddenly and dramatically.

Additionally an important characteristic of these social learning models and of the 'real world' is an asymmetric distribution of information. It is natural to associate symmetry with the divine[5] and an almost perfect world. The economic system that we pervade makes for a compelling antithesis of that very divine. Moreover, this could not be more prevalent than in the distribution of information in the economy. With an asymmetric distribution, there is reason to be 'nosy' and infer more about the underlying state of the world from the actions of others.

Thus the social learning literature[6] analyzes an economy where a sequence of individuals is supposed to make an once-in-a-lifetime decision under incomplete and asymmetric information. The typical conclusion is that, despite the asymmetry of information, eventually every individual imitates her predecessor, even though she would have chosen a different action on the basis of her own information alone. In

[2] Keynes (1965 p. 158) notes, "Worldly wisdom teaches that it is better for reputation to fail conventionally than to succeed unconventionally". Additionally See Jones (1984).
[3] See Akerlof (1980), Bendor and Mookherjee (1987), Coleman (1987), Hirshleifer and Rasmusen (1989), and Kuran (1989).
[4] See Arthur (1989), Farrell and Saloner (1985), Katz and Shapiro (1985), and Karni and Schmeidler (1989).
[5] Think Egyptians and their pyramids, Greeks and the Parthenon, Fibonacci and his magical series etc.
[6] Banerjee (1992) and Bikhchandani et. al (1992) introduced the basic concepts and stimulated further research in this area. Among others, Lee (1993), Chamley and Gale (1994), Gul and Lundholm (1995) and Smith and Sørensen (2000) provide extensions.

this sense, individuals rationally (in the Bayesian sense) 'ignore' their own information and 'follow the herd'[7].

It is important to note that almost all learning models assume *perfect information*: every agent is assumed to be informed about the entire history of actions that have already been taken. The individual is thus comparing her information with that of a large number of other individuals. This is an obvious shortcoming of these models as in reality, individuals have *imperfect information*. If each individual observes the actions of only a small number of other individuals, it is not clear that 'herd behavior' will arise.

Celen and Kariv (2004)[8] discuss an imperfect-information model. The model provides outcomes that are quite distinct from and in some ways more extreme than the perfect-information model. They claim that such a model provides "better answers" to such questions as: Why do markets move from 'boom' to 'crash' without settling down? Why is a technology adopted by a wide range of users more rapidly than expected and then, suddenly, replaced by an alternative? What makes a restaurant fashionable overnight and equally unexpectedly unfashionable, and so on?

The most important distinction between perfect and imperfect information is that learning under perfect information has a martingale property that allows one to establish convergence of beliefs and actions. Under imperfect information, however, the learning process is not a martingale. Celen and Kariv (2004) show that this has an important implication: beliefs and actions are not convergent but cycle forever. Despite this instability, however, over time, private information is increasingly "ignored" and decision makers become increasingly likely to imitate their predecessors. Consequently, the model predicts longer and longer periods of uniform behavior, punctuated by infrequent switches. Under perfect information, social learning eventually ceases as individual behavior becomes purely imitative and hence is uninformative.

[7] The reasoning is as follows: agents observe private signals of some underlying state and make public decisions. Subsequent decision makers need to not only calibrate the implications of their private signals, but also what the decisions of there predecessors implied. They face a dilemma if their own private signal is indicative of something, which is contrary or unlikely given the previously observed decisions. Bayesian reasoning will eventually lead to ignoring your own private signals and following others. The analysis becomes trivial in the absence of conflict between private signal and previous announcements.

[8] Smith and Sørensen (1996) also relax the perfect information assumption. The two papers however differ in two ways. First, Celen and Kariv (2004) show that behavior can be drastically different under perfect and imperfect information. Second, they are able to describe not only the long-run outcomes but also the medium-run properties and behavior in case of divergence.

Thus, the perfect- and imperfect-information versions of the model share the conclusion that individuals can, for a long time, make the same choice. The important difference however is that, whereas in the perfect-information model a herd is an "absorbing state", in the imperfect-information model, there are continued, occasional and quick shifts in behavior.

In spite the obvious shortcomings of the perfect information setting, this paper will restrict its attention to such a framework as the theoretical and experimental methodology of an imperfect setting are beyond its scope.

Finally, it is important to note that rational inferences and subsequent cascading need not always be a good thing from a welfare perspective. Moreover herd behaviour has been deemed pathological because erroneous outcomes may occur despite individual rationality, and they may in fact be the norm in certain circumstances. Particularly interesting in this regard is the possibility of "reverse cascades" (Anderson and Holt, 1997) or what is also referred to as "herd externalities" (Banerjee, 1992). A reverse cascade occurs when initial decision makers are unfortunate to observe private signals that indicate the incorrect state, and a large number of followers may join the resulting pattern of "mistakes," despite the fact that their private signals are more likely to indicate the correct state. This fact is reiterated by Choi et al (2005): "since actions aggregate information poorly, despite the available information, herds need not obtain an optimal action".

Such behaviour might also arise in financial markets, where trading decisions comes across a ticker tape in sequence. Even if early traders have no inside information, others may incorrectly infer that the previous trades reveal private information[9]. Camerer and Weigelt (1991) report some trading sequences in laboratory experiments that fit this pattern. Such behaviour can easily crop up in the real world as well. The problem seems to be that when there is so little available information about a particular event, the crowd ends up focusing on the few voices who supposedly know what is going on (i.e. the experts) – the near universal belief that Judge Clement would be Bush's nomination for the Supreme Court was wrong.

This paper proceeds as follows. Section II discusses some examples from the real world as well sampling the existing literature that would fit rational herding due to social learning. Section III starts with a survey and follows up with some discussion

[9] In this way, some randomness in initial trades might create a price movement that is not supported by fundamentals, as in a reverse cascade. Also see Welch (1992).

on rationality. The theoretical model is outlined in Section IV. Section V describes the experimental setup and discusses the results from the experiment. Section VI discusses a post-experiment survey, which was conducted with the participants to better ascertain their beliefs and the general rules used for decision-making. Additionally some econometric methodology that could have been used is also discussed. Section VII looks into further extensions for future research. Section VIII has a brief discussion on the criticisms of the methodologies used. The final section contains a conclusion.

II. Examples

Over the last ten years rational herding has become an important tool in analysing how and why economic agents learn through observation in groups. Economic agents constantly learn from others, through chat, newspapers, and typically in financial markets, also through observing price movements, or buy and sell decisions by others.

A. Financial Markets

Perhaps the most relevant examples exist in the world of finance. It was suggested by Keynes (1936), for example, that this is how investors in asset markets often behave (the "beauty contest" example). For example, when Warren Buffett buys a stock or commodity this news affects its price.

Some investors are influenced by broker cold-calls with statements that famous investors are holding a stock [see Lohse (1998) and Davis (1991)]. E.g.: When news came out that Warren Buffett had bought approximately 20% of the 1997 world silver output, according to The Economist (1998) silver prices were sent 'soaring'. Moreover, when Buffett's filings reported that his shareholding in American Express and in PNC Bank have increased, these shares rose by 4.3% and 3.6% respectively [Obrien and Murray (1995)]. According to Sandler and Raghavan (1996), 'Whether Warren Buffett has been right or wrong about a stock, investors don't like to see him get out if they're still in. There is also a family of literature on evidence regarding herding in trades; see Lakonishok et al. (1992), Grinblatt et al. (1995), Wermers (1999). Bikhchandani and Sharma (2001) critically review alternative empirical measures of herding.

Investors are also influenced by private conversations with peers. For example, Fung and Hsieh (1999) state that ‘a great deal of hedge fund investment decisions are still based on ‘‘recommendations from a reliable source’’ ’. There is also evidence that investors are influenced by implicit endorsements, as with default settings for contributions in 401(k) plans; see Madrian and Shea (2000).

There have also been studies analysing herding behaviour amongst analysts and forecasters[See Trueman (1994), Ashiya and Doi (2001), Ehrbeck and Waldmann(1996), Givoly and Lakonishok (1984), Brown et al. (1985) and Graham (1999) amongst others]. Lastly, D’Arcy and Oh (1997) apply the cascade model of insurance to explain the Lloyd’s of London experience. The patterns of Lloyd’s profitability fit the cascade model.

B. Financial Theory

There is also a vast literature on learning in prices [e.g. Grossman and Stiglitz, (1976), Glosten and Milgrom (1985) and Kyle (1985), and Christie and Huang (1995)], real options [Chamley and Gale (1994); see also Hendricks and Kovenock (1989), Bhattacharya et al. (1986), Zhang (1997) and Grenadier (1999)], retail deposits [Kahn et al. (2002)], financial innovation [Persons and Warther (1997)], real estate and stock markets [Zeira (1999)], corporate conservatism [Zwiebel (1995)], and delegated portfolio management [Maug and Naik (1995)].

Nelson (2001) offers a model of IPOs in which the decision to go public is more likely to be associated with informational cascades than the decision to hold off. This paper is off special relevance as similar to Moscarini et al. (1998) it allows the underlying value (market conditions or the relative valuation of a company) to change over time in a Markov fashion. This allows her to analyze decisions in a changing environment such as IPO decisions, analyst recommendations, the choice of dining in a restaurant etc. Nelson[10] goes on to state other financial applications of the changing world model such as adoption of new productions ranging from LBOs, ABS and convertible debt. Another similar application is Perktold (1996) where a Macroeconomic model is presented in which agents receive signals wherein the underlying is a binary Markov chain. His model predicts recurring informational cascades.

It is important to note that some negativity in the literature [Avery and Zemsky,

[10] An interesting area for further research mentioned in this paper is the impact the ordering of firms would have on the total percentage of firms that go public.

(1998)] had stalled the development of financial herding for some years. However, recent work [Chari and Kehoe (2004) and Park and Sabourian (2004)] suggests that herding might be a prime explanation for persistent price spikes and crashes, and often cited "crazy" behaviour on the markets.

C. Financial Experiments

Cascade-like behavior is sometimes even observed in asset market experiments in which some investors are informed about a state of nature and others are not [Plott and Sunder (1982)]. In these markets, the uninformed tend to follow the trading patterns of the insiders well enough to minimize earnings differences between the two groups. There have also been studies explaining information aggregation [Plott and Sunder (1988)] and bubbles [Smith et al (1988)] in asset markets. Lastly Sunder (1995) provides an exhaustive survey of experiments in asset markets.

D. Other examples

Other applications are suggested by the majority-voting treatment of Hung and Plott (2001). Some experiments that simulate sequential jury voting have been conducted, and strong patterns of cascade-like conformity are observed in many cases.

The 1976 US presidential campaign provides a classic case study in this respect as well[11]. Little known Jimmy Carter achieved an important early success by concentrating his efforts towards securing the Democratic nomination in the Iowa caucus (which preceded the first primary in New Hampshire). “Super Tuesday”, in which many southern states coordinated their primaries on the same date, was an attempt to avoid the consequences of sequential voting. Moreover, in a study of U.S. presidential nomination campaigns, Bartels [(1988), pg. 110] discusses “cue-taking”, in which an individual’s belief about a candidate are influenced by the decision of others. Voters are also known to be influenced by opinion polls to vote in the direction that the poll predicts will win [Cukierman (1989)][12].

There is evidence of such behaviour even in Zoology. There is evidence of imitative behaviour transmission among animals, especially in territory choice, mating and foraging [Galef (1976), pg. 78)].

[11] Mckelvey and Ordeshook (1985) discuss a model where opinion polls convey information that causes bandwagons.
[12] For an interesting counter-example see DellaVigna and Kaplan (2006).

Taylor (1979) and Robin (1984) discuss numerous surgical fads and epidemics of treatment-caused illnesses ("iatroepidemics"). Some operations that have come and gone in popularity are tonsillectomy, elective hysterectomy, internal mammary ligation, and ileal bypass. This point is succinctly made by Burnum [(1987), pg. 1222] when referring to physicians who, "like lemmings, episodically and with a blind infectious enthusiasm [push] certain diseases and treatments primarily because everyone else is doing the same".

Adoption of a scientific theory can also cascade. Very few people have carefully examined the evidence that the earth is round (e.g., Foucault's pendulum or anomalies on maps). But since many others have adopted the view, others accept it. Even among physicists, few can examine carefully the evidence on all major theories. Inevitably, individuals must accept the overall decisions of others rather than their arguments and evidence.

It has also been suggested that various fertility decisions (how many children to have, whether or not to use contraception etc.) are heavily influenced by what other people in the same area are doing [Watkins (1990)]. Kislev and Bachrach (1973) also suggest that the same kind of reasons influence decisions to adopt new technologies. The same kind of influence is at work when, for example, academic researchers are on a "hot" topic and in "bandwagon effects" in consumer purchases. Another classic example is hiring in the job market [Stern (1990)][13].

Hence even though in many cases, 'conformist' behaviour might be a result of other non-Bayesian/non-social learning factors (like the ones discussed earlier), there exist several examples where the sequential process of decisions under uncertainty can be the compelling force leading to cascade formation. Moreover this has a special reference for financial markets and interaction of financial agents.

[13] This paper presents an econometric study based on a model in which a longer duration of job search is interpreted by employers as evidence that a worker has low skills.

III. Survey

Before proceeding on to the experiment, a brief discussion on uncertainty and individual rationality is warranted. Uncertainty, it is assumed, can be described by specifying random events in a given set S of 'states of the world'. A contingent history specifies a (possibly) different history for each possible state of the world. The particular contribution of Ramsey (1926) and Savage (1954) was to show that, under certain hypotheses, an individual decision maker – would ascribe a probability distribution to the states of the world in the set S, and 'von Neumann-Morgenstern' utilities to the possible histories, so that a history would be chosen in order to maximize the mathematical expectation of the 'utility' of the contingent history in each possible state of S. This is often called *subjective expected utility maximisation* because the probabilities are subjective in the sense that they need not conform to any of the standard frequentist or other notions of 'objective' probability. More simply, following Harsanyi, we may simply call it *Bayesian rationality*.

People can be called *rational* only if they have a general tendency to perform justified inferences. Are human beings primarily rational or primarily biased and irrational? People sometimes fail to live up to the highest standards laid down by the normative theories of how they ought to reason. They seem to be prone to a number of *biases*: they have tendencies to perform various types of unjustified inference.

More specifically, Bayesian rationality has often been criticized, but mostly on the grounds that individual's actual behaviour is not in accord with it – see, for example, Drèze (1974) and Kahneman and Tversky (1979). Tversky and Kahneman were responsible for introducing the idea of *heuristics and biases* into the explanation of ordinary non-demonstrative inference (See Kahneman, Slovic and Tversky [1982], Camerer [1995] review this literature and provide additional references). The core of this approach is that ordinary inference is sometimes mediated by general mental strategies – the heuristics – that may provide utility is some domains, but which in others create tendencies to error – the biases.

Several well-known experiments have been reported which illustrate these effects. This paper reports the findings on two such examples. To motivate my discussion, I distributed a survey asking two basic probability questions[14]. Subjects were recruited

[14] Four other questions regarding subject demographics were also asked.

via an email sent to multiple academic departments at Cambridge[15]. I received a total response from 57 participants[16]. Additionally, it is important to note that both staff[17] (academic and administrative) and students (Undergraduates and Postgraduates) received the email from their respective departments.

The data set was split almost equally between Men (53%) and Women (47%). In terms of nationality, the survey comprised of 53% British respondents and a further 9% from the U.S.A. As an academic discipline, History (25%), Economics (21%) and M.B.A (12%) made up a large share of the data. Furthermore, 60% of respondents were pursuing a Postgraduate degree of some sort.

Turning attention to the actual probability questions, in previous research, most people in the first question [see below Q1)] wrongly estimated that the witness is about 80 percent likely to be right, i.e. somewhere near the estimate of his accuracy, ignoring the *base rate* statistic. In the second question [see below Q2)], a number of subjects in the research of Kahneman and Tversky (1982) wrongly thought that the second (b) is more likely. These subjects committed the conjunction fallacy: a conjunction cannot be more probable than one of its conjuncts, since the conjunction logically implies its conjuncts. In this particular case, there are a lot of bank clerks, one some of whom are feminists. (The very most that is possible is that all bank clerks are feminists.)

In both the questions, researchers have found that people seem to base their judgments on how *representative* each case is of its category, and ignore or violate relevant statistical principles. In the first case, people ignore base-rate probability. In the second, there is a tendency to match the salient information about the person described (her radical politics) to a stereotypical category (feminists), leading to the conjunction fallacy.

Q1) Eighty-five percent of the cabs in the city are Green and the rest are Blue. A witness claims that a taxi in an accident was blue. Under tests, the witness correctly identifies both blue and green taxis on eighty-percent of occasions. What is the probability that the taxi was, in fact, blue?

[15] Please see Appendix for a list of departments contacted.
[16] There might be a selection bias as not every person who may have received the email bothered answering the survey. This might bias my result, however the direction of the bias can go either way.
[17] However, I did not receive a single staff response.

The correct answer in this question is about 41 percent. In other words, it is more likely (59 percent) that the witness was *wrong*. Assuming each kind of taxi is equally likely to be in accidents, we can say that, for every 100 crashes, 85 will tend to be of green taxis and 15 of blue ones. The witness will wrongly identify 20 percent of the former (i.e. 17) and correctly identify 80 percent of (i.e. 12) are blue. Thus, he would report blue 29 times, but only be correct on 12 of them, which is approximately 41 percent. (See Pollard and Evans [1983] for the more intuitive analysis).

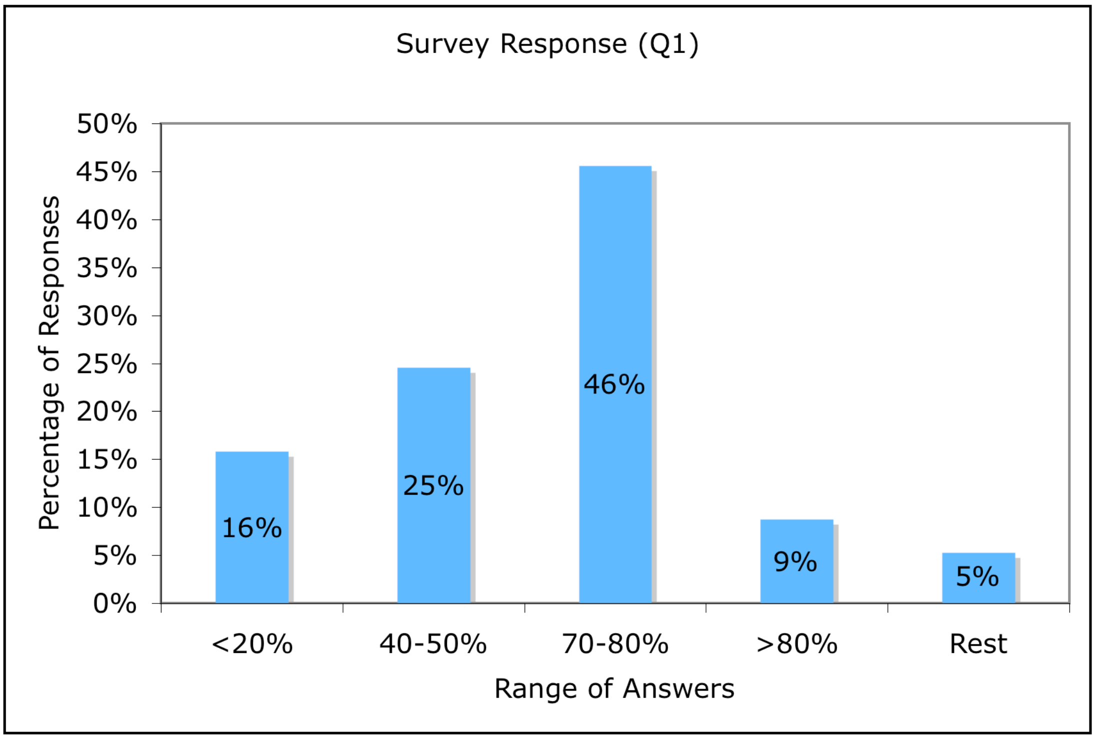

FIGURE 1. SURVEY RESPONSE (Q1)

Interestingly enough in my survey responses (see Figure 1 above), the Median and Mode answers were 80%. The mean[18] of all responses was 59%. Moreover only 12 respondents (21%) chose the correct answer.

Hence in total 31 out of 57 respondents (54%) chose an answer along the lines of previous research (70-80%, >80%). In this regard it is important to look back at the Bayes formulae for computing the correct answer.

P(Blue/ id Blue) = P(id Blue/ Blue)P(Blue)/ P(id Blue/Blue)P(Blue) + P(id Blue/Green)P(Green)

[18] One respondent inserted "Not knowable" as the response.

P(Blue/ id Blue) = (0.8)(0.15)/(0.8)(0.15)+(0.2)(0.85) = 0.41

Hence these results are not in contradiction what has been found in the past, showing that a substantial majority of people still got the answer wrong (with a majority of responses in the 80% range).

Q2) Linda is 31 years old single outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice and also participated in anti-nuclear demonstrations. Which is more likely:

a) Linda is a bank clerk

b) Linda is a bank clerk who is active in the feminist movement

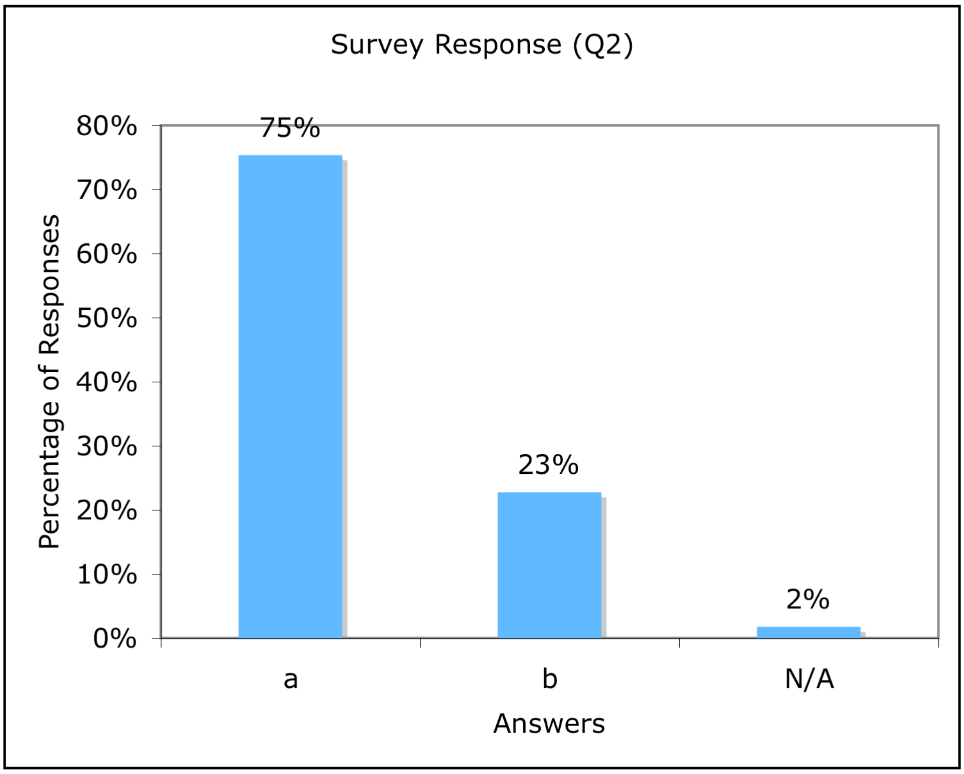

FIGURE 2. Q2) SURVEY RESPONSE

Surprisingly, the results from my survey of Q2) (see Figure 2 above) were quite different from what professional researchers have previously found. 43 out of 57 respondents rightly chose option a) as the correct answer (75 percent of all respondents) and option b) was chosen less frequently – by 13 people (23 percent of all respondents). One respondent was off the opinion that it could be "any or neither" - the entry has been recorded as N/A. The mode and median response were a).

However it is still important to note that 13 out of 57 individuals did suffer from the conjunction fallacy.

It is important to note that my results need to be viewed slightly carefully as there is a possibility of a selection bias[19] (and it would not be clear which direction the bias would go), my data set is not large enough and the survey was not conducted in a controlled laboratory environment. However some of my results [particularly for Q2)] do tend to support (to some extent) the existing criticism of the bias and heuristic literature.

The ideas of Tversky and Kahneman and their colleagues have not been accepted without question. Most critical attention has concentrated on what appears to be their rather pessimistic assessment of the ability of people to make accurate judgements under uncertainty, and recently, some attention has been directed not only at alternative explanations of the findings of heuristics and biases research, but also at the good statistical intuitions which people *do* exhibit.

For example, evidence. which confirms that people have, after all, some reasonable appreciation of statistical principles such as the law of large numbers comes from the work of Nisbett and his colleagues (see Fong, Krantz and Nisbett [1986]; Nisbett, Krantz, Jepson and Kunda [1983]; and a summary of this research in Holland, Holyoak, Nisbett and Thagard [1986]). These researchers have found that this apparently general understanding can be triggered – and, conversely, inhibited – by the way the problem is framed, and its relation to a person's existing knowledge.

Whether people can rightly be characterized as using heuristics has been questioned by several authors, e.g. Cohen (1981) and MacDonald (1986). Put briefly, Cohen's position is that since people must be rational in order to communicate, experiments which demonstrate otherwise do so either because they require knowledge outside the subject's experience, or because they have tricked the subjects into mistakes[20]. MacDonald (1986) is sympathetic to Cohen's point of view and tries to offer additional points in its support. One interesting argument he uses concerns the subjective weighting of the information presented in the kind of problems devised by researchers in heuristics and biases. The case of Linda (see above) will serve as an example. According to MacDonald, presenting the idea of Linda's being a bank clerk

---

[19] Not all who may have received the email replied.

[20] See the commentaries to his paper by Evans and Pollard (1981), Griggs (1981), and Kahneman (1981).

in the conjunction question is to invite one to *assume* it to be true – otherwise why mention it at all?

We thus have evidence that people are neither wholly at the mercy of irrational heuristics, nor naturally gifted naïve statisticians. Nevertheless, this section was meant as a spring board to initiative a discussion on some of the problems that may crop up when individuals are making decisions involving probability calculation.

IV. The Theoretical Model

The social learning model of Banerjee (1992) and Bikhchandani, Hirshleifer and Welch (1992) (BHW) describes the decision problem faced by a sequence of exogenously ordered individuals each acting under uncertainty about the state of the world. Each agent takes her decision within a Bayes-rational framework on both a privately observed informative signal, and the ordered history of all predecessors' decisions; she observes neither her predecessors' realized private signals, nor their realized payoffs. Private signals are assumed to be drawn from an identically and independently distributed random variable correlated with the state of the world.

Every agent in this structure settles with positive probability on a single action regardless and everyone rationally "herds" on the same (either good or bad) action. In this "informational cascade", socially valuable private signals are lost. Smith and Sørensen (2000) show that this result holds iff the quality of private signals is bounded.

Making a natural extension to the existing literature, Moscarini et al. (1998) consider a modified but natural model where it is common knowledge that the state of the world changes stochastically over time. They then ask how robust are the above findings to such an evolving environment. The authors find that because of the resulting information depreciation, only temporary informational cascades can arise. In fact BHW themselves discuss at length the fragility of cascades to the release of small amount of public information.

Here, they note that any cascade will eventually come to an end not due to informational input, but instead simply because of the fading relevance of old information. Moreover, cascades on a single action arise only if the state of the world is sufficiently persistent. When the underlying state changes sufficiently unpredictably no cascade ever arises, as prior information depreciates so fast that the belief can never be too extreme.

Finally, cascades on alternating actions arise when the state of the world is changing frequently enough, because here too information depreciates slowly. Hence to conclude, temporary cascades on single actions arise when the environment changes slowly. In this sense, the authors show that herding survives as a temporary phenomenon only.

This is the natural extension of the never-ending herd idea of BHW and Banerjee (1992) to a stochastic environment. Provided information does not depreciate too quickly, such inertia still arises. BHW themselves briefly discuss the possibility of the state changing in a non-stationary fashion. They provide a specific numeric example where the state changes with 5% chance after 100 periods; they find that cascade reversals are more likely than 5%.

It is important to note that the problem of optimal experimentation in a changing environment has been addressed by Kiefer (1991). Keller and Rady (1995) have developed a continuous-time version of the model. In the Moscarini et al. (1998) model the optimization problem and the aggregation is simple[21]. Information depreciation increases the instant value of new information but decreases its long-term value. In a social learning context, since agents are myopic, only the former effect is present, and the depreciation unambiguously discourages the non-informative cascade stage.

A finite set of individuals take sequentially a single action from a set of binary actions, a_0 and a_1. Payoffs to actions are contingent on an unknown state of the world, ω_0 or ω_1. Let q^1 be the common prior belief that the state is initially ω_1. Action a_1 is more rewarding than action a_0 in state ω_1, while the opposite is true in state ω_0: the payoff of action a_i in state ω_j is 1 if $i = j$ and 0 if $i \neq j$, with $i, j \in \{0,1\}$. For example, action a_i may be ``buy good *i*," and the state of the world $\omega_i$ : ``good *i* is better than the alternative good $j \neq i$."

After an individual's decision, the state of the world changes with chance ε, assumed for simplicity to be Markovian and independent of the current state of the world:

$$\Pr\left(\omega^n = \omega_i \mid \omega^{n-1} = \omega_j\right) = \Pr\left(\omega^n = \omega_j \mid \omega^{n-1} = \omega_i\right) = \varepsilon$$

[21] See Rustichini and Wolinsky (1995) for a long-run analysis in an individual experimentation problem.

for $i, j \in \{0,1\}$, $i \neq j$, and any n[22].

Before choosing an action, individual n both observes a private signal $\sigma^n \in \{\sigma_0, \sigma_1\}$ and the public history of action decisions of all preceding individuals $1,2,\ldots,n - 1$. The agent cannot see predecessors' signals. Private signals are drawn from a state-dependent Bernoulli distribution, and are independently conditional on the current state. The probability that the signal σ_i is realized in state ω_j is $\alpha > ½$ if $i = j$ and $1-\alpha < ½$ if $i \neq j$, with $i,j \in \{0,1\}$. The quality of the private signal is assumed bounded, i.e. $\alpha < 1$.

For $n \geq 2$ let $H^n \equiv \{a_0,a_1\}^{n-1}$ be the space of all possible period n histories of actions chosen by the n-1 predecessors of individual n. Let h^n denote an element of H^n. Let $q^n \equiv \Pr(\omega_1/h^n)$ be the public probability belief that the state is ω_1 in period n conditional on the publicly observed history of actions chosen by the predecessors of individual n.

Similarly let $r^n_i \equiv \Pr(\omega_1/h^n, \sigma_i)$ be the posterior belief that the state is ω_1 conditional on both the public action history h^n and the realization σ_i of the private signal observed by individual n. A simple application of Bayes' rule yields

$$r_i^n = \frac{\Pr(\omega_1 \cap \sigma_i | h^n)}{\Pr(\sigma_i | h^n)} = \frac{\Pr(\sigma_i | h^n, \omega_1)\Pr(\omega_1 | h^n)}{\Pr(\sigma_i | h^n)} ,$$

so that

$$r_0^n = r_0(q^n) = \frac{(1-\alpha)q^n}{\alpha(1-q^n) + (1-\alpha)q^n} \tag{2.1}$$

$$r_1^n = r_1(q^n) = \frac{\alpha q^n}{\alpha q^n + (1-\alpha)(1-q^n)} \tag{2.2}$$

[22] These results can be extended to a general two-state Markov transition matrix.

These posterior probabilities are used to compute the expected payoffs from taking the two different actions in the two states. The decision rule of agent *n* is to choose the action a^n which gives her the highest expected payoff. Given the simple payoffs, if individual *n* receives the private signal $\sigma^n = \sigma_1$, then it is optimal to take action $a^n = a_1$ if and only if $r_1^n \geq ½$. After substituting from (2.2), this becomes $q^n \geq 1-\alpha$. The decision rule can be summarized as:

$$\text{if } \sigma^n = 0, \text{ then } a^n = a_0 \Leftrightarrow q^n \leq \alpha, \text{ and } a^n = a_1 \Leftrightarrow q^n > \alpha$$
$$\text{if } \sigma^n = 1, \text{ then } a^n = a_0 \Leftrightarrow q^n < 1-\alpha, \text{ and } a^n = a_1 \Leftrightarrow q^n \geq 1-\alpha$$

where the action when indifferent WLOG minimizes the possibility of herding.

Were $\varepsilon = 0$, as in the models mentioned earlier, the public prior belief of individual n+1 equals the posterior belief that leads Ms. *n* to act according her signal $\sigma^n = \sigma_i$, i.e. $q^{n+1} = r^n_i$. There is an informational cascade (or cascade) on action a_i at time *n* whenever action a_i is taken by individual *n* regardless of the individual's private signal σ^n. Thus, a cascade on a_1 (respectively a_0) arises as soon as $q^k > \alpha$ (respectively $q^k < 1- \alpha$), for then the public belief swamps either private signal. A cascade once started would never end, because public belief would remain unchanged.

With $\varepsilon > 0$, the dynamics change drastically; however, the cascade region is unaffected by the state switching, since that event occurs only after the decision is made.

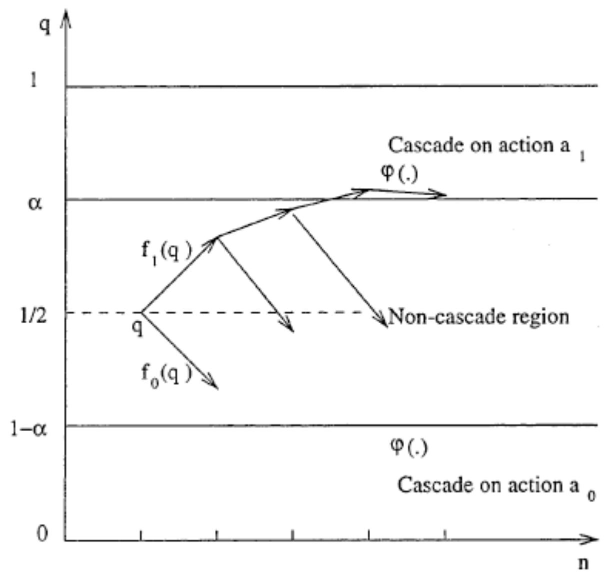

FIGURE 3. CASCADE AND NON-CASCADE REGION

When the possibility that the state of the world has changed in the meantime is accounted for, the public prior belief of individual $n+1$, coming after individual n who chose $a^n = a_i$ according to her signal $\sigma^n = \sigma_i$, satisfies $q^{n+1} = (1-\varepsilon)r^n_i + \varepsilon(1- r^n_i)$, which can be rewritten by (2.1) and (2.2) as

$$q^{n+1} = \begin{cases} f_0(q^n) \equiv \frac{(1-\varepsilon)(1-\alpha)q^n+\varepsilon\alpha(1-q^n)}{(1-\alpha)q^n+\alpha(1-q^n)} & \text{if } a^n = a_0 \\ f_1(q^n) \equiv \frac{(1-\varepsilon)\alpha q^n+\varepsilon(1-\alpha)(1-q^n)}{\alpha q^n+(1-\alpha)(1-q^n)} & \text{if } a^n = a_1 \end{cases} \qquad (3.1)$$

Now consider the case $q^k > \alpha$[23]. The action chosen will be $a^k = a_1$, regardless of the signal σ^k. The next individual $k + 1$ knows that $a^k = a_1$ is uninformative, and computes the public prior belief $q^{k+1} = (1-\varepsilon)q^k + \varepsilon(1- q^k)$. In general the following individual $n + 1$, as long as $q^n > \alpha$ or $q^n < 1- \alpha$, will update her prior belief during the cascade in the same manner, according to the uninformative cascade dynamics as shown below:

$$q^{n+1} = \varphi(q^n) \equiv (1-\varepsilon)q^n + \varepsilon(1-q^n) \qquad (3.2)$$

The public belief dynamics are stochastic and determined by (3.1) so long as $1-\alpha \le q^n \le \alpha$ (when not in a cascade) and are deterministic and follow (3.2) when either $q^n > \alpha$ or $q^n < 1-\alpha$ (during the cascade). Figure 3 above depicts these very dynamics.

The authors go on to argue that any cascade will eventually stop[24]:

Proposition 1. *For any $\varepsilon \in (0, 1)$, if a cascade exists, then for some $k = k(\varepsilon) < \infty$, the cascade must end in $k(\varepsilon)$ periods.*

Moreover, temporary cascades arise when information depreciates the least, i.e. when the next state is most predictable based on current beliefs. This happens not only when ε is small, but also when it is near 1[25]:

Proposition 2 (cascades on a single action). *For any $q^1 \in (1-\alpha, \alpha)$, with probability one, a cascade on some action arises in finite time if and only if $\varepsilon < \underline{\varepsilon}(\alpha) \equiv \alpha(1-\alpha)$.*

[23] Note that the other case can be treated symmetrically.
[24] Proofs of the following Propositions and Corollaries are not included in this paper and can be retrieved from the original Moscarini et al. (1998) paper.
[25] When the state of the world changes rapidly enough and individuals alternate between the two actions, the system may enter an alternating cascade in which individuals alternate between the two actions regardless of private signals.

Corollary 1 (cascades on alternating actions). *For $q^1 \in (1-\alpha, \alpha)$, with probability one, cascades on alternating actions arise in finite time if and only if* $\varepsilon > \bar{\varepsilon}(\alpha) \equiv 1 - \alpha(1-\alpha)$.

Finally, from Proposition 2 and Corollary 1, and α > ½ it follows that if the environment is changing in an unpredictable way there will never be a cascade:

Corollary 2. *No cascade ever arises for* $\varepsilon \in [1/4, 3/4]$.

Hence the above model makes clear predictions based on the relationship between ε and α. Using results from proposition 2 and corollaries 1 and 2, three versions of the same experiment were conducted in order to empirically test the same.

V. The Experiment

A. Physical Setup

In addition to the several alternatives to the Bayesian view of conformity discussed earlier, there are some which particularly may exist in an experimental setting. Psychologists and decision theorists have found a tendency for subjects to prefer an alternative that maintains the "status quo". Samuelson and Zeckhauser (1988) gave subjects hypothetical problems with several alternative decisions. When one of the alternatives was distinguished as being the status quo, it was generally chosen more often than when no alternative was distinguished. This systematic preference for the status quo is an irrational bias if the decision maker's private information is at least as good as the information available to the people who established the status quo. In answering a question about an unfamiliar decision problem, however, it can be rational for a subject to select the status-quo option if it is reasonable to believe that this status quo was initially established on the basis of good information or bad experiences with alternatives.

It is important to rule out all non-Bayesian reasons for herding – therein lies one of the strengths of a laboratory experiment involving non-hypothetical scenarios: it is possible to control information flows in the laboratory and therefore to determine whether subjects tend to follow previous decision (s) only when it is rational.

Moreover conformity due to interpersonal factors (as discussed earlier e.g. network externalities etc.) can also be minimized in a laboratory experiment with anonymity and careful isolation of subjects.

Anderson and Holt (1997) created a novel experimental set up which was based on a specific parametric model (binary-signal-binary-action) taken from Bikhchandani et. al (1992)[26]. This physical setup can be applied to the Moscarini et. al (1996) world by simply adding an additional layer of changing worlds. Hence it is important to note that the subsequent explanation of the physical set up is based on a simple case when the worlds are not changing.

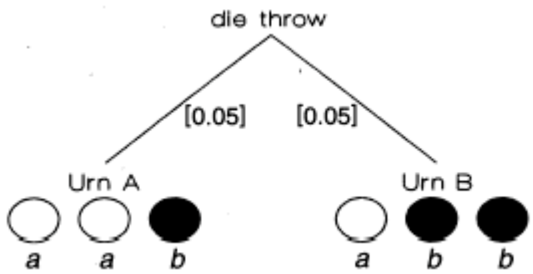

FIGURE 4. THE PHYSICAL SETUP

Consider the decision problem for an individual who observes a private signal that reveals information about which of two equally likely events has occurred. The events are denoted by A and B, and the signal is either *a* or *b*. The signal is informative in that the probability is 2/3 that the signal will match the label of the event. This setup can be implemented by putting balls labelled *a* or *b* in urns labelled A and B, as shown in Figure 4 above. Since the events (urns) are equally likely, each of the six balls in the figure are, *ex ante*, equally likely to be drawn. It is important to note that the posterior probability of event A given signal *a* is 2/3. Similarly, the posterior probability of event A given signal *b* is 1/3.

Suppose that individuals are approached in a random order to receive a signal and make a decision. The decisions (but not the signals) are announced publicly when they are made. If each individual earns a fixed cash payment for a correct decision (nothing otherwise), then an expected-utility maiximizer will always choose the urn with the higher posterior probability[27]. The first decision maker in the sequence, whose only information is the private draw, will predict event A if the signal is *a* and will predict event B if the signal is *b*. Hence, the prediction made by the first person will reveal that person's private draw.

[26] See Scharfstein and Stein (1990) for a similar application in an agency problem. This kind of principal agent problem (trying to con someone into believing that you know something) is common, especially in the context of asset markets.

[27] The payment method used in the experiments by me was slightly different. This is discussed in the next section.

If the second person's draw matches the label of the first person's prediction, then the second person should also follow the first person's prediction. But suppose that the first person predicts A and the second person draws *b*. The second person should infer that the first draw was *a*. This inference, combined with the *b* signal, results in posterior probabilities of ½ since the priors are ½ and the sample is balanced. I assume that the second person will choose the event that matches the label of the private draw when this label differs from the first decision[28]. In general, this rule is assumed to be the tiebreaker convention throughout this paper for all subjects.

Suppose that each subsequent individual assumes that others use Bayes' rule to make predictions[29]. For example, if the first two decisions are A and the third person observes a *b* signal, then this person is responding to an inferred sample of *a* on the first two draws, and *b* on the third draw. Since the events are equally likely a priori, and since the sample favours event A, the posterior probability of A is greater than ½. In this case, the third person should predict event A in spite of the private *b* signal[30].

Hence in such a set up, just the first two decisions can start a cascade in which the third and subsequent decision-makers ignore their own private information. Whenever the first and the second individuals make the same prediction, all subsequent decision-makers should follow, regardless of their own private information. In all cases, it takes an imbalance of two decisions in one direction to overpower the informational content of subsequent individual signals.

If individuals recognize that decisions made after the beginning of a cascade are not informative, they will ignore these "irrelevant" decisions in their probability assessments. But if someone breaks out of a cascade pattern and predicts the other event, then it is reasonable to assume that this deviant decision reveals a private signal that is contrary to the cascade, because the expected cost of deviating would be higher if the signal matched those inferred from previous decisions. Therefore, relevant signals are those inferred from decisions made before a cascade starts, from the two decisions that start a cascade, and from non-Bayesian deviations from a cascade.

[28] We could make an alternative assumption that decision is random, i.e., that it matches the label of the private signal with probability ½. This would not alter the analysis of cascade formation that follows, but it would alter some of the numerical probability calculations.
[29] This assumption is further discussed n Section VI.
[30] Here I have interpreted the two initial A decisions as indicating two *a* draws, i.e., that the second person would have announced B with a private *b* signal.

Specifically, if n is the number of relevant a signals and m is the number of relevant b signals. Then Bayes' rule can be used to calculate the posterior probability of event A, given any sequence of sample draws -

$$\Pr(\mathrm{A}|n, m)$$

$$= \frac{\Pr(n, m|\mathrm{A})\Pr(\mathrm{A})}{\Pr(n, m|\mathrm{A})\Pr(\mathrm{A}) + \Pr(n, m|\mathrm{B})\Pr(\mathrm{B})}$$

$$= \frac{(2/3)^n(1/3)^m(1/2)}{(2/3)^n(1/3)^m(1/2) + (1/3)^n(2/3)^m(1/2)}$$

$$= \frac{2^n}{2^n + 2^m}.$$

Finally it is important to draw parallels between this physical set up and Moscarini et. al (1996). The binary actions a_0 and a_1 can be compared to deciding whether the true state is Urn A or Urn B respectively. The unknown state of the world, ω_0 or ω_1 can be compared to event (urn) A or B respectively. q_1 is 50% in this set up due to the initial roll of the die. ε is the chance that the true state switches between A and B after an individual's decision and is assumed for simplicity to be Markovian and independent of the current state of the world. The value's of ε are allowed to vary across treatments and is discussed in part C of this section.

Individuals also observe private signals 'a' or 'b' ($\sigma^n \in \{\sigma_0, \sigma_1\}$) and the public history of action decisions (A or B) of all preceding individuals. α is interpreted as the probability that signal 'a' comes from Urn A or symmetrically the probability that signal 'b' comes from Urn B. The quality of the private signal is assumed bounded and is 2/3. When $\varepsilon = 0$ the Bayes' rule mentioned in Moscarini et. al (1996) would coincide with the Bayes' rule derived using n and m relevant signals. For non-zero values of $\varepsilon = 0$ Bayes' rule mentioned in section IV. should be used.

B. Procedures

A total of 23 subjects were recruited from the University of Cambridge and at least to my knowledge had no previous experience with this experiment. There was no restriction on participants and recruitment emails were sent out across the student and staff community. Each subject was paid out a John Lewis voucher worth £20 and each session lasted about 90 minutes. A total of 5 sessions were conducted with 5 subjects for each of the first three sessions and 4 subjects in each of the last two sessions. In each of the 5 sessions, subjects were asked to participate in all the three treatments. After the experiment subjects were also asked to fill out an online[31] post-experiment questionnaire. It is important to note that a compensation scheme based on the actual performance in the experiments (similar to Anderson and Holt [1997]) would have been a better incentive mechanism and might have made the subjects work harder. Unfortunately such a scheme was not possible due to logistical reasons beyond my control. However it is important that the compensation amount in this experiment (£20) exceeded the average earnings (including participation fee) in Anderson and Holt (1997) ($25) and in Celen and Kariv (2004) ($27). In fact this was one of the reasons for choosing $20 as the amount and therefore I have assumed this to be a "reasonable compensation" to ensure subject participation and attention. Additionally it is important to point out that I had initially secured funding for a total of 50 subjects, but due to constraints of time and logistics, only 23 participants could be recruited.

The experiment was set up online using an interactive PHP based programme designed by Nishanth Sastry and Eiko Yoneki based out of the Computer Laboratory at the University of Cambridge. Within any one session, all subjects participated in the first treatment and moved on to subsequent treatments only after finishing the first treatment. Written instructions were distributed for each treatment and each subject needed to answer some basic questions to ensure that they have clearly understood the instructions. To ensure complete clarity a visual demonstration was followed after every subject had finished reading the instructions and questions were taken subsequently. To ensure anonymity, only a single participant was in the computer

[31] The purpose of this questionnaire was to ensure (for post analysis) that the subjects clearly understood the aims and objectives of the experiment. Certain questions were also designed to suss out what mechanisms (if any) subjects relied on to answer the questions and how they perceived others rationality. Results to be discussed in subsequent sections.

laboratory at any one point in time. Every other participant was seated in an adjoining waiting room.

Each session consisted of 10 independent rounds per treatment. At the start of each round, the web-based program used an ‘electronic’ throw of the dice to determine which of the two urns would be used. Urn A was used if the throw of the die was one, two, or three; urn B was used otherwise. Subsequently a Markov Process decided the likelihood of the Urn being switched between subjects within any one round. This likelihood of state switching (which was Markovian and independent of the current state of the world) varied across the three treatments. The instructions (which were specific to each treatment) included and explained the entire structure and working of the treatment in question.

Within any one treatment the order of the subjects were chosen at random and remain fixed for all of the 10 rounds of that particular treatment. For e.g. - if subject S4 was the first decision maker in treatment 1, she would be the first decision maker in all the 10 rounds of treatment 1. Hence the first decision maker took all her 10 decisions (one per round) at the same time. Subsequently the second decision maker (second in all the 10 rounds) would take all her 10 decisions at the same time after having observed her own private signal and the first person’s decision in each of the 10 independent rounds. This process continued for all the 5 subjects in the session in any one treatment. Hence after the first treatment was completed (all 10 rounds for all 5 subjects) we moved on to the second treatment wherein a new order for the subjects was decided at random. The session would be complete after the third treatment.

It is important to reemphasise that the ordering however was changed at random across treatments (while remaining fixed within any one treatment). Within in any one treatment, in each round, subject’s private signals were drawn (again ‘electronically’) with replacement. After seeing a private draw for each of the 10 rounds, the subject would record his/her urn decision, A or B, and the decision would appear on the monitor on the subsequent subjects turn. In this way, each subject knew his or her own private draw and the prior decisions of others, if any, before making a prediction. This process continued until all subjects had made decisions. At the end of each subjects turn, the correct Urn being used was displayed on the screen.

It is important to note that in order to strictly follow the norm in Anderson and Holt (1997) and other similar studies, one should have also randomly changed the order of decision making between rounds within each treatment. However, the software designed for this study precluded doing the same. However, it is safe to assume that this particular caveat is not central to the analysis. Nonetheless the order for each treatment was decided at random and was also changed at random between treatments.

C. Results

An information cascade is possible if an imbalance of previous inferred signal causes a person's optimal decision to be inconsistent with his or her private signal. Moreover in the words of Moscarini et. al (1996) - "there is an informational cascade (or cascade) on action a_i at time *n* whenever a_i is taken by individual *n* regardless of the individual's private signal σ^n".

C.a Treatment 1 (ε = 0)

In treatment 1 the state of the world does not change between subjects within any one round. With $\alpha = 2/3$ this treatment corresponds to Proposition 2 in section IV. Hence the theory would predict a cascade on a single action in finite time. With 5 sessions and 10 per rounds per session, we had a total of 50 rounds to analyse. Moreover, with 23 subjects there were a total of 230 observations to study. Cascade behaviour was observed 23 out of the 26 rounds in which it was possible for cascades to be formed. Within this there were 6 instances of reverse cascades observed. Additionally there was a minority of cases when individuals did not follow the pattern of rational inference about other's signals. There are 3 such rounds where the individual could have cascaded (and ignored her private decision) but instead chose to make a decision based on private information but inconsistent with Bayesian updating. Over all the five sessions, 3 percent of the decisions were inconsistent with both Bayes' rule and private information. Lastly there were 7 instances where the tiebreaker convention was not followed i.e. - when Bayes' rule showed ½ as the chance of either event occurring, the individual does not follow her signal (as assumed) but simply forms a cascade.

TABLE 1 - DATA FOR SELECTED ROUNDS ACROSS SESSIONS 1,2 AND 3

Sessions 1,2, 3		Decision Sequence					
		Subject number: Urn decision (private draw)					Cascade outcome
Round	Urn Used	1	2	3	4	5	
2	B	S8:B(b)	S7:B(b)	**S6:B(a)**	S9:B(b)	**S10:B(a)**	cascade
7	B	S8:B(b)	S7:B(b)	**S6:B(a)**	S9:B(b)	**S10:B(a)**	cascade
8	A	S8:B**(a)	S7:A(a)	S6:B(b)	**S9:B(a)**	**S10:B(a)**	reverse cascade
8	B	S15:B(b)	S11:B(b)	S13:B(b)	**S14:B(a)**	**S12:B(a)**	cascade

Notes: Boldface - Bayesian decision, inconsistent with private information.

* - Decision based on private information, inconsistent with Bayesian updating.

** - Decision inconsistent with Bayes' Law and private information.

Table 1 shows the longest sequence of cascade behaviour selected from sessions 1, 2 and 3. Consider the third row from top (round 8). Although Urn A was used and the first decision maker (subject S8) saw a 'a' signal, subject S8 incorrectly decided to predict B. The second person saw a 'a' signal again and predicted A. Therefore the first two outcomes effectively cancel each other out. The third person rightly (in the Bayesian sense) predicted B. Now the fourth person experienced a tiebreaker and decided to follow the herd and ignore the private information. The last person in the sequence rationally cascaded. This is an instance of reverse cascade where the group settled on an incorrect outcome. The boldfaced characters indicate decisions that were consistent with Bayes' rule and inconsistent with private information. Similar patterns emerged in rounds 2,7 and 8 (this was round 8 from session 3).

TABLE 2 - DATA FOR SELECTED ROUNDS ACROSS SESSIONS 4 AND 5

Sessions 4, 5		Decision Sequence				
		Subject number: Urn decision (private draw)				Cascade outcome
Round	Urn Used	1	2	3	4	
3	A	S16:B(b)	S18:B(b)	**S20:B(a)**	**S19:B(a)**	reverse cascade
5	A	S16:B(b)	**S18:B*(a)**	**S20:B(a)**	**S19:B(a)**	reverse cascade
9	A	S16:B(b)	S18:B(b)	S20:B(b)	**S19:B(a)**	reverse cascade
2	B	S21:B(b)	S22:A(a)	S23:A**(b)	**S24:A*(b)**	cascade
10	B	S21:B(b)	S22:B(b)	S23:B(b)	**S24:B(a)**	cascade

Notes: Boldface - Bayesian decision, inconsistent with private information.

* - Decision based on private information, inconsistent with Bayesian updating.

** - Decision inconsistent with Bayes' Law and private information.

Similarly Table 2 above shows selected cascade data from sessions 4 and 5. It is important to note that this treatment ($\varepsilon = 0$) is exactly the same experiment that Anderson and Holt (1997) conducted. Although that experiment was conducted with a much larger data set, the results are qualitatively the same. Similar to treatment 1, they found a high rate of conversion to cascade outcomes (cascade behaviour was observed in 41 of the 56 rounds when such behaviour was possible), a relatively small frequency of individuals only following their private information (26 per cent). Lastly consistent with this paper they too found a very few cases where the decisions were inconsistent with both Bayes' and private information (4 per cent).

Lastly, it is encouraging to see that the outcomes in treatment 1 are not clashing with what theory would predict. Occurrence of rational cascades and the limited occurrence of inconsistent observations are not in contradiction with Proposition 2 in section IV (which predicted the occurrence of cascades in finite periods).

C.b Treatment 2 (ε = 0.5)

In treatment 2 the state of the world changes with 50 per cent chance between subjects within any one round. With $\alpha = 2/3$ this treatment corresponds to Corollary 2 in section IV. Hence the theory would predict that no cascade should ever arise. With 5 sessions and 10 per rounds per session, we had a total of 50 rounds to analyse. Moreover, with 23 subjects there were a total of 230 observations to study.

In general the broad results from treatment 2 are not in contradiction with Corollary 2. In majority of the cases, subjects ignore outcomes from their predecessors and simply follow their private information. However, there are a few cases that appear to be inconsistent. In a total of 10 observations (4 per cent of all observations) subjects have taken decisions, which would be inconsistent with their private information.

TABLE 3 - DATA FOR SELECTED ROUNDS ACROSS SESSIONS 2 AND 3

Sessions 2,3	Decision Sequence Subject number: Urn decision (private draw) (True Urn)					Cascade outcome
Round	1	2	3	4	5	
10	S6:A(a)(A)	S7:A(a)(B)	**S8:A(b)(A)**	S9:B(b)(A)	S10:B(b)(A)	Cascade possibility
3	S11:B(b)(B)	S15:B(b)(B)	**S14:B(a)(A)**	S13:B(b)(A)	S12:A(a)(A)	Cascade possibility
4	S11:A(a)(A)	S15:A(a)(A)	**S14:A(b)(B)**	S13:A(a)(B)	S12:B(b)(B)	Cascade possibility
6	S11:A(a)(A)	S15:A(a)(B)	S14:A(a)(B)	S13:A(a)(A)	**S12:A(b)(B)**	Cascade possibility
7	S11:A(a)(B)	S15:B(b)(B)	S14:B(b)(A)	S13:A(a)(A)	**S12:A(b)(B)**	Cascade possibility
9	S11:B(b)(A)	**S15:B(a)(A)**	S14:B(b)(B)	S13:B(b)(A)	S12:A(a)(A)	Cascade possibility

Notes: Boldface - Decision inconsistent with Bayes' Law/ private information.

Table 3 shows selected data from sessions 2 and 3 where individuals have taken decisions that are inconsistent with Bayes' Law (and private information). Consider the fourth row from above where subject S12 ignores the private signal 'b' and predicts (incorrectly) event A. It is interesting to note that this subject appears to have fallen into a trap where she/he believes that the sequence of A decisions before her/him are part of a cascade pattern. However the theory would tell us that any such thinking would be deemed incorrect as with (ε = 0.5) one should simply follow ones private signal and not follow any "herd". The other rounds in this table also appear to depict a similar pattern.

TABLE 4 - DATA FOR SELECTED ROUNDS ACROSS SESSIONS 4 AND 5

Sessions 4,5	Decision Sequence				
	Subject number: Urn decision (private draw) (True Urn)				Cascade outcome
Round	1	2	3	4	
8	S16:B(b)(A)	S18:B(b)(B)	S19:B(b)(B)	**S20:B(a)(A)**	Cascade possibility
10	S16:B(b)(B)	S18:B(b)(B)	S19:B(b)(B)	**S20:B(a)(B)**	Cascade possibility
5	S23:B(b)(B)	S21:B(b)(B)	S22:B(b)(B)	**S24:B(a)(B)**	Cascade possibility

Notes: Boldface - Decision inconsistent with Bayes' Law/ private information.

Looking at table 4 for sessions 4 and 5 a very similar pattern appears; inconsistent outcomes are based around situations when a 'cascade appears'. It is important to note that in treatment 2, the term cascade is used loosely as strictly speaking cascade refers to situations when individuals rationally ignore their private information and follow the decisions of those before them. However the broad results from treatment 2 are heartening as an overwhelming majority of the observations follow the theory. There are only a very few cases where inconsistent outcomes are observed.

C.c Treatment 2 (ε = 0.9)

Finally in treatment 3 the state of the world changes with 90 per cent chance between subjects within any one round. With $\alpha = 2/3$ this treatment corresponds to Corollary 1 in section IV. Hence the theory would predict that cascades on alternating actions could arise in finite time. With 5 sessions and 10 per rounds per session, we had a total of 50 rounds to analyse. Moreover, with 23 subjects there were a total of 230 observations to study.

Cascades on alternating actions were observed in 15 out of the 18 rounds when such behaviour was possible. A total of 7 were outcomes of reverse cascades. Additionally there were 3 rounds in which participants followed their private signals only and ignored Bayes' Rule. The tiebreaker convention was broken 15 times and a total of 12% of observations were inconsistent with both private information and Bayesian updating.

TABLE 5 - DATA FOR SELECTED ROUNDS ACROSS SESSIONS 1,2 AND 3

Sessions 1,2,3		Decision Sequence				
	Subject number: Urn decision (private draw) (True Urn)					Cascade outcome
Round	1	2	3	4	5	
2	S1:A(a)(A)	S2:B(b)(B)	S3:A(a)(A)	**S4:B(a)(B)**	S5:A(a)(A)	cascade
4	S8:A(a)(A)	S6:B(b)(B)	**S7:A(b)(B)**	**S10:B(a)(A)**	**S9:A(b)(B)**	reverse cascade
9	S8:A(a)(B)	S6:A(a)(A)	**S7:A(b)(B)**	S10:B(b)(B)	**S9:A(b)(A)**	cascade
7	S15:B(b)(A)	S11:A(a)(B)	**S14:B(a)(A)**	S13:A(a)(B)	S12:A(a)(A)	reverse cascade
9	S15:B**(a)(A)	S11:A(a)(B)	S14:B(b)(A)	S13:B(b)(B)	**S12:B(a)(A)**	reverse cascade

Notes: Boldface - Bayesian decision, inconsistent with private information.

* - Decision based on private information, inconsistent with Bayesian updating.

** - Decision inconsistent with Bayes' Law and private information.

Table 5 shows the longest sequence of cascade behaviour selected from sessions 1, 2 and 3. Looking at the second row (round 4) from above we can see alternating reverse cascade for subjects S8, S7 and S9 and alternating reverse cascades between S6 and S10. Similarly looking at round 9, we see S8, S7 and S9 cascading together - S7 has a reverse cascade whilst S9 has a normal cascade.

TABLE 6 - DATA FOR SELECTED ROUNDS ACROSS SESSION 4

Session 4	Decision Sequence				
	Subject number: Urn decision (private draw) (True Urn)				Cascade outcome
Round	1	2	3	4	
2	S16:B(b)(A)	S18:B(b)(B)	S19:A(a)(A)	**S20:B(a)(B)**	cascade
3	S16:A(a)(A)	S18:B(b)(B)	**S19:A(b)(A)**	S20:B(b)(B)	cascade
6	S16:B(b)(B)	S18:A(a)(A)	S19:A**(b)(B)	**S20:A(b)(A)**	cascade

Notes: Boldface - Bayesian decision, inconsistent with private information.

* - Decision based on private information, inconsistent with Bayesian updating.

** - Decision inconsistent with Bayes' Law and private information.

Similarly Table 6 shows some selected from round 4. In this treatment it is important to note that sessions with only 4 participants (sessions 4 and 5) are not ideal for analysis as alternating cascades can only happen in pairs and hence it is not possible to analyse longer sequences of potential cascade behaviour.

Comparing the results with the theory, Corollary 1 in section IV predicted alternating cascades in finite time. Treatment 3 witnessed 15 such instances out of a total potential of 18. Hence the "conversion rate" seems to fit the theory and is similar to what was observed in treatment 1. However there does seem to be a higher frequency of behaviour that is inconsistent with both private information and Bayes'. In Treatment 3 this stood at 12 percent (compared with only 3 percent in treatment 1). It is important to note that this implies that 12 percent of the times the theory's prediction was violated (assuming common knowledge of rationality). Instances where the outcome was inconsistent with only Bayes' (the * outcome) would also be technically counted within cases where the theory is violated. Counting the same this number jumps to about 13 percent in treatment 3 (compared with 4 percent in treatment 1).

Since treatment 3 is computationally more challenging than treatment 1, it is not entirely surprising to see a higher frequency of inconsistent outcomes. However we can still find refuge in the fact that the broad prediction of Corollary 1 was not violated; alternating cascades were observed almost always when they were supposed to.

VI. Further Post Experiment Analysis

A. Survey

In experimental economics, it is of paramount importance to insure that subjects have a full and correct understanding of the mechanisms involved in the experiment. Moreover there is an unwritten protocol that prohibits researchers from lying to subjects. Lastly, Bayesian outcomes are analysed assuming common knowledge of rationality, which, may not always be the case. After the three treatments were completed in any one session the participants were asked to fill out a questionnaire with the following -

Q5) Were the instructions clear?

Q6) What do you think was the purpose of this experiment?

Q7) How did you make your decisions?

Q8) What rule, if any did you use to make your decision?

Q9) How do you reckon the other participants in this room made their decision?

It is important to note that Q7), Q8) and Q9) were asked specific to each treatment. In addition certain demographic questions were also asked (Gender, Nationality, Subject of study and degree level). These have not been included in this analysis, as they are not considered central to the analysis.

Q5) and Q6) were incorporated to ascertain whether there was full understanding of the experiment. Q7) and Q8) were used to analyse what criteria subject's use to make decisions. Q9) was asked to better understanding the belief structure that was prevalent in the study. Questions 7,8 and 9 are important to get a deeper understanding behind the outcomes, which prima facie are considered inconsistent.

Lastly, it is important to mention that there are drawbacks in this questionnaire. Subjects have answered questions qualitatively (apart from Q5) which were then interpreted by me. In hindsight a more rigorous approach could have included similar questions with multiple-choice answers.

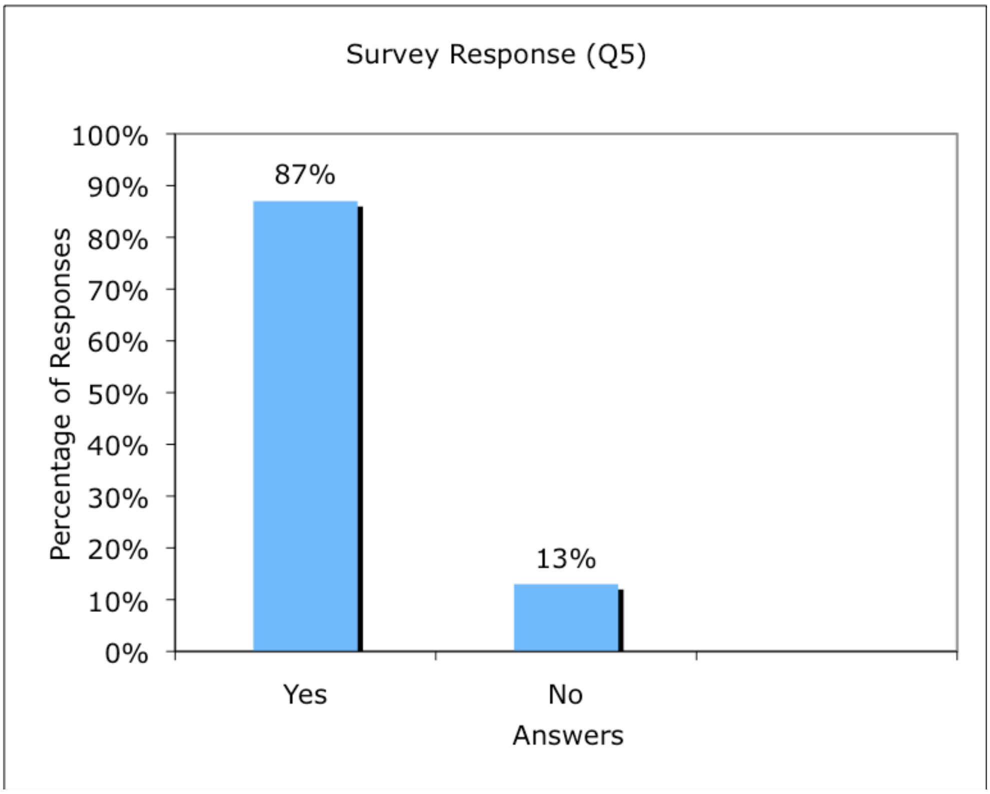

FIGURE 5. SURVEY RESPONSE (Q5)

Survey response from Question 5 reveals that 87% of the participants understood the instructions and were not confused by it. This fact is corroborated by the responses from Q6) where all subjects correctly described the purpose of the experiment.

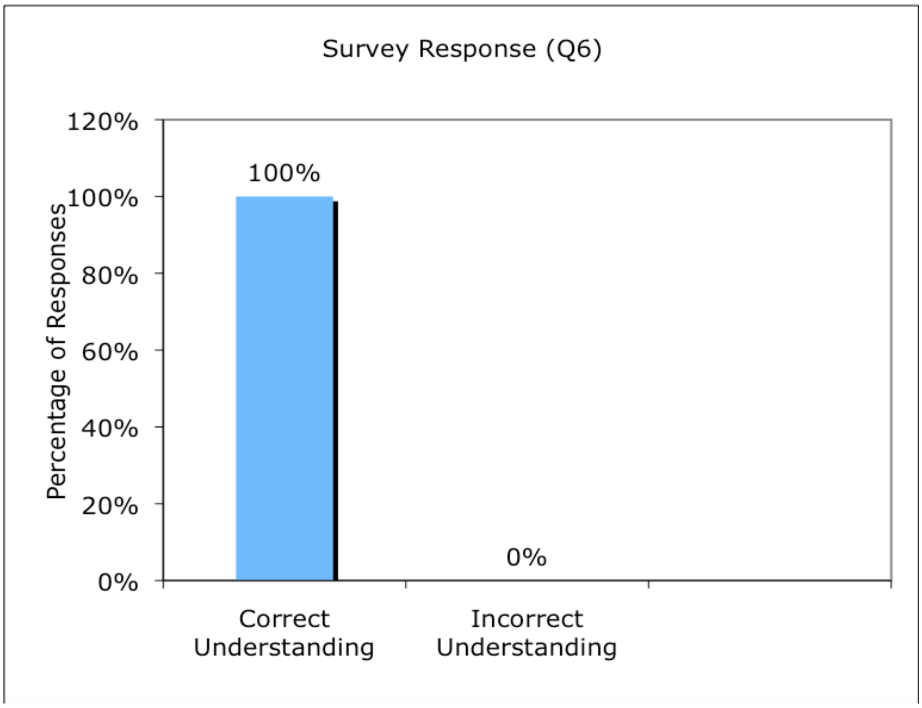

FIGURE 6. SURVEY RESPONSE (Q6)

For question 7, 87 percent and 91 percent of the respondents correctly described the decision-making criteria for treatment 1 and 2 respectively. Question 8 reveals the exact same numbers (87 and 91%) for the two treatments. This is not surprising given the low incidents of irrational or inconsistent outcomes in both the treatments. Moreover answers from Q9) reveal that in treatment 1 96 percent of all respondents expect others to use the same logic as theirs. This number shoots up to 100 percent for treatment 2.

In treatment 3 however these numbers appear to be less healthy - in Q7) 78 percent of the respondents correctly describe the decision-making criteria. The corresponding number for Q8) in treatment 3 is 74 percent. Hence about over 20 percent of the subject population (5-6 respondents) wrongly decided on how to tackle treatment 3. Furthermore subjects expect less rationality of others as well in treatment 3 - In question 9 13 percent of all respondents do not expect others to behave in a logical manner. These findings again seem to only reemphasize the data from treatment 3. Treatment 3 requires a more "involved" calibration and therefore participants struggled more in assessing the correcting criteria.

In conclusion the major findings from treatment 1 and 2 are more in accord with the theory in Moscarini et. al (1996) - the outcome from the survey also showed that the participants were broadly aware of the correct decision making criteria and expected others to do the same. Finally, most subjects were not confounded by the instructions.

B. An Econometric Methodology

Given the limited amount of data (23 subjects) it would not be feasible to conduct a meaningful econometric study. A larger data set (which was originally planned) would be needed. However, the methodology deserves a brief mention [See Choi et. al (2009), Weizsäcker (2008), Celen and Kariv (2004) and Anderson and Holt (1997)].

A dynamic model (similar to a random utility model) in which people calculate posteriors allowing for the possibility of errors in earlier decisions would allow for a more rigorous analysis of question 9 from the survey in the previous subsection. Error rates are econometrically estimated assuming a logistic distribution of independent shocks to expected payoffs. The equilibrium can be summarized by a choice probability function following a binomial logit distribution:

$$\mathbf{Pr}\left(a_{it}=1|I_{it}\right)=\frac{1}{1+\exp\left(-\beta_{it}x_{it}\right)},$$

where a_{it} is the action of agent i at time t (Urn A or B), I_{it} is agent *i's* information set at t (history of observed decisions from her predecessors), β_{it} is a coefficient, and x_{it} is the difference between the expected payoffs from actions $a = 1$ and $a = 0$, respectively (Urn A or B). Note that β captures the sensitivity to payoff differences. The tendency to make errors diminishes as $\beta \to \infty$, while behavior becomes random as $\beta \to 0$. For positive values of β the choice probability is an increasing function of the payoff difference.

This model does assume that subject's should rationally take into account the mistakes of others when drawing inferences from their behavior. Hence the error structure in this model is recursive: the β parameter for the first person in the sequence affects the second person's expected payoffs, which are used in turn to estimate a β parameter for the second-stage decision. At each stage, the β estimates for previous nodes are used to calculate the expected payoffs for each decision, conditional on the private signal and the decisions observed in the previous nodes.

Then the difference in expected payoffs for a period (node) constitutes the independent variable in the estimation for that particular period (node).

Finally the maximum likelihood method is used to estimate the parameter β. Anderson and Holt (1997) recommend the Newton-Raphson algorithm to minimize the negative of the log-likelihood function (as an alternative to the recursion they also mention that the β can be restricted to be the same for all periods. However the recursive method would provide a better fit as suggested by the likelihood ratio test). Lastly, Choi et. al (2009) also recommend some specification tests to check the restrictions that can be imposed on such a model.

It is important to note that this methodology is an adoption of the Quantal Response Equilibrium (QRE) model of McKelvey and Palfrey (1995, 1998) which allows for agents to make mistakes (occasional) and can be interpreted following Harsanyi and Seltens's "trembling hand".

VII. Further Extensions

The model presented by Moscarini et. al (1996) presents a stylized reference point which has several real world applications (in finance and otherwise). Real world data in finance is "noisy" and therefore it is difficult to test for very specific questions. Using simple experimental methodologies it is possible to test for a model (such as the changing world one), which can be broadly representative of the actual mechanisms at work in the market place. However this paper is only an initial attempt to test for herding in an uncertain world - several extensions and improvements are in order.

A. Heuristics

The physical set up of this experiment was symmetric (2 'a' balls and 1 'b' ball in Urn A and vice versa in Urn B). One implication is that subjects may suffer from a counting heuristic - it is possible to simply count the inferred signals to get accurate results. Hence observing consistent behaviour does not necessarily imply Bayesian updating on behalf of the participants. In essence the set up allows for a simple rule of thumb using which you are highly likely to get the right answer (more so in treatment 1 and 2). Anderson and Holt (1997) recommend an asymmetric setup to distinguish counting from Bayesian behaviour. Using an asymmetric set up one could test the model of changing worlds. It is important to note that with such a structure Bayes' law becomes more complicated and one could expect a greater deviation in behaviour from theory.

However, the presence of a counting heuristic does not necessarily reduce the credibility of the symmetric design. Anderson and Holt (1997) did conduct a series of experiments using an asymmetric distribution. Although they found a lower incident of rational cascades (normal) (70 percent vs. 73 percent), the overall conclusions do not change - a reasonably high occurrence of rational cascades was still observed. Moreover only a third of the deviations from Bayes' rule in the asymmetric design could be explained by counting.

Moreover the presence of certain heuristics need not be such a bad thing. Certain simple rules of thumb can provide a "sufficient statistic" for more complicated calculations. Off particular relevance in this area is the work of the German psychologist Gerd Gigerenzer. A critic of the work of Kahneman and Tversky, he focuses on how heuristics can be used to make optimal decisions rather than produce cognitive biases. Gigerenzer et. al (1999) point out that "Fast and frugal heuristics that embody simple psychological mechanisms can yield inferences about a real-world environment that are at least as accurate as standard linear statistical strategies embodying classical properties of rational judgment".

B. Herding vs. Cascades

It is important to point out that although the terms informational cascade and herd behaviour are used interchangeably in this paper, Smith and Sørensen (2000) emphasize that there is a significant difference between them. An informational cascade is said to occur when an infinite sequence of individuals ignore their private information when making a decision, whereas herd behavior occurs when an infinite sequence of individuals make an identical decision, not necessarily ignoring their private information. Thus, an informational cascade implies a herd but a herd is not necessarily the result of an informational cascade.

The practical importance of the distinction between the two is that in a cascade learning ceases, while in a herd the individuals become more and more likely to imitate but their actions may still provide some information. Hence in a herd the behavior can change suddenly and dramatically and may better explain why mass behavior is so fragile and prone to fads. Understanding such a distinction may be of importance to better streamline real world data (especially from asset prices).

In the Anderson and Holt (1997) discrete-signal-discrete-action setup all herds are cascades since once two consecutive decisions coincide no signal can lead to a deviation. In contrast, Celen and Kariv (2004) propose a continuous-signal-discrete action setup to distinguish between cascades and herds completely. Using this setup on the changing world model would enrich the literature and enable a deeper study of social learning in an uncertain world.

C. General Extensions

In addition to the above there are a host of other extensions that could provide interesting and important answers towards the study of learning in the financial world.

Within any one treatment, if subjects were made to play one round at a time and the ordering was changed between rounds, there could be a strong learning component that could be analysed. It is not unreasonable to expect questions of individual and group learning over time to have important theoretical and practical implications.

It may also be off interest and relevance to further analyse the data with respect to certain demographic fixed effects such as gender, nationality or degree of study. In this regard Fisman et. al (2009) talk about exposures to different "ideologies" and its impact on social learning. In another interesting application Celen et. al (2009) talk about studying social learning in the presence of information, signals and advice from other agents.

In another paper Kariv (2005) studies the impact of social learning in the presence of "overconfident agents". He describes them as those who "overweigh their private information relative to the public information revealed by the decisions of others." Therefore, when following a herd, they "broadcast" more of the information available to them. The paper goes on to show that such a situation increases the free-rider problem of rational agents. Extending this to the uncertain world in this paper would have obvious applications for trading in financial markets.

The functioning of financial markets have often been viewed as networks [See Allen and Babus (2008), Babus (2008), and Gale and Kariv (2007)]. Moreover, experiments have studied social learning within distinct network architectures [See Gale and Kariv (2003), and Choi et. al (2009)]. Additionally, Sørensen and Smith (2008) also study a model of social learning where “everyone only sees un-ordered samples from the action history”. Conducting experiments within specific networks with the additional layer of changing worlds would again provide a rich study for a comparison with the real world.

VIII. Criticisms

Experiments in general have several things in their favour. Amongst other things they “weed out” bad theories, predict behaviour in the field, establish limits to the models (“stress tests”), and discover new empirical regularities.

However, the methodologies involved in experimental economics are far from perfect. There tends to be a bias in the subject pool as most participants are from a University community and hence may not have any experience with the problem at hand. In general the stakes are also not high enough to draw parallels with the real world. Additionally there tend to be a limited number of participants in any experiment and hence statistical analysis is imperfect. Finally, there are also concerns of validity (internal and external) and psychology of subjects.

Moreover, specific to this paper there are some techniques used which could have been improved upon. A better payment mechanism should have involved paying participants for every correct decision. Additionally, randomising the ordering across rounds in any session would have also been a more rigorous approach. In general increasing the subject pool from 23 to 50 (as was originally planned) would have provided more data and an opportunity to conduct an econometric analysis of the observations. There is also room for some sensitivity analysis - the results may be very sensitive on values of ε (E.g. - $\varepsilon = 0.9$ vs. $\varepsilon = 0.8$). Finally, the post experiment survey should have been conducted with multiple-choice answers.

Hence although the results from this paper provide conclusions in a certain direction, they should be interpreted with a pinch of salt due to the imperfections embodied in the experimental methodologies used.

IX. Conclusion

In any market place there would exist innumerable utility maximizing agents with incomplete information. Moreover these agents are connected and transmit signals to each other via the their respective actions. Additionally there is always an inherent uncertainty in the system as the underlying state may change from one time period to another. Understanding the informational and welfare properties of such a system would provide powerful tools for prediction and policy formation. Moscarini et. al (1996) discuss one such model where agents receive discrete signals and take once in a lifetime discrete action. The agents also receive a complete history of decisions of their predecessors but do not receive their signals. Moreover the underlying state of the world can change between agents and this process is assumed to be Markovian and independent of the current state of the world.

Relying on the experimental techniques in Anderson and Holt (1997), this paper empirically tests the main propositions of the changing world model. When the state of the world does not change, the results are in accord with theory - the group converges to rational cascades almost always when they have such an opportunity. Moreover, not more than 4 percent of all observations would be deemed inconsistent with the theory. When the state of the world always changes with ½ chance then also the results closely follow the theory - only 4 percent of all observations could be labelled as inconsistent with the model. Finally, when the state of the world changes almost always (90 percent chance) then also rational alternating cascade behaviour is observed almost always when it should. However, in this version of the experiment there is a higher propensity for subjects to make mistakes - almost 13 percent of all observations were inconsistent with the model. The higher frequency in this setup could be attributed to the more demanding calibration that would be required. The broad results from the post experiment questionnaire were in sync with the data from the experiment themselves. It is important to note that the general methodologies used in this paper are not without fault and hence the results must be read with caution. Finally, the extensions mentioned in this paper deserve due attention for further research.

BIBLIOGRAPHY

Allen, F. and Anna Babus, 2008, “Networks in Finance”, Working Paper 08-07, Wharton Financial Institutions Center, University of Pennsylvania.

Babus, A., 2008, “The Formation of Financial Networks”, Discussion Paper 06-093, Tinbergen Institute.

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

Anderson, L. and C. Holt, 1997, “Information Cascades in the Laboratory.” *American Economic Review*, 87(5), pp. 847-62.

Arthur, W.B., 1989, “Competing Technologies, Increasing Returns, and Lock-In by historical Events,” *Economic Journal*, XCIX 116-31

Ashiya, M. and Doi, T., 2001, “Herd behavior of Japanese economists”, *Journal of Economic Behavior and Organization*,Vol. 46, pp. 343–346.

Avery, C. and P. Zemsky, 1998, “Multi-Dimensional Uncertainty and Herd Behavior in Financial Markets”, *American Economic Review*. 88, pp. 724-48.

Banerjee, A., 1992, “A Simple Model of Herd Behavior.” *Quarterly Journal of Economics*, 107(3), pp. 797-817.

Banerjee A.V.,1989, “The Economics of Rumours,” mimeo 1992 (revised version. “Herd Behaviour and the Rewards for Originality,” mimeo Princeton University.

Banerjee, A.V.,1992, “A Simple Model of Herd Behaviour,” *Quarterly Journal of Economics*, 107(3), pp 797-817.

Bhattacharya, S., Chatterjee, K. and Samuelson, L., 1986, “Sequential research and the adoption of innovations”, *Oxford Economic Review*, Vol. 38, pp. 219–243.

Bikhchandani S., Hirshleifer D. and Welch I.,1992, “A Theory of Fads, Fashion, Custom and Cultural Change as Informational Cascades,” *Journal of Political Economy*, 100(5), pp 992-1026.

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

Bikhchandani, S. and Sharma, S., 2001, “Herd behavior in financial markets”, *IMF Staff Papers,*Vol. 47, pp. 279–310.

Brown, P., Foster, G. and Noreen, E.,1985, “Security analyst multi-year earnings forecasts and the capital market”, in *Studies in Accounting Research*, no. 21. (Sarasota, FL: American Accounting Association).

Burnum, J.F., 1987, “Medical Practice a la Mode: How Medical Fashions Determine Medical Care.” *New England J. Medicine* 317, pp. 1220-22.

Camerer C., 1995, “Individual Decision Making,” in J. Kagel and A. Roth eds., *Handbook of experimental economics*. Princeton, NJ: Princeton University Press, pp.587-616.

Camerer C. and Weigelt K, 1991, “Information Mirages in Experimental Asset Markets,” *Journal of Business*, October 1991, 64(4) pp. 463-93.

Celen, B. and Shachar Kariv, 2004, “Observational Learning Under Imperfect Information,”*Games and Economic Behavior*, 47(1), pp. 72-86.

Celen, B. and Shachar Kariv, 2004,“Distinguishing Informational Cascades from Herd Behavior in the Laboratory”, *American Economic Review*.94(3), pp. 484-497.

Celen, B. and Shachar Kariv, 2005, “An Experimental Test of Observational Learning under Imperfect Information,” *Economic Theory*.

Celen, B., Andrew Schotter and Shachar Kariv, 2003, “The Advice Puzzle: An Experimental Study of Social Learning where Words Speak Louder than Actions,” NYU.

Celen, B., Andrew Schotter and Shachar Kariv, 2009, “An Experimental Test of Advice and Social Learning”, Working Paper.

Chamley, C. and D. Gale, 1994, “Information Revelation and Strategic Delay in a Model of Investment.” *Econometrica*, 62(5), pp. 1065-85.

Chari, V.V. and Patrick J. Kehoe, 2004, “Financial Crisis as Herds: Overturning the Critiques”, *Journal of Economic Theory*, 119, pp. 128-150.

Christie WG and Huang RD,1995, “Following the Pied Piper: do individual returns herd around the market?” *Financial Analysts Journal* 51: 31-37.

Choi, S., Douglas Gale and Shachar Kariv, 2005, “Behavioral Aspects of Learning in Social Networks: An Experimental Study”, in John Morgan eds., *Advances in Applied Microeconomics, Volume 13, Behavioral and Experimental Economics.*

Choi, S., Douglas Gale and Shachar Kariv, 2009, “Social Learning in Networks: A Quantal Response Equilibrium Analysis of Experimental Data”,Working Paper.

Cotts Watkins, S.,1990, “From Local to National Communities: The transformation of Demographic Regions on Western Europe 1870-1960,”, *Population and Development Review*, XVI, 241-72.

Cukierman A, 1989, “Asymmetric Information and the Electoral Momentum of Public Opinion Polls,” mimeo Princeton University.

D’ Arcy S.P., Pyungsuk O., 1997,“ The Cascade Effect in Insurance Pricing”, *The*

Journal of Risk and Insurance, Vol 64, No 3, 465-480.

Dasgupta A. "Social Learning with Imperfect Observation and Payoff Complementarities," (LSE), work in progress.

Davis, R., 1991, *Dick Davis Digest.*

Davis D. and Holt C., 1993 *Experimental Economics*, Princeton, NJ: Princeton University Press.

DellaVigna, S and Ethan Kaplan, 2006, "The Fox News Effect: Media Bias and Voter Behavior", Working Paper.

Douglas G. "Bayesian Learning in Social Networks," 2003. Forthcoming, *Games and Economic Behavior.*

Drèze, J.H., 1974, "Axiomatic Theories of Choice, Cardinal Utility and Subjective Probability: A Review", in *Allocation under Uncertainty: Equilibrium and Optimality,* edited by J.H. Drèze, London: Macmillan, ch. 1, pp. 3-23.

Dryden J.,1964, "The Satires of Dreyden: Absalom and Achitopel, the medal, MacFlecknoe., London: Macmillan.

Economist, "Warren silverfinger", *The Economist*, 7 February 1998.

Ehrbeck, T. and Waldmann, R., 1996, "Why are professional forecasters biased? Agency versus behavioral explanations", *Quarterly Journal of Economics*, Vol. 111, pp. 21–40.

Farrell J. and G Saloner,1985 "Standardization, Compatibility and Innovation," *Rand Journal of Economics*, XVI, pp.70-83.

Fisher, Eric O'N., 1998, "Explaining Bubbles in Experimental Asset Markets".

Fisman, R., Daniel Markovits and Shachar Kariv, 2009, "Exposure to Ideology and Distributional Preferences", Working Paper.

Fong, G.T., D.H. Krantz and R.E. Nisbett, 1986, "The effects of statistical training on thinking about everyday problems", *Cognitive Psychology*, vol. 18, pp 253-92.

Fung, W. and Hsieh, D. A., 1999, "A primer on hedge funds", *Journal of Empirical Finance*, Vol. 6, pp. 309–331.

Galef, B.G., Jr. 1976, "Social Transmission of Acquired Behavior: A Discussion of Tradition and Social Learning in Vertebrates." in Jay S. Rosenblatt et. al eds. *Advances in the Study of Behavior,* Vol. 6, New York: Academic Press.

Gale, D. and Shachar Kariv, 2007, "Financial Networks", *American Economic Review, Papers & Proceedings*.

Gale, D. and Shachar Kariv, 2003, “Bayesian Learning in Social Networks”, *Games and Economic Behavior*, 45(2), pp. 329-346.

Gigerenzer, G., Peter M. Todd, and the ABC Research Group, 1999, *Simple Heuristics That Make Us Smart*, Oxford, Oxford University Press.

Givoly, D. and Lakonishok, J., ‘1984, “The quality of analysts’ forecast of earnings”, *Financial Analysts Journal*, Vol. 40, pp. 40–47.

Glosten, L. R. and Milgrom, P. R., 1985, “Bid, ask and transaction prices in a specialist market with heterogeneously informed traders”, *Journal of Financial Economics*, Vol. 14, pp. 71–100.

Graham J.R., 1999, “Herding among Investment Newsletters: Theory and Evidence” *The Journal of Finance* Vol LIV, No. 1.

Grenadier, S. R., 1999, “Information revelation through option exercise’”, *Review of Financial Studies*,Vol. 12, pp. 95–129.

Grether D., 1980,“Bayes Rule as a Descriptive Model: The Representativeness Heuristic.” *Quarterly Journal of Economics*, 95(3), pp. 535-57.

Grinblatt, M., Titman, S. and Wermers, R., 1995, “Momentum investment strategies, portfolio performance, and herding: a study of mutual fund behavior”, *American Economic Review*, Vol. 85, pp. 1088–1105.

Grossman, S. J. and Stiglitz, J. E., 1976, “Information and competitive price systems”, *American Economic Review*, Vol. 66, pp. 246–253.

Gul, F. and R. Lundholm, 1995, “Endogenous Timing and the Clustering of Agents' Decisions.” *Journal of Political Economy*, 103(5), pp. 1039-66.

Hammond, P.J., 1976, “Changing Tastes and Coherent Dynamic Choice”, *Review of Economic Studies*, 43, pp. 159-73.

Hammond, P.J., 1981, “Consistent Dynamic Choice Under Uncertainty and Bayesian Rationality”, *Economics Technical Report*, Institute for Mathematical Studies in the Social Sciences, Stanford University.

Hammond, P.J., 1982,”Utilitarianism, uncertainty and information”, in *Utilitarianism and beyond*, edited by Amartya Sen and Bernard Williams, Cambridge, Cambridge University Press, ch. 4, pp 90-91.

Hendricks, K. and Kovenock, D., 1989, “Asymmetric information, information externalities, and efficiency: the case of oil exploration”, *Rand Journal of Economics*, Vol. 20, pp. 164–182.

Hirshleifer, D and Siew Hong Teoh, 2003, “Herd Behaviour and Cascading in Capital Markets: a Review and Synthesis”, *European Financial Management*, Vol. 9, No. 1, 25–66.

Holland, J.H., K.J. Holyoak, R.E. Nisbett and P.R. Thagard, 1986, *Induction*, Cambridge, Massachusetts, MIT Press.

Holt C., Anderson L., 1996,"Classroom Games: Understanding Bayes' Rule." *Journal of Economic Perspectives*,10 (2), pp. 179-87.

Hung A.A. and Plott C.R.,2001, "Information Cascades: Replication and an Extension to Majority Rule and Conformity-Rewarding Institutions" *American Economic Review*, 91: 1508-1520.

Kahn, C. M., Pennacchi, G. G. and Sopranzetti, B., 2002, "Bank deposit rate clustering; theory and empirical evidence", *Journal of Finance*, Vol. 54, pp. 2185–2214.

Kahneman, D. and Tversky, A., 1979, "Prospect Theory: An analysis of Decision under Risk", *Econometrica*, 47, pp. 263-91.

Kahneman D. and Tversky A., 1973, "On the Psychology of Prediction." *Psychological Review*, 80 (4), pp. 237-51.

Kahneman, D., P. Slovic and A. Tversky, A., 1982, *Judgement Under Uncertainty: Heuristics and Biases*, Cambridge, Cambridge University Press.

Kahneman, D. and A. Tversky, 1982, "Judgments of and by representativeness", in Daniel Kahneman et. al. (Eds) *Judgement Under Uncertainty: Heuristics and Biases*, Cambridge, Cambridge University Press.

Kariv, S. ,2005, "Overconfidence and Informational Cascades," Under Revision.

Karni E, and D. Shmeidler, 1989, "Fixed Preferences and Changing Tastes," mimeo.

Katz M. and C. Shapiro, 1985, "Network Externalities, Competition and Compatibility,"*American Economic Review*, LXXV, 420-40.

Keller, G., Rady, S., 1995, "Optimal Experimentation in a Changing Environment." The University of Edinburgh Discussion Paper.

Keynes J., 1965, "The General Theory of Employment, Interest and Money" New York: Harcourt Brace & World.

Kiefer, N., 1991, "A Dynamic Model of Optimal Learning with Obsolescence of Information", Cornell University mimeo.

Kislev Y., and V. Shchori-Bachrach,1973, "The Process of an Innovation Cycle," *American Journal of Agricultural Economics*, p. 28-37.

Kyle, A. S., 1985, "Continuous auctions and insider trading", *Econometrica*, Vol. 53, pp. 1315–1335.

Lakonishok, J., Shleifer, A. and Vishny, R. W.,1992, "The impact of institutional trading on stock prices”, *Journal of Financial Economics*, Vol. 32, pp. 23–43.

Lohse, D., ‘Tricks of the trade: ‘‘Buffett is buying this’’ and other sayings of the cold-call crew’, *Wall Street Journal*, 1 June 1998.

Lee, I. H., 1993, “On the Convergence of Informational Cascades.” *Journal of Economic Theory*, 61(2), pp. 396-411.

Madrian, B. and Shea, D., 2000, “The power of suggestion: Inertia in 401(k) participation and savings behavior”, *NBER* Working Paper 7682.

Manktelow, K.I., and D.E. Over, 1990, *Inference and Understanding: A Philosophical and Psychological Perspective*, London: Routledge.

Maug E. and Naik N., 1995, “Herding and Delegated Portfolio Management: The Impact of Relative Performance Evaluation on Asset Allocation,” Working Paper, Duke University.

MacDonald, R.R., 1986, “Credible conceptions and implausible probabilities”, *British Journal of Mathematical and Statistical Psychology*, vol. 39, pp 15-27.

Mckelvey, R.D. and T.R. Palfrey, 1995, “Quantal Response Equilibria for Extensive Form Games.” Games and Economic Behavior, 10, pp. 6-38.

Mckelvey, R.D. and T.R. Palfrey, 1998, “Quantal Response Equilibria for Extensive Form Games.” *Experimental Economics*, 1, pp. 9-41.

Moscarini, G., Marco Ottaviani and Lones Smith, 1998, “Social Learning in a Changing World”, *Economic Theory*, 11: 657–665.

Nelson, L.-B., 2002, “Persistence and reversal in herd behavior”, *Review of Financial Studies*, Vol. 15, pp. 65–95.

Newell, A., and H.A. Simon, 1972, *Human Problem Solving*, Englewood Cliffs, NJ. Prentice-Hall.

Nisbett, R.E., D.H. Krantz, C. Jepson and Z. Kunda, 1983, “The use of statistical heuristics in everyday intuitive reasoning”, *Psychological Review*, vol. 90, pp 339-63.

Obrien, T. L. and Murray, M., “Buffett boosts American Express stake to 9.8%, acquires 8.3% of PNC Bank”, *Wall Street Journal*, 15 February 1995.

Park, A. and Hamid Sabourian, 2004, “Herding in Models of Sequential Trade and Monotonic Signals”, Mimeo, University of Toronto.

Perktold, J., 1996, “Recurring information cascades”, Working Paper, University of Chicago.

Persons, J. C. and Warther, V. A., 1997, “Boom and bust patterns in the adoption of financial innovations”, *Review of Financial Studies*, Vol. 10, pp. 939–967.

Plott C and Shyam Sunder.,1982, "Efficiency of Experimental Security Markets with Insider Information: An Application of Rational- Expectations Models." *Journal of Political Economy*, 90 (4), pp 663-98.

Plott, Charles and Shyam Sunder, 1988, "Rational Expectations and the Aggregation of Diverse Information in Laboratory Security Markets," *Econometrica*, 56:5, pp. 1085-1118.

Pollard, P. and J.S.tB.T. Evans, 1983, "The role of "Representativeness" in statistical inference: a critical appraisal", in J.S.tB.T. Evans (ed.) *Thinking and Reasoning: Psychological Approaches*, London, Routledge & Kegan Paul.

Post, Emily, 1927, *Etiquette in Society, in Business, in Politics and at Home*, New York: Funk & Wagnalls.

Ramsey, F.P., 1926, "Truth and Probability", in *The Foundations of Mathematics and Other Logical Essays*, edited by R. Braithwaite, London: Kegan Paul, pp. 156-98.

Rips, J. L., 1994, *The Psychology of Proof*, Cambridge, Massachusetts, The MIT Press.

Robin, E.D., 1984, *Matters of Life and Death: Risks vs. Benefits of Medical Care.* New York: Freeman.

Rustichini, A., Wolinsky, A., 1995, "Learning about Variable Demand in the Long Run", *Journal of Economic Dynamics and Control*, 19, 1283±1292.

Samelson W. and Zeckhauser R.1988, "Status Quo Bias in Decision Making," *Journal of Risk and Uncertainty*, 1 (1), pp 7-59.

Sandler, L. and Raghavan, A., "Salomon holders watch for possible buffeting", *Wall Street Journal*, 23 September 1996, p. C1.

Savage, L.J., 1954, *The Foundations of Statistics*, New York: John Wiley. Second revised edition – Dover, 1972.

Scharfstein D. and J. Stein,1990, "Herd Behaviour and Investment," *American Economic Review*, LXXX, pp.465-79.

Smith, L. and P. Sørensen, 1996, "Rational Social Learning by Random Sampling." MIT, mimeo.

Smith, L. and P. Sørensen, 2000, "Pathological Outcomes of Observational Learning." *Econometrica*, 68(2), pp. 371-398.

Smith, L. and P. Sørensen, 2008, "Rational Social Learning with Random Sampling".

Smith, Vernon L., Gerry L. Suchanek, and Arlington W. Williams, 1988, "Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets," *Econometrica*, 56:5, pp. 1119-1151.

Stern S., 1990,"The Effects of Firm Optimizing Behavior in Firm Matching Models," *Review of Economic Studies*. 57(4), pp 647-60.

Strotz, R.H., 1956, "Myopia and Inconsistency in Dynamic Utility Maximization", *Review of Economic Studies*, 23, pp. 311-28.

Sunder S., 1995, "Experimental Asset Markets: A Survey", in John H. Kagel and Alvin E. Roth eds., *Handbook of experimental economics*, Princeton University Press.

Taylor, R., 1979, *Medicine out of Control: The Anatomy of a Malignant Technology.* Melbourne: Sun Books.

Trueman Brett, 1994, Analyst Forecasts and herding behavior, *The Review of Financial Studies* 7, 97-124.

Tversky A. and Kahneman D., 1982b, "Subjective probability: A judgment of Representativeness", in Kahneman D., Slovic P. and Tversky A. eds., *Judgment under uncertainty: Heuristics and biases,* Cambridge University Press, Cambridge, U.K., 32--47.

Weizsäcker, G., 2008, "Do we follow others when we should? A simple test of rational expectations", Invited at *Amercian Economic Review*.

Welch I., 1992, "Sequential Sales, Learning and Cascades" *Journal of Finance*, 47(2) pp 695-732.

Wermers, R., 1999, "Mutual fund herding and the impact on stock prices", *Journal of Finance*, Vol. 54, pp. 581–622.

Zeira, J., 1999, "Informational overshooting, booms and crashes", *Journal of Monetary Economics*, Vol. 43, pp. 237–257.

Zhang, J., 1997, "Strategic delay and the onset of investment cascades", *RAND Journal of Economics*, Vol. 28, pp. 188–205.

Zwiebel J. 1995, "Corporate Conservatism and Relative Compensation", *Journal of Political Economy* 103, 1-25.