What have we learned from social learning?

Douglas Gale *

Economics Department, Boston University, Boston MA 02215, USA

Abstract

Models of herd behavior and informational cascades often make strong assumptions about the information available to agents, the nature of the choices being made, the timing of decisions, and the symmetry of equilibrium. This note considers the robustness of some results from the literature on social learning and argues that the inefficiency of equilibrium in the presence of informational externalities and strategic delay may be the most important lesson of all.

JEL classification: C7; D6; D8

Keywords: Learning; Delay; Information; Cascades

Social learning occurs in any situation in which agents learn by observing the behavior of others. Such a broad definition encompasses any situation in which there is asymmetric information, but recent research has focused on a more specialized set of situations, in which the existence of informational externalities leads to a loss of social welfare. In these situations, the attempt by agents to take advantage of the information of others can lead to the failure to exploit their own information in a socially optimal way. These models are broadly characterized by simple information structures and choices, in dynamic settings which allow us to study the completeness of information revelation and the rate at which information is revealed.

The class of models that are in some way related to the topic of social learning is vast, so some restrictions of scope are necessary. In the first place, I restrict attention to models with pure informational externalities, that is, model's in which

---

* Tel.: +1 (617) 353-6675; fax: +1 (617) 353-4449; e-mail: gale@bu.edu

an agent's payoff is unaffected by the actions of other agents. In these models, the only reason that an agent should care about the behavior of another agent is because it reveals payoff-relevant information, which allows the observing agent to make a better decision. Secondly, I restrict attention to models of rational behavior. Numerous studies of ad hoc learning rules or boundedly rational learning (e.g., Ellison and Fudenberg, 1993; Ellison and Fudenberg, 1995) could also be classed as dealing with issues of social learning, but they are outside the scope of this note. Thirdly, I exclude models of 'learning from experience', focusing instead on models in which each agent gets to make precisely one decision, so that learning results exclusively from observing others. In particular, this excludes the literature on optimal experimentation (e.g., Bolton and Harris, 1995) and the growing literature on learning in games (Fudenberg and Kreps, 1991) which deals with the problem of learning how to play an equilibrium strategy. What is left is admittedly a restrictive class of models, but it produces a sufficiently rich set of phenomena to be worth studying.

One phenomenon which has excited a great deal of interest is the existence of what is called 'herd behavior' or 'informational cascades'. A cascade occurs when agents 'ignore' their own information and imitate the behavior of other, supposedly better informed, agents. Herd behavior and informational cascades have been proposed as explanations of a variety of social phenomena, such as manias, social customs, panics, etc., as well as more mundane economic phenomena such as the failure of agents to adopt the most efficient technology.

The models on which these claims are based tend to be rather simple ones, and it is not surprising that some have questioned the robustness of the claims. That is one of the issues I shall want to discuss in this note. Briefly, I want to argue that while the phenomenon of 'herd behavior' or 'informational cascades' is robust, subject to the inevitable qualifications, the emphasis on this aspect of the models may be misplaced. From an economic point of view, it matters little whether information is asymptotically fully revealed, if it is revealed so slowly that it has little economic value ex ante. From a welfare perspective, this is hardly distinguishable from an 'informational cascade' in which information is suppressed by individuals' decisions.

The rest of the paper is organized as follows. Section 1 briefly describes a class of models and their characteristics. Section 2 discusses the robustness of the results. In Section 3, I discuss the importance of the results which, as has already been suggested, lies in the welfare properties of the equilibria rather than their informational properties per se. Finally, in Section 4, I address the question of why we need to learn from actions at all.

1. The basic model

For simplicity, I suppose that there are only two agents, $i = A, B$. All of the results I shall discuss generalize readily to n agents. Each of the agents receives a

signal, θ^i, which is correlated with some payoff relevant variable. Suppose for concreteness that the $\{\theta^i\}$ are independent and uniformly distributed on the interval $[-1, 1]$. To begin with, assume that each of the agents has to make a binary choice $x^i \in \{0, 1\}$, where $x^i = 1$ means that agent i chooses to 'invest' and $x^i = 0$ means that agent i chooses 'not to invest'. The actual return to investment is equal to $\theta^A + \theta^B$ for each agent, but since the signals are private information, each agent makes a decision under incomplete information.

Note that under this specification of the stochastic structure of the model, an agent's payoff is by assumption independent of what the other agent chooses to do. However, both of the signals contain information that is relevant for determining the underlying state of nature which determines the profitability of investment. Therefore, given the choice, each agent would like to observe the other agent's signal before making his or her own decision. Failing this, each agent would like to observe the other agent's action, since this will in all likelihood reveal something about the signal on which it is based.

The simplest temporal structure I can think of is one in which agent A is forced to make his decision first and agent B, having observed agent A's decision, makes her decision afterward. Note that this sequencing of decisions is exogenous and not determined by the agents. Given this sequencing of decisions, agent A's information set is $\{\theta^A\}$ and agent B's information set is $\{\theta^B, x^A\}$.

The (first-best) optimal decision is for both agents to invest if and only if $\theta^A + \theta^B > 0$. Since they do not know both signals, they have to condition their decisions on the information available to them. In the case of agent A, this means that he will invest if and only if $\theta^A > 0$, since the expected value of θ^B conditional on agent A's information is always 0. Then if agent B sees agent A invest, she knows that $\theta^A > 0$ and conditions her own decision on this fact. She will invest if and only if

$$\theta^B + E[\theta^A | \theta^A > 0] = \theta^B + \tfrac{1}{2} > 0.$$

Clearly, this is not as good as having complete information, and we can easily imagine situations in which both agents invest even though $\theta^A + \theta^B < 0$. For example, if θ^A is positive but very small and θ^B is only slightly greater than $-\frac{1}{2}$ both agents will invest even though the true return to investment will be negative and may be quite large in absolute value.

Thus, agent B chooses to invest because agent A invests, even though, if she had to make a decision based solely on her own (negative) signal, she would choose not to invest. This is the sense in which agent B may be led to 'ignore' her own information and imitate agent A, even though agent A's information is, in a sense, much less informative than B's.

In a nutshell, this is an example of herd behavior or an informational cascade. In the models of Banerjee (1992) and Bikhchandani et al. (1992) (BHW) there is a sequence of n agents, each of whom has to make the same decision. In Banerjee's case, they have to choose a number from an interval. In BHW's case, they make a

binary choice. Both demonstrate that if two agents choose the same action, every subsequent agent will make the same choice, regardless of the signal she received. Imitation dominates private information. In terms of the present example, suppose we add agents $C, D,...$ who receive signals $\theta^C, \theta^D, ...$, independently and uniformly distributed on $[-1, 1]$, and make the return to investment equal to $\theta^A + \theta^B + ...$. The analysis of the decisions of agents A and B is unchanged and if we go through the same argument for agents $C, D,...$ we shall find that if agents A and B invest, agent C will invest as long as $\theta^C > -3/4$; that if agents A, B and C invest, agent D will invest as long as $\theta^D > -7/8$; and so on. The longer the sequence of agents who have invested, the harder it is for a single agent not to invest, even if his or her information is very negative. Morever, each successive agent who invests in spite of a negative signal, suppresses her own information and makes it harder for subsequent agents to invest.

Another observation made by BHW concerns the 'fragility' of behavior in the presence of cascades. Suppose, for example, that agents A, B, and C invest and agent D observes a signal $\theta^D < -7/8$, so that it is optimal for her not to invest. Then her action cancels out the accumulated experience of the preceding agents, so that agent E will be unwilling to invest even if she observes a small positive signal. No matter how many agents have invested previously, it is always possible for a very bad signal to one agent to cancel out the history up until that point. The reason, clearly, is that only the first few agents who decide to invest are revealing a substantial amount of the information available to them. For the rest, their decision reveals next to nothing, precisely because the region in which it is optimal to invest is almost the entire interval. For example, if agent D decides to invest, we learn only that her signal falls in the range $-7/8 < \theta^D < 1$. On the other hand, a decision not to invest reveals a lot, namely, that her signal falls in the narrow range $-1 < \theta^D < -7/8$. The longer the run of decisions to invest, the greater the asymmetry between the amount of information revealed by the decision to invest and the decision not to invest.

The question we need to consider next is the robustness of these results.

2. How robust is herd behavior?

Although the model sketched in the preceding section is very simple, a number of significant modeling choices have already been made. First, the timing of decisions is exogenous. Secondly, the choice made by agents is discrete; in the example it is a binary choice. A reader familiar with Banerjee's model will have noted that the choice variable is continuous, but under his maintained assumptions it shares the properties of a discrete choice model. Agents either observe the one profitable choice or observe a purely noisy signal, so agreement by two agents is taken as evidence that they both know the truth. Third, the signals received by agents have been assumed to be continuous, although most of the results discussed

made no essential use of this property. Fourth, time is implicitly treated as discrete. In a model with endogenous sequencing of decisions, we could choose to model time as either a discrete or a continuous variable. Finally, an issue that is not salient yet is whether to focus on symmetric or asymmetric equilibrium. Each of these modeling decisions raises the prospect of more complex models, with different properties, and raises issues about robustness and realism and the 'correct' choices to make. I discuss these options in turn.

2.1. Signals

From what we have seen so far, it is not clear that the nature of the signal received by the agents is crucial to the results and intuition suggests that it is not. The simple reason is that each agent's signal is filtered by his or her action. If the agent makes a binary choice, then the most that other agents can observe is a binary partition of the agent's signal space. The exact nature of the signal should not matter and acquaintance with a variety of models suggests to us that this is indeed the case.

One place where the continuity of signals does matter is in the example we offered above of the 'fragility' of cascades. If agents only receive two possible signals, one good and one bad, then as soon as two agents have made the same choice, all subsequent agents must make the same decision. In order to have any hope of breaking this cascade, one would have to introduce an agent with a better signal or something of that sort.

2.2. Actions

Unlike the choice of signals, the choice of the action space for the agents does seem to make a difference. It is easy to think of examples in which agents have preferences defined over a continuous variable and any change in their information is reflected in the choice they make. However much they may be influenced by the behavior of the other agents, there is no reason why they should not take full advantage of their own information. Gul and Lundholm (1992) have explored a model in which agents make a continuous choice and exhibit an equilibrium in which the agent's action is a sufficient statistic for his or her information. Eventually all the information is revealed. This aspect of social learning needs to be explored further.

It is interesting to note that the discreteness of an agent's choice set in our basic example plays two roles. In the first place, the discreteness of the action set prevents an agent from revealing all of his or her information to other agents. This will always be the case if the signal is continuous and the action set is discrete. Secondly, the discreteness of the action set prevents the individual from making use of all the information at her disposal, for example, 'ignoring' her own

information and basing her decision on imitation of the other agent's action. These are two sides of the same coin, but it is helpful to distinguish them.

Even if we allow agents to make continuous choices, there is still an important difference between action and inaction. Because there may be fixed costs of taking an action, the choice between action and inaction involves a kind of lumpiness or indivisibility, even when the action is represented by a continuous variable. An important feature of the Gul-Lundholm model is that inaction is not allowed: agents must make a choice from the continuum, which then reveals their information. Introducing the possibility of inaction would allow us to produce herd behavior or cascades, for some parameter values, just as in the binary choice model.

2.3. Endogenous sequencing

An obvious feature of the basic example is that an agent who moves later will do better on average. Additional information can only help. If the agents were allowed to choose their position in the decision-making queue, they would all want to be last. Since someone has to go first, the stage is set for a conflict in which each agent maneuvers to get the best place in the decision-making queue. Strategic delay, motivated by informational externalities, has been studied by Chamley and Gale (1994) and Gul and Lundholm (1992). The analysis of models with endogenous timing raises a number of issues not found in models with exogenous timing and I return to these later. But first it is interesting to ask what would happen to our earlier example if the timing of investment decisions were made endogenous.

Let us go back to the original two-agent example and suppose that time is divided into two periods and that agents can invest in either period, subject only to the restriction that their decisions are irreversible. If an agent invests in the first period, the decision must be based on private information only. If the agent waits until the second period, there is a possibility that the other agent will have invested in the first period, and that will reveal some information to the second agent about the first agent's signal. There is also a cost of waiting, which is modeled by assuming that both agents discount the future using a common discount factor $0 < \delta < 1$.

If an agent chooses to invest in the first period, the expected return to investment is equal to θ^i. If the agent were to wait one period, the cost of this delay in the form of discounting of future utilities, would be $(1 - \delta)\theta^i$. Agents with high values of θ^i are more impatient to invest and therefore less likely wait. Agents with low values of θ^i are less impatient. Obviously, if $\theta^i < 0$ then agent i will not invest in the first period. However, even if θ^i is slightly positive, the agent may decide to wait because of the option value of delay.

In order for there to be any option value at all, the decision made at the second date must depend in a non-trivial way on the outcome at the first date. This means that an agent who waits until the second date will invest only if the other agent

invested at the first date. Suppose that an agent i invests at the first date if and only if $\theta^i > \bar{\theta}$. If agent i waits and agent $-i$ does not invest, the return to investing at the second period is

$$\theta^i + E\left[\theta^{-i}|\theta^{-i} < \bar{\theta}\right] < 0.$$

The probability of this happening is $P[\theta^{-i} < \bar{\theta}]$. The option value of delay is the expected loss that agent i avoids by not investing in the event that agent $-i$ does not invest at the first date. The marginal type $\bar{\theta}$ is indifferent between investing in the first period and waiting until the second. For this type the costs of delay just balance the option value of delay:

$$(1 - \delta)\bar{\theta} = -P\delta\left[\theta^i < \bar{\theta}\right]\left\{\bar{\theta} + E\left[\theta^i|\theta^i < \bar{\theta}\right]\right\}.$$

It is straightforward to show that there is a unique value of $\bar{\theta} > 0$ satisfying this equation and that an agent will prefer to invest in the first period or wait until the second according to whether her signal is greater than or less than $\bar{\theta}$. Furthermore, an agent who chooses to wait will find it optimal to invest if and only if someone invested in the first period, that is, $\theta^{-i} > \bar{\theta}$.

Although this is a very simple equilibrium, it has a number of interesting features. First, we see again the fact that information is not fully revealed. Secondly, agents are likely to make mistakes and they may ignore their own information. Even if both agents receive positive signals, neither will invest if the signals are less than $\bar{\theta}$. Third, the game really does end in two periods. If we added an infinite number of periods, the play of the game would be exactly the same. By the end of the second period, every agent has either invested or decided never to invest. If no one invests in the first period, no new information is revealed in the second period and so the agents' decisions will be exactly the same in every subsequent period, that is, it is not optimal to invest.

If there were more than two agents, the play of the game could continue for several periods. But it must still be the case that if no one invests in one period, investment stops forever. This possibility of investment collapse is always present and is a necessary condition of equilibrium. In order to have delay there must be a positive option value and there can only be a positive option value if there is a positive probability that the agents will never invest.

Since the play of the game ends in two periods (or more generally in a finite number of periods), decisions are resolved very quickly when the period length is very short. This does not lead to efficiency. Although there is no delay in resolving decisions, there is still a welfare loss from 'bad' investment decisions.

2.4. Time

In the Gul and Lundholm (1992) model, time is continuous and an agent's choice of when to act perfectly reveals his information. An interesting question is

what happens in the present example if time is continuous? Unfortunately, there is no symmetric equilibrium in this case. To see why this is so, suppose that we could find an equilibrium in which type $\theta(t)$ invests at time t. At that moment, an agent with a signal $\theta^i = \theta(t)$ must be indifferent between investing immediately and waiting until $t + dt$. The cost of delay is $r\theta(t)dt$, where r is the agent's instantaneous rate of time preference. The gain is the option value of seeing whether the other agent will invest in the interval $(t, t + dt)$. If the other agent invests, we learn that $\theta^{-i} = \theta(t)$; if not, we learn that $\theta^{-i} < \theta(t)$. Now agent i already knows that $\theta^{-i} < \theta(t)$ at time t and he is willing to invest. But if he is willing to invest when he knows that her signal is $\theta^{-i} < \theta(t)$ he must strictly prefer to invest when he learns that $\theta^{-i} = \theta(t)$. What this means is that he will invest whatever happens, so the option value is zero. Thus, he is not indifferent to a short delay, and in fact, would rather invest earlier. The equilibrium unravels.

The problem here is that when the agent has a binary choice to make, the option value is either zero or large. To support an equilibrium in continuous time there must be a small option value to balance the small cost of (a short) delay. For this, agents must be able to make small adjustments in their choices, that is, we have to model their actions as continuous variables. For example, suppose that agent i's utility function is $u(\theta^i + \theta^{-i}) = -(\theta^A + \theta^B - x^i)^2$ and x^i can be any real number. Then as time passes, the interval in which agent i believes the other agent's signal must lie is changing gradually and agent i will change his value of x^i accordingly. The ability to make small adjustments to x^i generates exactly the small option value that we need to support an equilibrium like that of Gul and Lundholm.

This is a technical point, but as long as the outcome of the analysis depends on such points, care has to be exercised. In some sense, the choice of whether to model time as discrete or continuous should not matter to the economic results. The fact that the choice of continuous time pushes us in the direction of a revealing equilibrium, because it seems a more natural or tractable equilibrium in that context, is not an indication that this is the 'truth'. Clearly, a deeper analysis of these issues is required before we can claim with any confidence to understand all the subtleties behind these modeling issues or address the robustness issue with authority.

2.5. Symmetric or asymmetric equilibria?

So far, we have considered only symmetric equilibria. Does this matter? It does. In the preceding example, with the continuous choice variable, there always exist asymmetric equilibria in which there is no delay. Suppose that agent A always expects agent B to go first, that is, at any point in time t, agent A thinks that agent B will make her choice in the interval $(t, t + dt)$ if she has not already done so. Given this belief, it will never be optimal for agent A to make his

decision until he has observed agent B's, and given that agent A will not make a decision until he has observed agent B's move, it is never optimal for agent B to delay. So in equilibrium, agent B makes her decision at time 0 and agent A makes his decision ϵ periods later.

The existence of this equilibrium depends crucially on the assumption of continuous time. If agent A expects agent B to reveal her information immediately, it is always worth an infinitesimal delay to obtain the information.

It is also important that in this example there is no possibility of inaction. If, for some values of θ^i, agent i did not want to choose any value of x^i at all, things would not work out so simply. If there are states of nature in which agent B does not choose any action, it is possible that agent A will sometimes want to choose x^A without observing agent x^B. In that case, agent B might have an incentive to delay in hopes of deceiving agent A into revealing his choice.

I have some doubts about the relevance of asymmetric equilibria in practical applications, but there is no doubt that, in some environments, it makes a great deal of difference to the outcome whether one assumes that equilibrium is symmetric or asymmetric.

3. The importance of social learning

From the preceding discussion, it appears that certain indivisibilities are crucial for herd behavior or informational cascades, at least in the class of models considered. A more generous way of putting the matter might be to say that the signal space has to be large relative to the action space. This is something that has been clear since the development of the rational expectations literature and hardly counts as the central contribution of the social learning literature. What is new is the observation that, with indivisibilities, the suppression of information is amplified. Each additional agent who is absorbed into the herd or cascade takes with him his or her information, thus depriving the subsequent decision makers of its benefit. The fact that indivisibilities are required to make the story work does not strike me as a fatal limitation. There are many situations in which agents do indeed make 'lumpy' or indivisible decisions and the model of choice found in much of the social learning literature is a sensible one for these applications. Thus, even if indivisibilities are in some sense essential for these results, the results may nonetheless be considered robust within an interesting class of models.

Another issue is whether the possibility of herd behavior or informational cascades is economically important. As we have already mentioned, herd behavior and cascades have been offered as explanations of a variety of phenomena, and if they are accepted as plausible explanations, they may have extended our understanding of these phenomena in important ways. While not dismissing the importance of herd behavior or informational cascades, we would choose to

emphasize a different aspect of the literature. From the point of view of economic welfare, it matters little whether there is an informational cascade or whether information is fully revealed after a long delay. Once we endogenize the timing of decisions and recognize that delay is costly, we encounter a new possibility, that all the information initially possessed by agents will be revealed asymptotically, but so slowly that economic welfare is much less than the first best. The work of Caplin and Leahy (1993), Caplin and Leahy (1994), Chamley and Gale (1994), and Gul and Lundholm (1992) makes this point clearly. The important lesson is not so much that delay is an important phenomenon in its own right, though I believe it is and that was the initial motivation for my research in this area; rather, it is that informational externalities result in serious failures to achieve a desirable social outcome. Whether the outcome is delay, or incomplete revelation of information, or some other market failure, the important ingredient is the free rider problem and the failure to internalize an informational externality.

4. Why learn from observing actions?

Stepping back from the social learning models themselves, we can ask some broader questions about the value of this research program. One question often asked is "Why do the agents need to learn by observing each others actions? Why can't they just exchange the information directly?" In other words, why not allow 'cheap' talk? Would this solve the problem? The quick answer is that there may be many situations in which communication is anything but 'cheap'. However, even if we take the possibility of cheap talk seriously, it is not clear that this will alter the conclusions of the analysis. Let's go back to the basic example and ask what would happen if agents were allowed to engage in one round of cheap talk before the beginning of the game. In one equilibrium both agents announce their true signals and the first-best decisions are made. In another equilibrium, however, both agents 'babble', in which case cheap talk has no effect on the equilibrium outcome at all. In general, the full communication equilibrium is only one among many.

When we consider the robustness of these equilibria, there is little reason to think that cheap talk improves matters. Glazer and Rubinstein (1994) have analyzed a model in which it is impossible to implement the full communication equilibrium in Bayes–Nash equilibria. In other words, one cannot get rid of the multiplicity of equilibria by designing an appropriate mechanism. In an unpublished analysis of cheap talk, I have shown that, under some mild regularity conditions, the only equilibrium which is robust to the introduction of an arbitrarily small cost of communication, is the equilibrium in which there is no communication.

The models we have reviewed are very simple and, in particular, they do not

admit prices. Economic theory assigns an important role to prices as purveyors of information, so the absence of prices may be a serious limitation. Avery and Zemsky (1995) have argued that allowing prices to aggregate public information can prevent informational cascades. To make this point, they use a variant of the Glosten and Milgrom (1985) asset-market model, in which a competitive market maker sets bid and ask prices equal to the expected value of the asset conditional on all publicly available information. Agents arrive randomly and buy or sell one unit of the asset at the posted prices. Some agents buy or sell for liquidity reasons (noise traders); others have private information which they want to exploit for profit (informed traders). The market maker tries to infer the information available to the informed traders from the buy and sell orders, but the presence of noise traders gives the informed traders an informational advantage. Eventually, the preponderance of trades on one side of the market or the other will tell the market maker whether his price is too high or too low and asymptotically the true value of the asset is revealed. Avery and Zemsky conclude that the presence of a continuous variable (price) that aggregates public information prevents the occurrence of a cascade.

This is a very important observation, but not necessarily a critical one for the theory. In the Avery–Zemsky model, there are no gains from trade. The only reason that informed traders want to buy or sell is because their valuation of the asset differs from that of the market maker. However small the difference in their valuations, the informed trader will buy (sell) if and only if her valuation is higher (lower) than that of the market maker. Regardless of how much information has been revealed in the past, the informed trader's decision reveals (some of) her own information. However, if traders had other motives for trade in addition to private information, a hedging motive, for example, the analysis would be quite different.

More research is undoubtedly needed before we shall understand fully the impact of prices on social learning. In an important series of papers, Xavier Vives has explored the dynamics of information revelation in asset markets using competitive models of rational expectations equilibrium (see, e.g., Vives (1993)). Strategic interaction is ruled out by the competitive assumption, indivisibilities do not play a role, and agents only learn from prices. However, the same kinds of issues concerning the speed and efficiency of information revelation are addressed, in a class of models that allow a role for prices and much richer behavior on the part of agents. In Vives' model, trade is motivated by asymmetric information (informed traders) and liquidity shocks (noise traders). Because of the presence of noise traders, it takes time for the truth to be revealed. However, standard sampling theory suggests that the speed of convergence to the true asset value should be inversely proportional to the square root of the number of observations. What Vives shows is that the rate of convergence will be much slower because informed traders do not have the right incentives to reveal their information. The challenge to social learning theorists is to emulate Vives' work in developing richer models of social learning.

5. For further reading

Bala and Goyal (1993),Hendricks and Kovenock (1989),Rob (1991).

Acknowledgements

References

Avery, C. and P. Zemsky, 1995, Multi-dimensional uncertainty and herd behavior in financial markets, Unpublished.

Bala, V. and S. Goyal, 1993, Learning from neighbors, Unpublished.

Banerjee, A., 1992, A simple model of herd behavior, Quarterly Journal of Economics 107, 797–817.

Bikhchandani, S., D. Hirshleifer and I. Welch, 1992, A theory of fads, fashion, custom, and cultural change as informational cascades, Journal of Political Economy 100, 992–1026.

Bolton, P. and C. Harris, 1995, Strategic experimentation, Unpublished.

Caplin, A. and J. Leahy, 1993, Sectoral shocks, learning and aggregate fluctuations, Review of Economic Studies 60, 777–794.

Caplin, A. and J. Leahy, 1994, Business as usual, market crashes, and wisdom after the fact, American Economic Review 84, 548–566.

Chamley, C. and D. Gale, 1994, Information revelation and strategic delay in a model of investment, Econometrica 62, 1065–1085.

Ellison, G. and D. Fudenberg, 1993, Rules of thumb for social learning, Journal of Political Economy 101, 612–643.

Ellison, G. and D. Fudenberg, 1995, Word-of-mouth communication and social learning, Quarterly Journal of Economics 110, 93–125.

Fudenberg, D. and D. Kreps, 1991, Learning and equilibrium in games, Unpublished.

Glazer, J. and A. Rubinstein, 1994, The design of organizations for collecting information from conformist agents, Unpublished.

Glosten, L. and P. Milgrom, 1985, Bid, ask and transaction prices in a specialist market with hetergeneously informed traders, Journal of Financial Economics 14, 71–100.

Gul, F. and R. Lundholm, 1992, On the clustering of agents' decisions: Herd behavior versus the timing of actions, Unpublished.

Hendricks, K. and D. Kovenock, 1989, Asymmetric information, information externalities, and efficiency: The case of oil exploration, Rand Journal of Economics 20, 164–182.

Rob, R., 1991, Learning and capacity expansion under demand uncertainty, Review of Economic Studies 58, 655–675.

Vives, X., 1993, How fast do rational agents learn? Review of Economic Studies 60, 329–348.