June 6, 2002

Taking the Road Less Traveled:

Does Conversation Eradicate Pernicious Cascades?

H. Henry Cao*

David Hirshleifer**

*Kenan-Flagler School of Business, UNC, Chapel Hill, NC, 27599-3490, caoh@bschool.unc.edu, http://www.kenan-flagler.unc.edu/faculty/directory/189.html

**Fisher College of Business, The Ohio State University, 740A Fisher Hall, 2100 Neil Avenue, Columbus, OH 43210-1144; hirshleifer_2@cob.osu.edu, http://www.cob.ohio-state.edu/fin/faculty/hirshleifer/

We thank Abhijit Banerjee, Jonathan Berk, Steven Durlauf, Drew Fudenberg, Jack Hirshleifer, Hayne Leland, Seongyeon Lim, Siew Hong Teoh, Ivo Welch and seminar participants at the University of California, Berkeley, the University of Chicago, and the University of Wisconsin for helpful comments.

Taking the Road Less Travelled:

Does Conversation Eradicate Pernicious Cascades?

We offer a model in which sequences of individuals often converge upon poor decisions and are prone to fads, despite being able to communicate both past payoff outcomes and the private signals underlying past choices. This reflects direct and indirect action-based informational externalities; and conversational externalities—the failure of individuals to take into account the benefits their conversations confer upon later individuals. In contrast with previous cascades literature, cascades here are spontaneously dislodged and in general have a probability less than one of lasting forever. Furthermore, the ability of individuals to communicate can reduce average decision accuracy and welfare.

Two roads diverged in a wood, and I—
I took the one less traveled by,
And that has made all the difference.

—Robert Frost, "The Road Not Taken," from *Mountain Interval*. New York: Henry Holt and Company, 1920; Bartleby.com, 1999. www.bartleby.com/119/. [1/15/02].

1 Introduction

Social observers have often commented on the 'folly,' 'fickleness,' or 'madness' of mass behavior. Ideally, however, if large numbers of rational individuals were to communicate their imperfectly correlated information signals, we would expect convergence toward a very high quality of decisions.

A popular explanation for the failure to achieve this benchmark is that people are imperfectly rational. An alternative explanation that has been explored in a literature on informational cascades or herding is that the rational choices of individuals cause information to be aggregated poorly (see, e.g., Banerjee (1992), Bikhchandani, Hirshleifer and Welch (1992)).

In this approach, an individual's use of information gleaned from observation of predecessors makes his own action less informative to later observers. Such an individual will sometimes find it optimal to choose an action consistent with the choices or experience of others regardless of his own (possibly opposing) private signal. In such a situation he is in an *informational cascade* or *herd*, and his action is uninformative to later individuals. Under appropriate conditions, cascades/herds eventually form with probability 1. Furthermore, if individuals are ex ante identical, and the only sources of information available to an individual is his private signal and past actions, then all subsequent individuals are also in a cascade, and information aggregation ceases. Thus, the system reaches an equilibrium in which the great mass of individuals make relatively ill-informed decisions. As a result, decisions are fragile: individuals are prone to shifting their behavior in the event of even a modest external shock.

The analysis of cascades/herding has been extended to address various applications.[1]

[1] These include the purchase of initial public offerings (Welch (1992)), delay and fluctuations in investment (Chamley and Gale (1994), Zhang (1997), monopolistic product pricing (Ottaviani (1999), Bose, Orosel and Vesterlund (2001)), Chamley (2001)), macroeconomic fluctuations (Caplin and Leahy (1994), optimal managerial contracting (Khanna 1997), Chalkley and Lee (1998)), the winner's curse in auctions (Neeman and Orosel (1999)), the design of organizations (Khanna and Slezak (2000)), financial crises (Chari and Kehoe (2000), Dasgupta (2000)), the evolution of overconfident beliefs (Bernardo and

In the basic cascades/herding approach, individuals can observe the actions of predecessors but do *not* observe either past private information signals, or payoffs from past actions. Thus, models of cascades/herding rule out by assumption some potentially profitable communication activity.

Such an assumption is reasonable in some contexts, especially when the relevant individuals are not personally acquainted, when decision outcomes take a long time to resolve, or when conflicts of interest interfere with communication. However, in many situations verbal communication of the information or experiences of previous decisionmakers is important, from everyday choices people make among restaurants or auto mechanics to the diffusion of new technologies among potential adopters. There is evidence that verbal communication between individuals exerts a strong influence on purchasers' judgments about a variety of different kinds of consumer products and business factor inputs.[2] Conversation also influences individuals' investment decisions.[3]

However, if conversation perfectly conveyed *all* information (and individuals were rational), everyone would be equally informed, and everyone would make equally good decisions. The unrealism of this implication suggests that assuming *perfect* communication may be too extreme.

In this paper, we take as given that there are some barriers to communication, and explore whether these barriers clog the information pipeline so tightly that large numbers of individuals make ill-informed decisions. We have in mind such applications as the spread of agricultural innovations in a neighborhood, the creation of 'buzz' about a new movie or initial public offering of stock, the choice of whether to invest retirement funds in stocks or bonds, or the choice of medical procedures by doctors in a geographical region.

Welch (2000)), order of speech in debates (Ottaviani and Sorensen (2001)), and securities market fluctuations (Lee (1998), Avery/Zemsky (1998), and Cipriani and Guarino (2001)). An experimental literature has verified the empirical relevance of cascades/herding (see, e.g., Anderson and Holt (1996), and Celen and Kariv (2001)). For recent reviews of informational cascades/herding theory, social learning, and applications see Bikhchandani, Hirshleifer and Welch (1998) and Brunnermeier (2001).

[2]These include, for example, the adoption of agricultural techniques and the adoption of new computer technology. Several empirical studies are discussed in Herr, Kardes and Kim (1991). Shiller (2000; ch. 8) discusses the importance of the reactions of early viewers of new movies and the resulting word-of-mouth 'buzz' that they create.

[3]A survey by Shiller and Pound (1989) asked individual investors what first drew their attention to the company whose stock they had purchased most recently. Almost all investors named sources which involved direct personal interaction; direct personal interaction was also important for institutional investors. Shiller (2000, ch.8) cites several other studies indicating the crucial role that conversation plays in security investor decisions. Furthermore, Kelly and O'Grada (2000) and Hong, Kubik and Stein (2001) provide evidence that social interactions between individuals affects decisions about equity participation and other financial decisions.

We extend the cascades model of Bikhchandani, Hirshleifer and Welch (1992) to permit full communication of information about payoffs, or some degree of communication of past private signals.

Even with limited communication, there is some reason to suspect that individuals will eventually converge upon correct outcomes. If at any point in time an inefficient cascade starts, the information blockage can spontaneously be dislodged by either the repeated arrival and communication of new private signals or of new payoff information. It seems plausible that such information arrivals would sooner or later make decisions arbitrarily accurate.

In contrast with this intuition, our analysis shows that under a fairly broad set of circumstances, moderate limits to observability and communication cause individuals to fall into inefficient cascades. Even individuals very late in the decision queue make inaccurate decisions frequently. The reason that efficiency is not achieved is that information aggregation is limited by three types of informational externalities.

The first type, *direct action-based informational externalities*, is also present in the basic cascades/herding model: individuals fail to take into account the benefit that adjusting their actions in response to private signals confers upon later decision makers. In addition, in our model there are also *indirect action-based informational externalities*, and *conversational externalities*.

In the former, an individual's action choice affects the information of later individuals indirectly by generating informative payoff outcomes.[4] As William Blake put it, "If others had not been foolish, we should be so." (*The Marriage of Heaven and Hell*, 1793?).

In the latter, conversational externalities, an individual's decison to expend resources acquiring private information from others through conversation confers an external benefit. The more private information he obtains from predecessors, the more he has to pass on to others. Owing to this externality, we find that even if all information could be communicated at low cost, it may not be, leading to inefficient outcomes.

In contrast with the basic cascades/herding model, we find that cascades/herds have a probability of being spontaneously dislodged— endogenously, not owing to an external shock. Thus, cascades in general have a probability between zero and one of lasting

[4]The information about payoff may be acquired by direct observation, or by conversation; our model does not distinguish these alternatives. Previous models in which there are informational externalities associated with observation of past payoffs include Vives (1993), Caplin and Leahy (1993) and Persons and Warther (1997). However, these papers do not explore how indirect action-based informational externalities influence the formation of cascades/herds.

forever.[5] In other words, there is a significant probability of realizations in which individuals choose one of the possible actions only a finite number of times. If this happens, uncertainty about the quality of the project chosen infinitely often disappears, but uncertainty about the alternative project persists forever. So individuals in the model will always have reason to wonder wistfully whether the other project, "the road less traveled by," was in fact the better one.

Even if communication of past payoffs alone does not always dislodge inefficient cascades, communication by newly arriving individuals with private information about both project alternatives may do so. However, we find that the introduction of (limited) communication of private signals does not suffice to eradicate inefficient cascades.

If each individual can, through conversation, acquire the private signal of his immediate predecessor, we find that inefficient cascades/herds still have a positive probability of lasting forever. Furthermore, we consider a setting where each individual can, at a cost, learn through conversation *all* the information possessed by his immediate predecessor–including any information the predecessor has acquired about earlier individuals' signals. So long as the cost of obtaining this information is positive, no matter how small, the informational chain is eventually broken. Thus, even virtually free speech results in cascades/herds and inefficient outcomes in the long-run.

Finally, we explore by means of numerical examples whether the ability to observe past payoffs in addition to actions improves welfare and decision accuracy. We find that observation or communication of past payoffs can trigger inefficient cascades/herding earlier, by rendering the decisions of early individuals less informative (both directly and indirectly) to subsequent individuals. In consequence, delay in communication of payoffs can improve average decision quality and welfare.

A previous literature on word-of-mouth learning has explored the conditions under which different individuals converge upon similar outcomes, and under which they achieve efficient outcomes. The assumptions of the word-of-mouth literature are different in several ways from those in the herding/cascades literature, making comparisons non-obvious. Much of the word-of-mouth learning literature is based on 'rule-of-thumb' decision making, which leaves open the question addressed here: do fully rational individuals converge to efficient outcomes?[6] In addition, in continuum models aggregate

[5] Realized payoffs are stochastic given ex ante project quality, so a single observation does not in general resolve project quality.

[6] Ellison and Fudenberg (1993, 1995) examine settings in which successive continua of individuals make simultaneous decisions each generation. Players observe a subset of past actions and signals about past payoffs, and make choices using simple rules of thumb that are potentially consistent with

social outcomes are non-stochastic and path-independent, whereas the cascades/herding approach emphasizes the effects of early chance signal realizations on social outcomes.

The remainder of the paper is structured as follows. Section 2 provides a simple example to illustrate how inefficient cascades when there is conversation about or observaton of past payoffs. Section 3 examines a more general model with transient shifts in payoffs and/or noisy observation/communication of payoffs. Section 4 provides an example to illustrate how observation/communication of payoffs can, ex ante, reduce expected welfare and decision accuracy. Section 5 examines the effect of conversation that conveys previous individuals' private signals. Section 6 concludes. Except where otherwise noted, proofs are in the Appendix.

2 Conversation about Payoffs and Cascades: Definitions and an Illustrative Example

This section provides a simple example in which the start of a cascade/herd ends all experimentation. If such a cascade starts, one of the choice alternatives, 'the road less traveled by,' will (in this simple example) never be tried. Thus, inefficient cascades can occur with positive probability, and once started last forever (in the absence of exogenous shocks). In the more general model of Section 3, we will see that inefficient cascades can occur even if there is a great deal of experimentation. We will see there that when payoffs are subject to transient shocks or are observed with noise, inefficient cascades can start even when all choice alternatives have been tried any number of times; and have only a probability of lasting forever.

There is reason to think that it is easier to convey through conversation information about past payoffs than about past private signals.[7] Thus, until Section 5, we will assume that conversation conveys information about past action choices or payoffs, but not the private reasons for decisions (i.e., the spontaneously arriving private signals).

rationality. The aggregate state of the population evolves deterministically. Players can converge to the efficient choice even for rules of thumb that are quite naive. Banerjee and Fudenberg (1999) explore herding and efficiency among rationally optimizing individuals in a setting similar to that of Ellison and Fudenberg (1993). In their model individuals observe distinct subsets of signals and do not observe the dates at which previous actions were taken. It is therefore striking that in equilibrium different individuals still converge upon the same behaviors (whether efficient or otherwise). The main focus of our paper is not on cascades/herding per se, but on possible convergence to inefficient outcomes.

[7]Payoff consequences of past actions are objective and measurable, whereas reasons for past actions (i.e., private information signals) are internal and subjective. Subjective information may be less credible because it is hard to verify directly. Furthermore, it may be less salient than payoffs, and is often hard to describe succinctly.

We assume that there is an exogenous sequence of individuals, each of whom decides which among a set of projects to adopt. Each individual learns (either through direct observation, or through conversation) the decisions of all those ahead of him. Each individual privately observes an identically distributed and conditionally independent signal about *expected* profitability of different action choices (projects). The true expected gain from each project is the same for all individuals. The realized value of a given project is subject to transient idiosyncratic shocks, i.e., two different individuals may gain different amounts from the same project.

In addition, depending on the scenario, the individual may observe a signal about the payoff outcomes resulting from the action choices of previous individuals. Since our setting allows for greater observation and/or communication than in the original cascades/herding models, we need to generalize slightly the definition of a cascade.

Definition: *An* **informational cascade** *occurs if, based on his observation and communication with others, an individual's action does not depend on his private signal.*

This definition generalizes the notion of an informational cascade to apply in a setting where individuals can observe more than just past actions, such as the payoff outcomes resulting from past action choices. In the special case of our setting where there is no acquisition of a signal about past payoff outcomes, Bikhchandani, Hirshleifer and Welch (1992) showed that if the number of individuals is sufficiently large, the probability that an informational cascade/herding eventually starts approaches one. (See also the related results of Banerjee (1992) and Welch (1992).)

Definition: *An observation/communication structure is* **long-run inefficient** *if in the long run the choices made are on average inferior to those made when all information is aggregated optimally.*

Implicit in the phrase 'on average' is both a probabilistic expectation and averaging over many individuals. In a scenario in which conditionally independent signals about project payoff are aggregated perfectly, in the long run everyone makes the correct decision, so if the structure is long-run efficient then almost surely almost everyone chooses the better action.

If cascades/herding lead to a positive probability that wrong decisions are made in the long run, then decisions are long-run inefficient, or in the terminology of Bikhchandani, Hirshleifer and Welch (1992), idiosyncratic. They also show that cascades/herding are fragile with respect to public shocks in the sense that an exogenous additional public disclosure with modest precision will with substantial probability shift the behavior of the next individual.

An important feature of our model is that individuals learn only about the payoffs of alternatives that are actually adopted. In addition, we allow payoffs to be subject to transient shocks and/or observation of payoffs to be noisy.[8] Owing to action dependence, we will see that cascades/herding can form in a regime with perfect observation/communication of past payoffs, even if these payoffs are perfect indicators of the future values of action alternatives.

The basic idea is simple. Suppose that there are two projects Y and Z. The payoff of project Y is either $v_Y = 0$ or 2, with probabilities μ and $1 - \mu$; the payoff of project Z is known to all, 1. Successive individuals receive a series of direct private signals about the value of project Y; conditional on the value of the project these value signals are i.i.d.. Each signal has two possible realizations, H and L, with $Pr(H|v_Y = 2) = p, Pr(H|v_Y = 0) = 1 - p$, $p > 1/2$. Let $v_Y(\emptyset)$ be the prior mean payoff on Y, and let $v_Y(S)$ denote the posterior mean payoff on Y when a sequence of signals S is observed. We assume that μ is such that $v_Y(\emptyset) < 1 < v_Y(H)$ where $v_Y(H)$ is posterior mean after receiving one H signal.

If the first individual obtains an L signal, which suggests that project Y is inferior to project Z, his action, Z, reveals his signal. Clearly, if the second individual receives an L signal, he will also choose project Z. Individual 2 will also choose project Z if he receives an H signal, because the posterior mean payoff of project Y after observing an LH sequence, $v_Y(LH)$, is the same as the prior mean, $v_Y(\emptyset) < 1$, the payoff of project Z. So if Individual 1 chooses Z, a cascade on project Z forms; the next individual chooses Z no matter what signal he receives. Learning about past payoffs does nothing to break this cascade, because observation of the non-stochastic payoff on Z tells successors nothing about the profitability of project Y. Thus, in the absence of interfering shocks, inefficient cascades/herding on project Z persist forever. If, on the other hand, the first individual selects project Y, then all uncertainty about Y is resolved and the ensuing cascade is always efficient.

If the first individual invests in project Z, the resulting cascade is idiosyncratic. The true payoff of project Y could be higher than Z, whereas if information of many individuals were aggregated optimally, they would converge upon the correct decision

[8] Actual conversation is noisy for several reasons, including discreteness in conceptual categories, difficulty in putting ideas into words and in interpreting words, and conflicts of interest between the communicators.

with probability 1.

Furthermore, the inefficient cascades/herding outcome is fragile with respect to exogenous shocks. If we introduce to the model a release of a public signal of modest precision at a given date, this can break the cascade and affect the behavior of many later individuals. Consider, for example, a single public signal with the same precision as the private signals that individuals receive. The arrival of a public signal realization of H on Y can shatter a cascade on Z. If the next individual also observes a private H signal, he chooses Y. If $v_Y = 2$, then thereafter all individuals choose Y instead of Z.

In this simple example there can only be an inefficient cascade upon one of the projects, Z, and once an inefficient cascade is broken the correct choice is always made thereafter. In the model of Section 3, there is a positive probability of an inefficient cascade on either project. Furthermore, there exist sample paths on which both projects are tried any number of times, yet inefficient cascades occur and have a positive probability of lasting forever. We will also see in Section 5 that even the communication of past private signals at low cost does not eliminate inefficient cascades.

3 A Model with Transient Shocks to Payoffs and Noisy Observation/Communication of Payoffs

In the example in Section 2, all individuals who adopt the same project receive the same payoff. We now offer a general model that allows for transient or individual-specific payoff shocks, i.e., payoffs that are stochastic conditional on the project states. The project states are unknown to individuals but are constant over time. We will show that under mild distributional assumptions, even though payoffs are observable, cascades always arise, with positive probability are incorrect, and with positive probability last forever. Thus information aggregation is inefficient relative to a situation in which private signals are aggregated perfectly. For notational simplicity, we first model the case in which observation or communication of past payoffs is perfect. We then will show that this analysis generalizes immediately to the case in which signals about past payoffs are noisy.

3.1 Transient or Individual-Specific Payoff Shocks

3.1.1 The Economic Setting

Let there be a sequence of individuals $i = 1, 2, 3, \ldots$. Each must choose between two project alternatives. The state of project $n = Y, Z$ is u_n or d_n (up or down). The underlying state is unobservable to individuals and does not change over time. The project has two possible payoffs, $v_{nl} < v_{nh}$. The project payoff $\tilde{v}_n^i$ is stochastic conditional on the project's state, and its distribution depends on the state but not on the individual. Let μ_{s_n} denote the prior probability that the state of project n is s_n. An i superscript denotes the payoff of the project adopted by Individual i. The prior probability that the payoff of project n when adopted by Individual i, $\tilde{v}_n^i$ takes on value v_{nk}, $k = l, h$, is denoted $\mu(v_{nk}|s_n)$. The expected payoff of project n conditional on state s_n is $\bar{v}_n(s_n) = \mu(v_{nh}|s_n)v_{nh} + \mu(v_{nl}|s_n)v_{nl}$. The payoffs $\tilde{v}_n$ of different projects and the project states s_n are uncorrelated across projects.[9] Throughout the paper, we assume that the likelihood ratio of payoffs from the two projects are not identical and that the likelihood ratio of signals from the two projects are not identical.

We apply the concept of perfect Bayesian equilibrium.[10] Individuals observe predecessors' actions and learn the payoff experiences for all previously-selected projects. Individual i observes a vector of conditionally independent and identically distributed direct private signals $\tilde{\sigma}^i = (\tilde{\sigma}_Y^i, \tilde{\sigma}_Z^i)$, where $\tilde{\sigma}_n^i$ is a signal about the state of project n, with possible values $\sigma_{nL} < \sigma_{nH}$. These signals are independent across projects, and (conditional on state) are independent across individuals. Let $p_q(s_n)$ be the probability that an individual observes signal realization σ_{nq} given that the true state of project n is s_n. We assume for notational convenience that $p_q(s_n) > 0$ for all n, q and s.

Let a^i be Individual i's action and let $A^i = (a^1, ..., a^i)$ represent the history of actions taken by Individuals $1, 2, ..., i$. Let b^i be the *payoff* from Individual i's action and let $B^i = (b^1, ..., b^i)$ represent the history of payoffs from the action of Individuals $1, 2, ..., i$. Given history (A^{i-1}, B^{i-1}), let $J^i(A^{i-1}, B^{i-1}, a^i)$ be the set of signal realizations that lead Individual i to choose action a^i given the information publicly available to him. His decision a^i communicates to others that he observed a signal in the set J^i. Let Σ denote

[9] It is straightforward to extend the model, including Proposition 1, its corollary, and Proposition 2 to the case in which the number of projects, the number of possible values from each project and the number of signal values are more than two; we have verified this explicitly (proofs available upon request). However, for notational simplicity, we present here the binary case.

[10] Since an individual's payoffs do not depend on what later individuals do, there is no incentive to make an out-of-equilibrium move to try to influence a later player. We therefore do not need to analyze what later players would do along off-equilibrium paths.

the set of possible signal realizations. Then if $J^i = \Sigma$, the individual is in a cascade, and Individual $i's$ action conveys no information about his signal realization.

Individual $i+1$'s conditional expected payoff from adopting project n given his own signal realization $\sigma^{i+1} = \sigma = (\sigma_Y, \sigma_Z)$ and the history A^i, B^i is

$$V_n^{i+1}(A^i, B^i, \sigma) \equiv E[V|\sigma^{i+1} = \sigma, A^i, B^i].$$

Individual $i+1$ adopts project n if the conditional expectation of project n is larger than that of the other project. When there is a tie, we follow the convention that Individual $i+1$ chooses the project favored by his private signal, and if there is still a tie, Individual $i+1$ chooses project Y. (The tie-breaking convention is not essential for the results; similar results would obtain under randomization among tied projects.)

If all information is aggregated efficiently, then by the law of large numbers, the correct decision will be made by everyone except a finite number of early individuals. Thus, we will use the term 'long-run efficient' to refer to a situation where with probability one, all but a finite number of individuals make the correct choice.

We impose a No-Long-Run-Tie assumption that if individuals learn enough about value by observing or conversing with predecessors, then they are not indifferent between the two project alternatives.

Assumption 1: No long-run ties If $n \neq n'$, then $\bar{v}_n(s_n) \neq \bar{v}_{n'}(s'_{n'})$, where $\bar{v}_n(s_n)$ denotes the expected payoff of project n given that the state of project n is s_n.

3.1.2 Implications

In this setting inefficient cascades/herds occur.

Proposition 1 *If Assumption 1 holds, and individuals can observe both actions and payoffs, then as the number of individuals increases, (i) the probability that cascades/herds eventually start approaches one; (ii) inefficient cascades/herds occur with positive probability; (iii) there is a positive probability that inefficient cascades/herds last forever; and (iv) if payoffs are subject to nondegenerate transient shocks, the probability that cascades/herds last forever can be less than one.*

Intuitively, for Part (i), if individuals try two different projects and thereby observe payoffs many times, by the law of large numbers the public information pool eventually becomes so informative about the desirability of the choice alternatives that some individual will start to 'ignore' his private signals about project states (i.e., his decision will be independent of his private signals). This starts a cascade.

For Part (ii), an inefficient cascade is possible because at the point at which an individual optimally conforms with the evidence about the actions and payoffs of others despite a private opposing private signal, he is still uncertain about which alternative is better. For example, early payoffs may indicate a high expected value for project Y relative to project Z even if 2 is better.

Part (iii) indicates that with positive probability the cascade continues forever. For example, if individuals are cascading on project Y, they do not generate new information about the true mean value of project Z (given the actual state). This information blockage can prevent people from ever discovering the superiority of project Z. If the state on project Y is favorable, individuals will tend to continue receiving good news about project Y, in ignorance of the even higher true mean payoff of the alternative project. There is a positive probability that a sequence of low payoffs on project Y will trigger a temporary or permanent switch to project Z. However, at any point in time there is also a positive probability that no further switch ever occurs.

Part (iv) indicates that that the cascade may be broken endogenously, because project payoff provides a noisy indication of the underlying state. With a long enough sequence of low payoff outcomes, a cascade, whether correct or not, will be broken.

The following example illustrates how inefficient cascades/herding can occur on either project, and how there is a probability between zero and one of an inefficient cascade on either project lasting forever.

Numerical Example of Inefficient Cascades with Heterogeneous Payoffs

Individual 1 observes private signals about Y and Z. He therefore has 4 possible information outcomes based on high versus low signals about each project. Individual 2 learns the payoff received from the project that Individual 1 adopted, as well as his own private signals about Y and about Z. Therefore, conditional on Individual 1's project choice, Individual 2 has 2^3 possible observation/communication outcomes.

Projects Y and Z each have possible payoffs of 0 or 1. Each project can be in two states, which are equally likely. Conditional on the state, the payoff distribution is: $Pr(v_Y = 1|u_Y) = 0.8$, $Pr(v_Y = 1|d_Y) = 0.2, Pr(v_Z = 1|u_Z) = 0.7, Pr(v_Z = 1|d_Z) = 0.3$. If the state is known, the conditional expectation of payoff from the projects are $\bar{v}_Y(u_Y) = 0.8, \bar{v}_U(d_Y) = .02, \bar{v}_Z(u_Z) = 0.7, \bar{v}_Z(d_Z) = 0.3$. Thus the no tie condition is satisfied. Project Y's ex ante expected payoff is more sensitive to state than Z's. Prior to his project choice, each individual receives signals regarding the state of the two projects. $Pr(\sigma_{YH}|u_Y) = 0.8$, $Pr(\sigma_{YH}|d_Y) = 0.2, Pr(\sigma_{ZH}|u_Z) = 0.6, Pr(\sigma_{ZH}|d_Z) = 0.4$.

Since the signal on project Y is more precise than that of project Z, he chooses Y if and only if his Y signal is high.

Now suppose that the true states are u_Y and u_Z, so that Y is superior. If the first individual receives a low signal about Y, he chooses Z. Suppose he obtains a payoff of 1 from Z. The second individual infers from the first individual's action that the first signal on project Y was low. So even if the second individual's private signal on project Y is high, his conditional expectation of Y is 0.5. Thus, the second individual chooses project Z even if he receives a low private signal on project Z; the favorable payoff of the first individual is such a favorable indicator (precision .7) that it outweighs his own private signal (precision .6). Therefore the second individual is in an inefficient cascade— his action is incorrect (relative to the true state), and it is independent of his private information signal.

There is a positive probability that the cascade will be overturned as later individuals learn about prior payoffs from the adopted projects. If a sufficiently long sequence of 0 payoffs from Z occurs, the cascade/herd will be overturned since posterior expectation of project value will approach 0.3. Nevertheless, since the true mean of Z is 0.7, the probability that the posterior mean of Z will stay above 0.5 forever is positive.[11] With positive probability, the inefficient cascade lasts forever. $\|$

Fixation upon one project has been analyzed in models without private information. In the basic single-person multi-arm bandit model of Rothschild (1974), it can be socially optimal to desist from experimenting between choice alternatives even though substantial uncertainty remains about which is more desirable. In the multi-person model of Bolton and Harris (1999), the amount of experimentation performed is suboptimal, because of the positive informational externality that such experimentation confers upon later observers.

In contrast with multi-arm bandit models, here individuals do not just view past payoffs, they also spontaneously receive conditionally independent private signals about value regardless of what actions were taken. If these signals were reasonably well aggregated, society would converge to accurate decisions. However, we find that even though both information sources are present (private signals and information about payoffs), there is still a failure to converge to efficiency.[12]

[11] Beliefs evolve according to the net number of 1 versus 0 profit signals. This number follow a random walk with a positive drift, which need never cross a fixed lower bound— see, e.g., Chung (1974), p. 263.

[12] Also in contrast with the multi-arm bandit literature, in our model each individuals makes only a single decision, so individuals do not design strategies for the purpose of generating new payoff

3.2 No Transient Shocks

We have shown that inefficient cascades/herding can occur and can last forever when individuals can observe actions and payoffs, and that these cascades/herds have a positive probability of being broken. The possibility of cascades/herds being dislodged spontaneously contrasts with the result of Banerjee (1992) and Bikhchandani, Hirshleifer and Welch (1992). In their settings, if all individuals are identical and there are no interfering shocks, a cascade once started lasts forever. A similar outcome applies in our model only in the special case in which payoffs depend on the state nonstochastically. In this special case, once a cascade starts, observing further identical payoffs does not reveal any new information. With no further information arriving, all succeeding individuals join the cascade. This proves the following corollary.

Corollary *If payoffs are non-stochastic functions of state and individuals observe all past actions and project payoffs, then cascades/herds once started last forever.*

When payoffs are non-stochastic functions of state, inefficient cascades/herding are extremely fragile in the sense that a small shock (such as a new public information arrival) can easily shifts the long run outcome. If the shock persuades just one individual to try the alternative project, then everyone thereafter shifts to the better project. An individual can be sure that he is in a correct cascade only if both projects have been tried previously or if the payoff of current project is higher than the highest possible value from the untaken projects.

3.3 Noisy Observation/Communication of Payoffs

So far we have assumed that payoffs are perfectly observable. Proposition 1 and the Corollary generalize immediately to the case in which individuals observe noisy signals about past project payoffs. We describe the model informally here, and provide formal details and proof in the appendix.

Suppose that each individual obtains (through conversation or direct observation) a set of noisy signals about each of the payoffs of previously-chosen projects. Each signal is equal to the payoff plus independent noise. After obtaining information about the past payoff of a project, for two reasons an individual may still not know the true mean payoff

information. Furthermore, in our analysis individuals possess private signals which are aggregated in a socially suboptimal fashion. Thus, in our setting there are potential inefficiencies both in generating and in aggregating information signals. Our analysis therefore brings together aspects of both the social learning literatures (private information) and the multi-arm bandit literatures (experimentation).

of the project (given the true state). First, the realized payoff can be a noisy indicator of the true mean. Second, the signal obtained by the individual may be a noisy indicator of the realized past payoff. As in the setting with noiseless observation/communication of payoffs, a given state and investment choice implies a probability distribution for possible signal realizations. For any two-stage garbling of the state resulting from stochastic payoffs and signal noise in payoffs, there is a corresponding setting with no signal noise but where the probability distribution of payoffs creates an equivalent garbling. In such a setting individuals will therefore behave in a precisely identical fashion. This intuition is developed formally in the appendix. We therefore have:

Proposition 2 *Suppose that Assumption 1 holds, individuals can observe past actions, decision payoffs are stochastically heterogeneous and that the observation or communication of payoffs is noisy. Suppose further that the distribution of the payoff signals is not identical across states. Then as the number of individuals increases: (i) the probability that a cascade eventually starts approaches one; (ii) inefficient cascades occur with positive probability; (iii) there is a positive probability less than one that cascades/herding last forever; and (iv) the probability that cascades/herding last forever can be less than one.*

As pointed out by Banerjee and Fudenberg (1999), observation of a small subsample of past actions causes subsets of observers to have different beliefs. If some individuals by chance observe a subset of actions that includes the less-popular action, they may be inclined to make use of signals about payoffs, improving information aggregation. In a sequential decision setting, an individual who observes a subset of past actions and/or payoffs that is relatively adverse to the currently-most-popular action may break a cascade/herd. Again, this would cause further information to be aggregated. This raises the question of whether information blockages and cascades can persist in the long run when there is random sampling of previous decision makers. We do not attempt to model this question here.[13]

[13] It is clear that if the probability of each predecessor being in the subsample is close to 1, then cascades/herding will still last a long time. More generally, in a setting where individuals only take subsamples among the most recent K predecessors, it is plausible that the system may reach a point where all recent individuals choose the same action, suggesting that cascades may persist forever. (This assumption differs from the settting of Banerjee/Fudenberg (1999), in which individuals do not have information about the timing of past decisions and in which sampling is not biased with respect to timing.) We have verified in a special example in which there is partial sampling that inefficient cascades still occur and with high probability last forever.

4 Observation/Communication of Payoffs Can Reduce Average Welfare and Decision Accuracy: An Example

Long-run efficient decision making is a stringent benchmark. It requires that almost everyone do the right thing with near certainty. One might hope that even if a structure in which payoffs are observable is not long-run efficient, that observing payoffs at least increases ex ante average welfare relative to not observing payoffs. We now show that this is not always the case.

Consider two projects Y and Z. The payoff of Y is 0 and $2+\epsilon$ with equal probability, and the payoff of Z is $1+\epsilon, 1-\epsilon$ with equal probability, where $\epsilon > 0$ is a sufficiently small quantity. Each individual observes a signal about Y with precision p and no signal about Z.[14] We compare the ex ante expected welfare outcomes of two scenarios. In the first, individuals do not observe any payoffs from the adopted projects. In the second scenario, individuals can observe payoffs from the Z project if it is adopted but individuals cannot observe payoffs from project Y.[15] Since there is no observation noise, once the payoff of Z is observed it is known precisely thereafter.

Assumption A: ϵ is sufficiently close to zero,

$$\epsilon < \frac{2p-1}{2-p}.$$

Assumption A ensures that if all individuals had to make decisions independently, they would adopt project Y without any signals or with an H signal, and adopt Z with an L signal. Define the social welfare as discounted expected utility of all individuals. Let δ be the discount factor over time, W_i denote the payoff of individual i, and let the average realized payoff be defined as $W(\delta) \equiv (1-\delta)\sum_{i=1}^{\infty} \delta^i W_i$.[16]

Result 1 *In the numerical example, allowing individuals to observe payoffs can on average make individuals worse off and decisions less accurate. Outcomes with observable payoffs are long-run inefficient.*

It seems reasonable that allowing an additional means of information transmission, conversation, would improve information flow and lead to improved decisions. Result

[14]We can allow individuals to observe a very noisy signal of Z without affecting any of the results.

[15]Similarly, we can allow individuals to observe a very noisy signal of the predecessor's realized payoff from adopting project Y without affecting any of our results here.

[16]It is easy to show that $\lim_{\delta\to 1} E[W(\delta)] = \lim_{i\to\infty} E[W_i]$; see the appendix.

1 shows that observing payoffs can potentially lock individuals in an inefficient cascade/herd. While the observation or communication of payoffs provides an individual with valuable information, it can reduce the informational content of the earlier individual's action. Whether later individuals are better off or worse off with learning about payoffs thus depends on the tradeoff between the information conveyed directly by payoffs and the possible reduction of information resulting from the alteration in previous actions.

Previous papers have pointed out that greater information flow can reduce ex ante welfare. In Hirshleifer (1971) disclosure can reduce welfare by reducing risk-sharing. Teoh (1997) finds that disclosure can reduce ex ante welfare by inhibiting contributions in a public good game. In settings with informational cascades/herding, improved information flows do not necessarily improve overall decisions. For example, Bikhchandani, Hirshleifer and Welch (1992) find that the prospect of a public information disclosure can reduce average welfare. A difference between our model and previous papers is that the amount of additional information generation is endogenous. The ability to observe past payoffs could potentially resolve all uncertainty about which action is superior and thereby bring about the correct action choice in the long run, if only there were sufficient experimentation.

The reason that observation or communication of past payoffs can be harmful is that this can cause individuals to settle into cascades/herding even earlier than they would have otherwise. This advance can impair information aggregation. This suggests that if observation/communication of payoffs is delayed, later individuals may sometimes obtain the benefit of observing payoffs without the cost of reduced informativeness of predecessors' actions and payoffs, making the average individual better off. We now modify the example by assuming that the payoffs of project Z can be observed with one period delay.

Result 2: *In the numerical example, delaying the observation/communication of payoffs can on average improve the accuracy of almost all individuals' decision and average welfare.*

According to Alexander Pope[17]

> A little learning is a dangerous thing;
> Drink deep or taste not the Pierian spring.

[17] *Essay on Criticism*, 1711, Part ii, Line 15, from John Bartlett, *Familiar Quotations*, 10th ed., 1919.

There shallow draughts intoxicate the brain,
And drinking largely sobers us again.

Taking observation of past actions as a starting point, 'drinking largely' in our setting would be observation of all past private signals, which leads to correct decisions in the long run. A shallow draught of learning, observing past payoffs (but not past signals), can be harmful. Observing payoffs only with delay is an even shallower draught of learning. In our setting this shallower draught can be the better one.

The rise of mass media and of interactive communication technology (printing, telegraph, telephone, television, email, the web) have made it easier to observe the payoff outcomes of others (or summary statistics of past payoffs).[18] The analysis here suggests that the resulting improvements in decisions may not be as great as might have been expected. The recipient of extra information (beyond what he would have obtained by observing past actions) gains directly from the ability to make a more accurate decision. However, this benefit may be opposed by the fact that the actions of the recipient may be less informative to later decisionmakers.

5 Communication of Private Information

Previous literature on cascades/herding focuses on learning by observation of either past actions or the payoffs resulting from past action choices. However, in conversation individuals can convey pre-decision information about alternative choices. We consider cases in which individuals observe past actions, but information about past signals is conveyed incompletely.

In Subsection 5.1 we consider a case in which the immediate predecessor's private signal is communicated. In Subsection 5.1 we consider the case where, for a cost, each individual can observe all information possessed by his immediate predecessor (not just the predecessor's private signal, but any information the predecessor has obtained by observing or conversing with his predecessors).

5.1 Communication of Immediate Predecessor's Private Signal

Suppose that Individual i costlessly observes all previous individuals' actions, and can costlessly learn by means of conversation only his immediate predecessor's private signal.

[18]These advances have also made it easier to observe past actions and to communicate private information. Our discussion here does not try to address the full scope of these changes.

The realized payoffs can be 2 or -2 for both projects Y and Z. There are two state of project Y, u_Y, d_Y representing up and down states with equal prior probabilities. Project Y provides an conditional expected payoff of 1 in state u_Y, and 0 in state d_Y. There is only one state for project Z and its expected payoff is 0.5. Private signals about states are conditionally independent, binary and symmetric, with values of either H or L. Thus, $Pr(H|u_Y) = Pr(L|d_Y) = p > 0.5$ and $Pr(H|d_Y) = Pr(L|u_Y) = 1 - p$.

We first present an example in which, for simplicity, payoff information is not observed or communicated. (This is the special case of the noise assumption in Section 3 in which the noise in observing payoffs is infinite.) We interpret this as a situation in which project outcomes require a considerable period of time to resolve, so that later individuals need to make their decisions before earlier individuals have learned the outcomes of past project choices. We then establish similar conclusions when payoffs are observable in Proposition 3.

We assume that an individual always follows his own signal if he is indifferent. Thus each of the first two individuals always acts in accordance with his own signal. We describe the choices of individuals for all possible cases. In Case I, if the first two individuals receive opposing signals, then the third individual will be in a situation informationally equivalent to his situation had he been first in the sequence.

Next consider Case II in which Individuals 1 and 2 receive the same signal realizations. Since the signals and payoffs are symmetric, suppose that both individuals receive H signals. In this case Individual 3 is in a cascade on project Y, as his signal is dominated by the information revealed through the first two individuals' actions. In Case IIa, if Individual 4 directly observes L_3L_4 (i.e., his own signal is L and through conversation he has learned about an L signal of his predecessor), Individual 4 breaks the cascade. When Individual 4 breaks the cascade, his action reveals that both his signal and the third signal were L, so Individual 5 is in a situation informationally equivalent to one where he is first in the sequence.

Finally consider Case IIb in which Individual 4 also follows the same action as his three predecessors. This implies that the third and fourth signals were not L_3L_4. In consequence, we will see that even when Individual 5 directly observes L_4L_5 (his own signal and that of Individual 4) he still follows the same action: he is in a cascade. In this case, Individual 5 can infer from the fact that the fourth signal was L that the third signal must have been H— otherwise Individual 4 would have switched actions. So Individual 5 can infer that there were in total three H signals and two L signals. The likelihood ratio is

$$\frac{Pr(d_Y|HHHLL)}{Pr(u_Y|HHHLL)} = \frac{Pr(HHHLL|d_Y)Pr(d_Y)}{Pr(HHHLL|u_Y)Pr(u_Y)} = \frac{(1-p)^3p^2}{p^3(1-p)^2} = \frac{1-p}{p} < 1,$$

so Individual 5 joins the cascade regardless of his signal.

We further show that in Case IIb, the cascade cannot be broken. Individual 6 is also in a cascade, because even if he observes L_5L_6, this merely offsets the first two H signals. From his perspective, the signals of Individuals 3 and 4 could have had at most one L, so either one was H and one was L, or else both signals were H. Thus state u is more likely than state d. Finally, Individual i will follow the cascade for all $i \geq 7$ because even when he directly observes LL, he will be in a situation exactly like that of an Individual 6 who observed L_5L_6.

To summarize, either Individual 3 or Individual 5 will be in a situation informationally equivalent to their being first movers, or else there will be a cascade starting with Individual 5. Thus after the actions of six individuals, with positive probability a cascade has formed, and if not then the beliefs of either the third or the fifth individuals are formed based on their own signals as if there were no prior history.

Thus, our analysis indicates that for any Individual $i = 2j+1$, there are three possible scenarios. In the first, his informational status is equivalent to that of a first mover, as in the case of Individual 3 in Case I or Individual 5 in Case IIa. In the second scenario, the information of predecessor's actions and his immediate predecessor's private signal is equivalent to two signals of H or two signals of L, as in the case of Individual 3 in Case IIa. Finally, in the last scenario, an individual is in a cascade that cannot be broken, i.e., even when Individual $2j+1$ and Individual $2j+2$ both observe signals opposite of Individual $2i$'s action, Individual $2j+1$ will still follow Individual $2j$'s action. In the last scenario, Individual $2j+1$ is like Individual 5 in Case IIb.

Next, we derive the probability of each individual to being in a cascade. An individual is in a cascade if his decision is independent of his signal. Thus Individual $i = 2j+1$ will be in a cascade only under scenarios 2 and 3 and Individual $2j+2$ will be in a cascade only if Individual $2j+1$ is in scenario 3. Let π_i denote the ex ante probability that Individual i is not in a cascade. The probability that Individual $2j+1$ is not in a cascade is the probability that Individual $2j+1$ is in scenario 1. This means that Individual $2j+1$ is like Individual 3 in Case I or Individual 5 in Case IIa. When Individual $2j+1$ is like Individual 3 in Case I, Individual $2j-1$ must be in the first scenario and the next two signals must be opposite of each other which occurs with probability $\pi_{2j-1}2p(1-p)$.

When Individual $2j+1$ is like Individual 5 in Case IIa, this implies that Individual

$2j-3$ must be informationally equivalent to Individual 1. In addition, the four signals preceding Individual $2i+1$ must be either $HHLL$ or $LLHH$. The probability that Individual $2j+1$ is like Individual 5 in Case IIa occurs with probability $\pi_{2j-3}2p^2(1-p)^2$. Thus the probability of Individual $2j+1$ not being in a cascade is $\pi_{2j-1}2p(1-p)+\pi_{2j-3}2p^2(1-p)^2$.

There are two exclusive cases in which Individual $i=2j$ is not in a cascade. In the first case, Individual $2j-1$ is in scenario 1. In the second case, Individual $2j-3$ is in scenario 1 and the three signals preceding Individual $2j$ are either HHL or LLH, which occurs with probability $\pi_{2j-3}[p^2(1-p)+(1-p)^2p]$. So for Individuals $2j+1$ or $2j$, we have n,

$$\begin{aligned} \pi_{2j+1} &= \pi_{2j-1}2p(1-p)+\pi_{2j-3}2p^2(1-p)^2 \\ \pi_{2j} &= \pi_{2j-1}+\pi_{2j-3}[p^2(1-p)+(1-p)^2p], \end{aligned} \tag{1}$$

with the initial conditions

$$\pi_1=1, \pi_2=1, \pi_3=2p(1-p).$$

Solving the iterative equations, for $i \geq 2$ we obtain

$$\begin{aligned} \pi_i &= \alpha_1(\beta_1)^j+\alpha_2(\beta_2)^j, \quad i=2j+1 \\ \pi_i &= \pi_{2j-1}+\pi_{2j-3}p(1-p), \quad i=2j, \end{aligned} \tag{2}$$

where

$$\begin{aligned} \alpha_1 &= \frac{3-\sqrt{3}}{2} \\ \alpha_2 &= 1-\alpha_1 \\ \beta_1 &= (1+\sqrt{3})p(1-p) \\ \beta_2 &= (1-\sqrt{3})p(1-p). \end{aligned} \tag{3}$$

Taking the limit as $i \to \infty$, we see that the probability of a cascade forming approaches 1 for later individuals. Furthermore, informational cascades aggregate information inefficiently. This follows immediately from the fact that once a cascade is formed, the public pool of information stops accumulating. This result is a special case of the following proposition:

Proposition 3 *If each individual can observe all past actions, possibly past payoffs, and through conversation learns the private signal only of his immediate predecessor, then informational cascades/herds form with probability one and are long-run inefficient.*

Thus, the problems of imperfect information aggregation associated with observational learning remain when individuals converse about past signals, so long as this discussion is sufficiently limited.

Of course, more generally conversation may not be as limited as this simple model. For example, an individual who learns his predecessor's private signal may be able to communicate two signals to his successor– his own signal and his predecessor's. Or, he may be able to combine these two signals and communicate a summary statistic. In either case, information aggregation may be improved. However, there are also other respects in which communication may in practice be more limited than assumed in our analysis. For example, there can be occasional breaks in the conversational chain, so that even though Individual 3 observes Individual 2's action, Individual 2's private information is not conveyed conversationally at all. Furthermore, since conversation takes time and energy, the information conveyed by Individual 2 to Individual 3 in conversation may be noisier than the information Individual 2 uses in his own decisions. Individual 2 may forget Individual 1's signal once Individual 2 has made his decision, in which case when asked Individual 2 may convey only his original signal. Individual 2 may also forget his own private signal after he makes his decision. Our simple assumptions about the limits to conversation illustrate how such limits can cause problems of information aggregation. The next subsection considers the possibility that communication is perfect is the individual obtaining information is willing to incur a fixed cost.

5.2 Costly Communication of Immediate Predecessors' Entire Information Set

Suppose that instead of conveying only an individual's immediate predecessor's private signal, conversation conveys the predecessor's entire information set. This includes any knowledge derived from the next predecessor in turn. We assume that there is a fixed cost of conversing which is incurred entirely by the party who is acquiring information. In other words, if an individual initiates a conversation with his predecessor, the predecessor responds by freely providing all his information.

Thus, there is potentially a conversational chain that aggregates all information from the first individual onward. Intuitively, if the communication cost is low, it is plausible that there would be highly accurate decisions in the long run. Communication is fairly unconstrained, so at low cost a fairly long communicative chain that aggregates many private signals becomes feasible. Consistent with this intuition, if conversation is completely costless, there exists an equilibrium in which each individual observes the previous individual's information, no information is lost, and decisions are efficient.

However, if obtaining information through conversation is costly, we will show that information is lost permanently and individuals are locked in a cascade with probability one. Furthermore, we provide an example in which even in the limit as the cost approaches zero, *substantially* inferior decisions are made. Thus, the outcome is socially inefficient.

Intuitively, in addition to the direct-action-based informational externality of the basic cascades/herding model, there is a conversational externality. An individual who pays to communicate with his predecessors does not reap the full benefit that his improved information and decisions confer upon later decisionmakers. If the individual benefit to conversation is small, even a small cost can have a large social consequence.

An Example

We first illustrate these points in an example similar to that of the previous subsection. The two projects are the same as defined earlier. However, here we assume that each individual can observe his immediate predecessor's information set (not just the predecessor's private signal) at a cost ϵ. The cost ϵ can be arbitrarily small.

Without loss of generality, suppose that the first signal is H, so that Individual 1 accepts project Y and his action perfectly reveals his signal. With this revelation, there is no gain to Individual 2 from conversing to learn his predecessor's information. If Individual 2's signal is H, he takes project Y. If his signal is L, he is indifferent between projects Y and Z, so he follows his own signal and takes project Z. Thus, the second individual's action is also perfectly informative. Thus, there is no benefit to Individual 3 from conversing to learn Individual 2's information. When the first two actions are different, they reveal that the first two signals are opposite, so that the informational status of Individual 3 is equivalent to that of a first mover. When the first two actions are the same, then the first two signals are the same and these dominate Individual 3's signal. Thus Individual 3 is in a cascade and will choose the same action. Individual 4 will also not pay to learn Individual 3's information because the same first two signals weakly dominate the next two signals, so he has nothing to gain from doing so. Similarly, each subsequent individual has no gain from observing his predecessor's information, and the cascade persists forever. This indicates that no matter how small the cost of conversation is, as long as the cost is positive, no one will ever observe his immediate predecessor's information. $\|$

This example illustrates the mismatch between the private value of communication and the social value. However, the binary example involves decisions by indifferent individuals, so there is a question of how robust is this effect. More generally, error-prone cascades must always form.

Proposition 4 *Assume that for a fixed cost each individual can converse with his immediate predecessor to learn that individual's information set. No matter how small the cost is, with probability one, eventually there will be an individual who will choose not to obtain his predecessors' information, and who will follow his predecessor's action in an informational cascade/herd. Thereafter all subsequent individuals choose not to obtain their predecessors' information and follow the same action. Moreover, cascades are long-run inefficient.*

Long-run efficiency was defined in Section 2 solely in terms of the accuracy of action choices, rather than as a balance between accuracy and conversation costs. If conversation costs are pure transfers to other individuals then long-run efficiency corresponds to approximate economic efficiency– approximate because it is a condition on average performance over many individuals. This might be the case, for example, if the recipients of conversational initiatives enjoy the process.

However, if conversation costs are pure dead weight losses, then a failure to achieve long-run efficiency as defined here does not imply even approximate social inefficiency; it is socially desirable to accept lower quality decisions in order to reduce the incurred costs. We now verify that even with very low conversation costs, decisions can be substantially inaccurate. In consequence, the social outcome is not even approximately socially efficient.

In the earlier binary example, no individual ever pays to learn his predecessor's information, regardless of the cost of observation. In the numerical example that follows, some individuals do learn through conversation, but a substantial loss of welfare still results even when the costs of communication are small. In this example, we assume that the distribution of signals is asymmetric. Project Z has a fixed payoff of 1/2. Project Y has two possible payoffs, 1 in state u and 0 in state d. The signal realizations H and L have probabilities $Pr(H|u) = p, Pr(H|d) = 0.5$. Thus, the likelihood ratios given the signals are

$$\frac{Pr(d|H)}{Pr(u|H)} = \frac{1}{2p}, \qquad \frac{Pr(d|L)}{Pr(u|L)} = \frac{1}{2-2p}.$$

Unlike the symmetric case, here one H signal does not exactly offset one L signal. The likelihood ratio between states u and d for a HL signal series is $2p(2-2p) < 1$, so one L signal dominates one H signal. The larger p is, the more H signals are needed to offset one L signal. Each individual can learn his immediate predecessor's entire information set for a cost of ϵ. As before, we assume that when an individual is indifferent between obtaining or not obtaining information, he will always choose to obtain information. Furthermore, as before we assume that an individual will always follow his own signal when he is indifferent between the two projects.[19]

Let l^i denote the public likelihood ratio between the bad state and the good state of project Y after Individual i's action. As earlier, let social welfare be defined as the discounted sum of the expected utilities of all individuals.

Let $W(\epsilon, \delta)$ denote the expected discounted payoffs averaged over individuals given cost ϵ and discount factor δ, i.e., $W(\epsilon, \delta) \equiv (1-\delta)\sum_{i=1}^{\infty} \delta^i E^{\epsilon}[W_i]$, where as before W_i is the payoff of individual i.

Result 3: *In the asymmetric numerical example, for arbitrarily small positive cost of conversation, the public likelihood ratio l^i of all individuals lies strictly in*

$$\left(\frac{3(1-p)^2}{p(2-p)}, \frac{3p^2}{1-p^2}\right),$$

which lies inside the interval between the likelihood ratio of three L signals and the inverse of the likelihood ratio of three L signals. In contrast, with zero observation cost, the likelihood ratio approaches 0 or infinity depending on the state of project Y. Let $W(\epsilon, 1) \equiv \lim_{\delta\to 1} W(\epsilon, \delta)$. As the discount factor δ approaches one, the welfare loss due to the conversation cost can be substantial relative to ϵ, i.e., $\lim_{\epsilon\to 0}[W(0,1) - W(\epsilon, 1)] > 0$.

Result 3 indicates that the ability of conversation to propagate information along the informational chain is delicate. Even with an arbitrarily small friction, such propagation eventually stops, at which point individuals are limited to observing something much less informative, the history of past actions.

Since the likelihood ratio is bounded, this means that there exists a number $\pi < 1$ such that the expected payoffs of all individuals is $(2\pi - 1)/4$. This is strictly less than $1/2$, the expected payoff under full information. As $\delta \to 1$, the social welfare under the rule in which all investors acquires information is in the neighborhood of $1/2 - \epsilon$.

[19] Similar results hold regardless of the tie breaking rules. For example, we can allow an individual to randomize or to follow his predecessor's project whenever he is indifferent.

The loss resulting from insufficient conversation can be $O(1)$ as $\epsilon \to 0$. In other words, even an arbitrarily small cost can cause a substantial reduction in the average quality of decisions. Thus, the social welfare (both gross and net of incurred conversation costs) actually attained can be lower than that under socially optimal decisions.

6 Conclusion

We have examined how efficiently society aggregates information when individuals make decisions in sequence, observe past actions, and observe or communicate either the payoffs resulting from past action choices, or their spontaneously-arriving private information signals. In previous cascades/herding models, society is trapped in inefficient choices until a shock arrives to dislodge the cascade. Here, individuals can *endogenously* talk their way out of inefficient cascades by learning about the past payoff outcomes, or the private signals of their predecessors. Thus, undisturbed cascades do not last forever with probability one.

If inefficient cascades are with probability one eventually dislodged, then in the long run individuals make accurate choices. However, our analysis indicates that it can be hard for groups of individuals to "talk their way out of trouble," because of the three types of informational externalities. The first two are externalities in action choice. Direct action-based informational externalities occur because, as in the basic cascades/herding model, each individual ignores the informational benefit to others of basing his action on his private signal. The indirect action-based externality occurs because each individual ignores that taking a little-chosen action provides very useful payoff-outcome information to others. This third type, conversational externalities, occurs because an individual ignores the benefit to later individuals of talking to earlier ones. When an individual converses with predecessors in order to come to a decision, he becomes more knowledgeable, which confers a benefit upon later decisionmakers who have may want to converse with him.

Nevertheless, we find that there is still a positive chance that an inefficient cascade does last forever. Even when project payoffs are communicated, there can be insufficient learning about the project that is less frequently undertaken. Society is harmed by a lack of experimentation on 'the road less traveled by.'

Furthermore, we provide examples which show that the ability to observe/communicate past payoffs can, ex ante, reduce expected average welfare and the expected accuracy of action choices. Similarly, delay in observation or communication of payoffs can increase ex ante expected welfare. Delay or non-communication can help because they can encourage individuals to aggregate more information before cascades clog the flow of information.

When individuals can, through conversation, acquire past spontaneously-arriving private signals, they can potentially learn a great deal about even the less-popular project choice. However, we find that owing to the conversational externality, even arbitrarily small costs of conversation cause substantially inefficient cascades in the long run. Furthermore, if the cost is zero but each individual can only acquire his immediate predecessor's private signal, inefficient cascades persist. Taken together, these findings indicate that the problem of information blockage through informational cascades/herding requires surprisingly little in the way of limits to observation or communication.

Our model of course does not capture all aspects of conversation. We comment upon only one further direction for research. In practice it seems that some gregarious individuals specialize in internalizing conversational externalities. These people are 'good conversationalists,' are knowledgeable about the past and current workings of their organization. Our approach suggests that such gregarious individuals may be very valuable to their friends and the organizations or societies to which they belong. That value may also be a source of power for individuals who maintain a rich network of conversational contacts.

Appendix

Proof of Proposition 1:

(i): Assertion (i) states that a cascade/herd will eventually form with probability one. We consider the case of two projects; the same conclusions hold for any number of projects, by similar reasoning. For clarity, we restate our assumption here. Let there be a sequence of individuals $i = 1, 2, 3, \ldots$. Each must choose between two project alternatives. The state of project $n = Y, Z$ is u_n or d_n (up or down). The underlying state is unobservable to individuals and does not change over time. The project has two possible payoffs, $v_{nh} < v_{nl}$. The project payoff $\tilde{v}_n^i$ is stochastic conditional on the project's state, and its distribution depends on the state but not on the individual. Let μ_{s_n} denote the prior probability that the state of project n is s_n. An i superscript is used to denote the payoff of the project if adopted by Individual i. The prior probability that $\tilde{v}_n^i = v_{nk}$ is denoted $\mu(v_{nk}|s_n)$, $k = h, l$. The expected payoff of project n conditional on state s_n is $\bar{v}_n(s_n) = \sum_k v_{nk}\mu(v_{nk}|s_n)$. The project payoffs $\tilde{v}_n$ of different projects and the project states s_n are uncorrelated across projects.

We now prove the first claim in Proposition 1, that with probability one a cascade starts. Without loss of generality, suppose that the true state is (u_Y, u_Z). Let l_n^i denote the public likelihood ratio for the two states of project n before Individual i's action, and let F_i denote the public information set available to individual i,

$$l_n^i = \frac{Pr(d_n|F^i)}{Pr(u_n|F^i)}.$$

Given this likelihood ratio, the probability that project n is in the up state is $1/(1 + l_n^i)$.

Using public information, individual i with likelihood ratios l_Y, l_Z will be indifferent between the two projects Y, Z if

$$f(l_Y, l_Z) \equiv \frac{\bar{v}_Y(u_Y)}{1 + l_Y} + \frac{l_Y \bar{v}_Y(d_Y)}{1 + l_Y} - \left(\frac{\bar{v}_Z(u_Z)}{1 + l_Z} + \frac{l_Z \bar{v}_Z(d_Z)}{1 + l_Z}\right) \tag{4}$$

is zero. The first two terms in f represent the conditional expected payoff from project Y based on public information and the next two terms in f represent the conditional expected payoff from project Z based on public information. The quantity $f(0, 0) \neq 0$ since $\bar{v}_Y(u_Y) > \bar{v}_Z(u_Z)$.

Since l_Y, l_Z is a martingale given state u_Y, u_Z, by the martingale convergence theorem (see Chung (1974)), there exists a random vector (l_Y^∞, l_Z^∞) such that (l_i^Y, l_i^Z) converge to (l_Y^∞, l_Z^∞) almost surely.

At least one project will be taken infinitely often. Let ω denote the a sample path such that (l_Y^i, l_Z^i) converges. This means that for any $\epsilon > 0$, there exists an i' such that $|l_Y^i - l_Y^\infty| < \epsilon$. Suppose that Y is taken infinitely often. Then $l_Y^\infty = 0$, because otherwise, l_Y^i will change by a finite amount to deviate from the ϵ neighborhood of l_Y^∞ whenever we observe an informative payoff from project Y. Similarly if Z is taken infinitely often, then $l_Z^\infty = 0$. Moreover, suppose that there are sample paths on which both projects are taken infinitely often and (l_Y^i, l_Z^i) converges. Then $(l_Y^i, l_Z^i) \to (0, 0)$. Since $f(0,0) > 0$, this implies that for all i sufficiently large, Individual i will take project Y. This is inconsistent with the assumption that both projects Y and Z will be taken infinitely often. Thus on sample paths that both projects are taken infinitely often, (l_Y^i, l_Z^i) will not converge and such sample paths have a measure of zero. Thus with probability one all individuals eventually cascade/herd on the same decision regardless of their private information.

(ii): Assertion (ii) states that an inefficient cascade can occur with positive probability. Without loss of generality, we assume project Y has the highest possible of the state-conditional expected payoffs, i.e., $\bar{v}_Y(u_Y) > \bar{v}_Z(u_Z) > \bar{v}_Y(d_Y)$.

With probability one, individuals will start to cascade on either project Y or Z. Suppose that individuals cascade on project Y given a series of signals and project payoffs. The same set of signals and payoffs has a positive probability of occuring in state (d_Y, u_Z). Similarly, suppose individuals cascade on project Z given a series of signals and project payoffs. The same combination of signals and payoffs has a positive probability of being observed in state (u_Y, u_Z). Consequently, an inefficient cascade/herd occurs with positive probability.

(iii): In this part, we prove that inefficient cascades/herds can start and last forever even when all payoffs resulting from adopted projects are publicly observable. From parts (i) and (ii), we know that with positive probability a cascade starts. Suppose the true state is (u_Y, u_Z), and a cascade has formed on project Z. We prove that such an inefficient cascade can last forever. At the beginning of the cascade with Individual i, let μ denote the highest expected value for Individual $i+1$ of project Y given the most favorable possible signal values,

$$\mu = E[v_Y | A^i, B^i, \sigma^{i+1} = \sigma^M],$$

where σ_M corresponds to the most favorable possible signals for project Y.

Now clearly $\mu_Z^i > \mu$, and

$$\mu_Z^i = \pi_Z^i(u_Z)\bar{v}_Z(u_Z) + \pi_Z^i(d_Z)\bar{v}_Z(d_Z). \tag{5}$$

where $\pi_Z^i(u_Z), \pi_Z^i(d_Z)$ are the conditional probabilities that project Z is in the high or low state. Let l_Z^i denote the likelihood ratio of the down state versus the up state of project Z. We can rewrite equation (5) as

$$\mu_Z^i = \left(\frac{1}{1+l_Z^i}\right)\bar{v}_Z(u_Z) + \left(\frac{l_Z^i}{1+l_Z^i}\right)\bar{v}_Z(d_Z) = \bar{v}_Z(d_Z) + \left(\frac{1}{1+l_Z^i}\right)\left[\bar{v}_Z(u_Z) - \bar{v}_Z(d_Z)\right]. \tag{6}$$

The expected payoff is a decreasing function of the likelihood ratio. Notice the likelihood ratio evolves according to the payoffs received from project Z. Suppose that the payoff from project Z is v_{Zk},

$$l_Z^{i+1} = l_Z^i\left(\frac{\mu(v_{Zk}|d_Z)}{\mu(v_{Zk}|u_Z)}\right),$$

which implies that

$$\log(l_Z^{i+1}) = \log(l_Z^i) + \log\left(\frac{\mu(v_{Zk}|d_Z)}{\mu(v_{Zk}|u_Z)}\right).$$

Letting x_Z^{i+1} denote the change in the log likelihood ratio,

$$\begin{aligned}
E[x_Z^{i+1}|u_Z] &= E\left[\log\left(\frac{\mu(v_{Zk}|d_Z)}{\mu(v_{Zk}|u_Z)}\right)\middle| u_Z\right] \\
&< \log\left(E\left[\frac{\mu(v_{Zk}|d_Z)}{\mu(v_{Zk}|u_Z)}\middle| u_Z\right]\right) \\
&= \log\left(\left[\mu(v_{Zh}|u_Z)\frac{\mu(v_{Zh}|d_Z)}{\mu(v_{Zh}|u_Z)} + \mu(v_{Zl}|u_Z)\frac{\mu(v_{Zl}|d_Z)}{\mu(v_{Zl}|u_Z)}\right]\right) \\
&= \log\left[\mu(v_{Zh}|d_Z) + \mu(v_{Zl}|d_Z)\right] \\
&= \log[1] \\
&= 0,
\end{aligned} \tag{7}$$

where the inequality follows from Jensen's inequality. Thus, the log of the likelihood ratio follows a generalized random walk with a downward drift. This implies (see Chung (1974) p.263) that with positive probability the log likelihood ratio may never return to the initial point. Since the expected payoff is a decreasing function of the likelihood ratio, with positive probability the expected payoff is larger than μ forever and an inefficient cascade can last forever.

(iv): When payoffs are stochastic and the conditional payoff from the down state of the adopted project k is strictly smaller than the maximum expected payoff of all other projects given the best signals, a long series of adverse payoffs can cause individuals to believe that project k is in the down state. Thus, a cascade can break when the most favorable signals for other projects are received. Thus, the probability a cascade lasts forever is less than one. ||

Proof of Proposition 2: By assumption, instead of directly learning payoffs from project n, individuals obtain a signal e_n about the payoff which can be e_{nG} (good) or e_{nB} (bad). The probability of observing e_{nq} given the true payoff of v_{nk} is $p(e_{nq}|v_{nk})$. Thus, the probability of observing e_{nq} given the true state of project n is s_n is

$$p(e_{nq}|s_n) = p(e_{nq}|v_{nh})\mu(v_{nh}|s_n) + p(e_{nq}|v_{nl})\mu(v_{nl}|s_n).$$

By assumption, for $s_n \neq s_n'$, $p(e_{nq}|s_n) \neq p(e_{nq}|s_n')$ for at least one q. Note that individuals are risk neutral and their decision rule depends only on the expected payoff of project n which is $\bar{v}_n$. The public signal e_n is a noisy signal of the expected payoff $\bar{v}_n$. Therefore the model is isomorphic to the model in Proposition 1. Thus, Proposition 2 follows immediately. ||

Proof of Result 1: We first show that $\lim_{i\to\infty} E[W^i] = \lim_{\delta\to 1} E[W(\delta)]$. Let us define $W_0 \equiv \lim_{i\to\infty} E[W^i]$. Since $E[W^i] \to W_0$, for all $\epsilon > 0$, there exists an N such that $|E[W^i] - W_0| < \epsilon$ for all $i \geq N$. Thus,

$$\begin{aligned} |E[W(\delta) - W_0]| &= \left|(1-\delta)\sum_{i=1}^{\infty} \delta^{i-1}(E[W^i] - W_0)\right| \\ &\leq \left|(1-\delta)\sum_{i=1}^{N-1} \delta^{i-1}(E[W^i] - W_0)\right| + \delta^{N-1}\epsilon. \end{aligned} \tag{8}$$

Let $\bar{M} \equiv \sum_{i=1}^{N-1} \delta^{i-1} |E[W^i] - W_0|$, and let $\delta' \equiv 1-(\epsilon/\bar{M})$. Then there exists a δ such that for all $1 > \delta > \delta'$, $|E[W(\delta) - W_0| < 2\epsilon$. Since ϵ is arbitrary, we have $\lim_{\delta\to 1} E[W(\delta)] = W_0$. Let R_1, R_2 denote two economic regimes. If the limit of expected expected payoffs of individuals in R_1 strictly dominates that in R_2, there exists a δ', that for all $\delta > \delta'$, the social welfare function $E[W(\delta)]$ dominates that in R_2.

No Communication/Observation of Payoffs: If individuals cannot observe or communicate past payoffs, then consider the subcase in which project Y's payoff is $2 + \epsilon$. If the first individual receives an H signal, then he will choose project Y. The next individual will choose project Y even if he receives an L signal since the prior mean of Y is higher than the prior mean of project Z. There is a Y cascade immediately. If the first two individuals receive two L signals, then the next individual will choose project Z

independent of his signal and a Z cascade forms. Finally, if the sequence of the signals is LH, then Individual 3 faces the same situation as the first individual. Starting after the first two individuals, the probability of a project Y cascade is p, the probability of a Z cascade is $(1-p)^2$, and the probability of no cascade is $p - p^2$. Consequently, the ratio of the likelihoods of a project Y cascade and a project Z cascade is $p/(1-p)^2$. Since a cascade will form with probability one, the long-run probability of a project Y cascade is

$$\frac{p}{1-p+p^2},$$

and the long run probability of a project Z cascade is

$$\frac{(1-p)^2}{1-p+p^2}.$$

Alternatively, when the payoff of project Y is zero, the long-run probability of a Y cascade is

$$\frac{1-p}{1-p+p^2},$$

and the long-run probability of a project Z cascade is

$$\frac{p^2}{1-p+p^2}.$$

Thus, as $i \to \infty$, the expected payoff of an individual i converges to

$$W_0 = 0.5\left[\left(\frac{p}{1-p+p^2}\right)(2+\epsilon) + \frac{(1-p)^2}{1-p+p^2}\right] + 0.5\left(\frac{p^2}{1-p+p^2}\right).$$

Communication/Observation of Payoffs: In contrast, if individuals can learn the payoffs of project Z, a cascade forms immediately after one round if the payoff of project Z is $1+\epsilon$. This is because if the signal is H, the first individual will choose Y and there will be no observation of project Z thereafter. If the signal is L, Z will be adopted and its payoff $1+\epsilon$ will stop anyone who wants to try Y. If the payoff of project Z is $1-\epsilon$, the analysis is the same as the case without payoff observations. Consequently, as $i \to \infty$, the expected payoff of an individual i converges to

$$\begin{aligned} E[W_p] &= 0.25[p(2+\epsilon) + (1-p)(1+\epsilon)] + 0.25p(1+\epsilon) \\ &\quad +0.25\left[\left(\frac{p}{1-p+p^2}\right)(1+2\epsilon) + 1 - \epsilon\right] + 0.25\left[\left(\frac{p^2}{1-p+p^2}\right)(1-\epsilon)\right]. \end{aligned}$$

We therefore have

$$W_0 - E[W_p] = 0.25\left[\frac{p(1-p)^2}{1-p+p^2}\right]\left(\frac{p}{1-p} - 1 - \epsilon\right) > 0.$$

The last inequality follows by the assumed parameter restriction.

||

Proof of Result 2: The assumption here is that the payoffs of project Z can be observed with one period delay. Let (v_Y, v_Z) denote the payoffs of project Y and Z. We calculate the conditional expected payoffs given the states and then determine the unconditional expectations. There are four possible payoff combinations:

Case I, $(2+\epsilon, 1+\epsilon)$: If the first signal is H, there is a Y cascade since the payoffs of Y are not observable. The probability of such a cascade is p.

If the first signal is L, the first individual adopts project Z. The second individual adopts project Z if he receives another L signal. A cascade forms since the third individual further observes the past payoffs of project Z and adopts project Z regardless of his signal. The probability of this cascade is $(1-p)^2$.

If the signal sequence is LH, then Individual 3 infers the signal sequence and he also observes that the payoff of project Z is $1+\epsilon$. The conditional probability of a Z cascade is $(1-p)/(1-p+p^2)$ and the conditional probability of a Y cascade is $p^2/(1-p+p^2)$. Consequently, the probability of a Y cascade is

$$p_Y = p + \frac{p(1-p)p^2}{1-p+p^2} = \frac{p - p^2(1-p)^2}{1-p+p^2}.$$

Thus, the payoff of an individual sufficiently far along in the queue is, in the limit,

$$E[W_I] = \frac{p - p^2(1-p)^2}{1-p+p^2} + 1 + \epsilon.$$

The expected payoffs of individuals far down the line for the other three cases can be derived similarly.

Case II, $(0, 1+\epsilon)$:

$$E[W_{II}] = \frac{1 - p - p^2(1-p)^2}{1-p+p^2}(-1-\epsilon) + 1 + \epsilon.$$

Case III, $(2+\epsilon, 1-\epsilon)$:

$$E[W_{III}] = \frac{p}{1-p+p^2}(1+2\epsilon) + 1 - \epsilon.$$

Case IV, $(0, 1-\epsilon)$:

$$E[W_{IV}] = \frac{p^2}{1-p+p^2}(1-\epsilon).$$

Finally, the expected payoff of an individual who is late in the queue in the limit is

$$\begin{aligned} E[W_d] &= 0.25E[W_I + W_{II} + W_{III} + W_{IV}] \\ &= 0.25\left\{\frac{p - p^2(1-p)^2}{1-p+p^2} + 1 + \epsilon + \left[\left(\frac{1-p-p^2(1-p)^2}{1-p+p^2}\right)(-1-\epsilon) + 1 + \epsilon\right]\right. \\ &\quad \left. + \left[\left(\frac{p}{1-p+p^2}\right)(1+2\epsilon) + 1 - \epsilon\right] + \left[\left(\frac{p^2}{1-p+p^2}\right)(1-\epsilon)\right]\right\}. \end{aligned}$$

It is straightforward to show that

$$E[W_d] - W_0 = 0.25\left[\frac{p^2(1-p)^2}{1-p+p^2}\right]\epsilon > 0.$$

Therefore,

$$E[W_d] > W_0 > E[W_p].$$

||

Proof of Proposition 3: When payoffs are communicated or observed, the proof is similar to that of Proposition 1 and is omitted. We now consider the case without observation or communication of payoffs. Following the same notation as in the proof of Proposition 1, let l_n^i denote the likelihood ratio of project n for individual i, F_i denote the public information set of individual i,

$$l_n^i \equiv \frac{Pr(d_n|F^i)}{Pr(u_n|F^i)}$$

Given this likelihood ratio, the probability that project n is in the up state is $1/(1+l_n^i)$.

Using public information, Individual i with likelihood ratios l_Y, l_Z will be indifferent between the two projects Y, Z if the expression (4) is zero. Without loss of generality, we assume that the true state is (u_Y, u_Z). Then l_Y^i, l_Z^i are martingales. By the martingale convergence theorem (see Chung (1974)), there exists a random vector (l_Y^∞, l_Z^∞) that (l_Y^i, l_Z^i) converges to (l_Y^∞, l_Z^∞) almost surely.

Suppose that $f(l_Y^\infty, l_Z^\infty) > 0$. Then along the sample path that leads to (l_Y^∞, l_Z^∞), there exists an N such that for all $i > N$, Individual i will adopt Y. If there is a break in the cascade on project Y, (l_Y^i, l_Z^i) must change by a finite amount, which is inconsistent with the fact that (l_Y^i, l_Z^i) converges. Similarly, when $f(l_Y^\infty, l_Z^\infty) < 0$ is negative, along the sample path that leads to (l_Y^∞, l_Z^∞), there exists an N such that for all $i > N$ Individual i adopts Z. Let $\Gamma \equiv \{\omega : f(l_Y^\infty, l_Z^\infty) \neq 0\}$. Then on any sample paths of Γ, cascades will be shattered only a finite number of times. For any $\epsilon > 0$, there exists an i' such that for $i > i'$, $\left|l_Y^{\infty} - l_Y^{i-2}\right| < \epsilon$, $\left|l_Z^{\infty} - l_Z^{i-2}\right| < \epsilon$, $\left|l_Y^{\infty} - l_Y^{i-1}\right| < \epsilon$, $\left|l_Z^{\infty} - l_Z^{i-1}\right| < \epsilon$. Since ϵ can be arbitrarily small, this indicates that on any sample paths in Γ, there exists an i' such that for $i > i'$, individual i will make the same decision as individuals $i-2, i-1$ regardless of his own signal or his predecessor's signal. Cascades eventually occurs with probability one if $f(l_Y^{\infty}, l_Z^{\infty}) \neq 0$ with probability one.

Next we show that $f(l_Y^{\infty}, l_Z^{\infty}) \neq 0$ with probability one, which implies that cascades form with probability one. Suppose otherwise and let ω denote such a sample path that $f(l_Y^{\infty}, l_Z^{\infty}) = 0$. We have $l_Y^{\infty} + l_Z^{\infty} \neq 0$ since $f(0,0) \neq 0$. In addition, an individual would be indifferent between the two projects based on the public information in the limit. When information makes Individual i almost indifferent between the two projects based on public information, favorable signals that Individual $i+2$ has about Y will cause the likelihood ratio to favor Y strictly over Z. In this case, since $l_Y^{\infty} + l_Z^{\infty} \neq 0$, the likelihood ratio will be updated by a finite amount between individual i and $i+2$ with positive probability bounded away from zero. Thus (l_Y^i, l_Z^i) can change infinitely often. Since (l_Y^i, l_Z^i) converges with probability one, (l_Y^i, l_Z^i) can change infinitely often only on a set of zero probability measure. Let O denote the set of sample paths such that the limit of likelihood ratio exists and the value of function f evaluated at the limit is zero, i.e., $O \equiv \{\omega : \lim_{i\to\infty}(l_Y^i, l_Z^i) = (l_Y^{\infty}, l_Z^{\infty}), \text{ and } f(l_Y^{\infty}, l_Z^{\infty}) = 0\}$. Then the probability measure of O is zero. Thus, with probability one, $f(l_Y^{\infty}, l_Z^{\infty}) \neq 0$. The proof that cascades are long-run inefficient is similar to that in the proof of Proposition 1 and is omitted. $\|$

Proof of Proposition 4: We divide the proof into several steps. First, we show that with probability one, it occurs infinitely often that an individual decides not to converse with his predecessor. Let (Ω, F, P) be the associated probability space.

Let C_i denote the index variable for conversation, i.e., $C_i = 1$ if Individual i converses and $C_i = 0$ otherwise. Let Q denote the set of sample paths with the following property. On each sample path belonging to Q, there exists an i', such that all individuals later than i' will converse with their predecessor. That is, $Q \equiv \{\omega :$ there exists an i' such that $C_i = 1$ for all $i > i'\}$. Since all late individuals on Q are going to converse with their predecessors, the probability that individuals makes the wrong decision infinitely often is zero. Let d_n^i denote the indicator variable if individual i take project $n, n = Y, Z$. Let $\bar{d}_n^i$ denote the histories of $d_n^{i'}$ for all individuals $i' \leq i$. Let $\bar{v}_n$ denote conditional expectation variable that gives $\bar{v}_n(s_n)$ on state s_n.

Let O_{ni} denote the set of sample paths with the following property. On each sample path belonging to O_{ni}, project n dominates project n' and individual $i-1$ does not choose project n, but all individuals later than $i-1$ do choose project n. Let $i-1$ denote the index for the last individual who does not take the correct project. Since there is no individual with index number 0, for the ease of exposition, we define $d_n^0 = 0$ for all projects n. Then, $O_{ni} = \{\omega : \bar{v}_n > \bar{v}_{n'} \text{ and } d_n^{i-1} = 0, d_n^{i'} = 1, \text{ for all } i' \geq i\}$. Then $Q/(\cup_{i,n} O_{ni})$ has a measure of probability zero. Consider a sample path in $Q \cap O_{Yi}$. Based on public information $F_{i'}$, we have

$$\frac{Pr(\bar{v}_Y > \bar{v}_Z | F_{i'})}{Pr(\bar{v}_Y < \bar{v}_Z | F_{i'})} = \frac{Pr(\bar{v}_Y > \bar{v}_Z)}{Pr(\bar{v}_Y < \bar{v}_Z)} \frac{Pr(\bar{d}_{Yi}, d_{Yi'} = 1 \text{ for all } i' > i | \bar{v}_Y > \bar{v}_Z)}{Pr(\bar{d}_{Yi}, d_{Yi'} = 1 \text{ for all } i' > i | \bar{v}_Y < \bar{v}_Z)} \to \infty.$$

Thus, the value of conversing approaches zero as it becomes almost certain that $\bar{v}_Y > \bar{v}_Z$, and Individual i' will not converse for i' very large. This contradicts the fact that $\omega \in Q$. Thus $Q \cap O_{Yi}$ has a probability measure of zero which implies that $Q \cap (\cup_{i,n} O_{ni})$ has a probability measure of zero. Moreover, $Q/(\cup_{i,n} O_{ni})$ also has a probability measure of zero. Thus Q has a probability measure of zero.

Without loss of generality, suppose the true state is (u_Y, u_Z). Then the likelihood ratio $(l_Y^i,\ l_Z^i)$ defined in the proof of Proposition 3 must converge to a random vector (l_Y^∞, l_Z^∞). Since, (l_Y^i, l_Z^i) converges, if it converges to a point (l_Y^∞, l_Z^∞) such that $f(l_Y^\infty, l_Z^\infty) > 0$, all late individuals will cascade/herd on project Y. Similarly all individuals will cascade/herd on Z if (l_Y^i, l_Z^i) converges to a point (l_Y^∞, l_Z^∞) such that $f(l_Y^\infty, l_Z^\infty) < 0$.

Moreover, we show that $f(l_Y^\infty, l_Z^\infty) = 0$ occurs with probability zero. Suppose otherwise and let ω denote a sample path such that $f(l_Y^\infty, l_Z^\infty) = 0$. We have $l_Y^\infty + l_Z^\infty \neq 0$ since $f(0,0) \neq 0$. In addition, one would be indifferent between the two projects based on the public information in the limit. Suppose the opposite is true. Notice that with probability one, it occurs infinitely often that individuals will not converse with his immediate predecessor. Let i_t be the index of such individuals. For a large enough i_t, based on public information confronting Individual $i_t + 1$, $f(l_Y^{i_t+1}, l_Z^{i_t+1})$ will be close to zero, so an individual would be almost indifferent between the two projects based on public information. This means that with positive probability, individual $i_t + 1$ will take a project different from i_t and such probability is bounded below by a positive number from zero. Thus (l_Y^i, l_Z^i) can change infinitely often. Since (l_Y^i, l_Z^i) converges with probability one, (l_Y^i, l_Z^i) can change infinitely often only on a set of zero probability measure. Thus, with probability one, $f(l_Y^\infty, l_Z^\infty) \neq 0$.

Finally, we show that with probability one individuals will eventually not observe predecessors' information. With probability one, the likelihood ratio converge to (l_Y^∞, l_Z^∞) such that $f(l_Y^\infty, l_Z^\infty) \neq 0$. Along almost all such sample paths, it is infinitely often that individuals will not converse. Let i_t be an individual who decides not to converse but

there are individuals after i_t who will converse. Let i'_t be the individual who is the first to converse again after i_t. Individuals will converse only if conversation will change their decision with positive probability, which means that the likelihood ratio will change by a finite amount to switch. Thus with positive probability, Individual i' will cause the likelihood ratio to change by a finite amount. However, (l^i_Y, l^i_Z) converges so that (l^i_Y, l^i_Z) cannot change by a finite amount bounded away from zero infinitely often. Thus, individuals will not converse infinitely often, and with probability one individuals will eventually not obtain his predecessors' information. The proof that cascades are long-run inefficient is similar to that in the proof of Proposition 1 and is omitted.
||

Proof of Result 3: Let l^i denote the public likelihood ratio between bad state and good state of project Y after Individual i's action becomes public. Let S denote the interval

$$\left(\frac{3(1-p)^2}{p(2-p)}, \frac{3p^2}{1-p^2}\right),$$

let H denote the set composed of i such that Individual i does not acquire information and $l^i \in S$. We will require the following lemmas.

Lemma 1 *If Individual $i+1$ decides not to converse with his predecessor and $l^i \in S$, then $l^{i+1} \in S, i+1 \in H$.*

Proof: If Individual $i+1$ is in a cascade, then l^{i+1} will not change and is thus in S. If Individual $i+1$ is not in a cascade, then l^i cannot be too large to dominate one bad signal and cannot be too small to dominate one good signal. Thus,

$$l^i \in [2-2p, 2p]$$

and

$$l^{i+1} \in \left[\frac{1-p}{p}, \frac{p}{1-p}\right] \in S.$$

Consequently, $i+1 \in H$ follows by definition.
||

Lemma 2 *If individual i_1 decides not to converse with his predecessor, and $l^{i_1} \in S$, then there is always a later individual $i_2 > i_1$ who will not converse with his predecessor and $l^{i_2} \in S$.*

Since the first individual has no predecessor to converse with and his likelihood ratio is in set S, Result 3 holds immediately by Lemma 2. Next we prove Lemma 2.

Proof: Let i be an individual who decides not to converse with his immediate predecessor, and $l^i \in S$. If the next individual also does not converse with Individual i, by Lemma 1, $l^{i+1} \in S$ and Lemma 2 holds. Thus, we can focus on the case in which the next individual converses with his predecessor. We consider several cases:

Case I: $l^i \leq 2p(2-2p)$.

Let

$$K \equiv -\frac{\ln(2-2p)}{\ln(2p)},$$

i.e., K is the smallest integer such that K H signals dominate one L signal and

$$(2p)^K(2-2p)^2 > 1.$$

We will require the following further lemma.

Lemma 3 *Suppose that Individual i has decided not to converse with his predecessor and $l^i \in S$, $l^i \leq 2p(2-2p)$. Then for all $m \geq i+1$, there are two possibilities: (i) For all $j \leq m$, Individual j converses with his predecessor; all individuals $j < m$ take action Y; and for all $j \leq m, l_j \in S$, there is at most one L signal among the signals between individuals $i+1$ to $i+m$, and*

$$l^i > (2p)^{m-N-1}(2-2p)^2;$$

or

(ii) there exists a $j \leq m$ such that j will not converse with his predecessor and $l^{j'} \in S$ for all $j' \leq j$.

Proof: We prove this by induction. First, the conclusion of Lemma 3 holds for $m = 1$.

Suppose that Individual $i+1$ has decided not to converse with his predecessor. Then $l^i \in S$ implies $l^{i+1} \in S$ by Lemma 1. Suppose Individual $i+1$ decides to converse with his predecessor. Then we must have $l^i > (2-2p)^2$. Otherwise the individual will cascade/herd even when he observes 2 L signals. Individual i takes project Y since $l^i < 1$. If an individual breaks the cascade, then his action reveals that he observed $2L$ signals and

$$l_{i+1} \in \left(1, \frac{2p}{2-2p}\right) \in S.$$

If Individual $i+1$ cascades/herds, he observes at most one L signal and

$$l^{i+1} \in \left(\frac{3(2-2p)^2}{2p(4-2p)}, \frac{3(2-2p)}{4-2p}\right) \in S.$$

We next prove that if the conclusion of Lemma 3 holds for Individual $m \geq i+1$, it must also hold for Individual $m+1$. Suppose that part (ii) is true for Individual m. Then it must be true for Individual $m+1$. Suppose instead part (i) is true for Individual m. If Individual m decides to break the cascade/herd and take project Z, this means that he must have observed exactly two L signals, because one L signal combined with other H signals and the prior is not enough to break the cascade/herd. In addition, the number of L signals cannot be larger than two because previous signals include at most one L signal. Individual $m+1$ will not converse with his predecessor and $l_{m+1} \in S$ by Lemma 1. If m still cascades/herds and $m+1$ decides not to obtain information, by Lemma 1, $l^{m+1} \in S$.

If Individual m cascades/herds and Individual $m+1$ decides to converse, then

$$l^i > (2p)^{k+1-N-1}(2-2p)^2.$$

In this case, if he breaks the cascade, his action will fully reveal his signal and implies that

$$l^{m+1} \in \left(1, \frac{2p}{2-2p}\right) \in S.$$

If Individual $m+1$ cascades/herds, he observes at most one L signal and

$$l^{m+1} \in \left(\frac{(m-i+2)(2-2p)^2}{2p(2p+(m-i+1)(2-2p)}, \frac{(m-i+2)(2-2p)}{2p+(m-i+1)(2-2p)}\right) \in S.$$

Thus Lemma 3 holds for $m+1$ and by induction it holds for all $m \geq i+1$. $\|$

Now for $m = i+K+2$, Part (i) of Lemma 3 cannot hold because if it did, then

$$l^i > (2p)^{K+1}(2-2p)^2 < 2p(2-2p),$$

which contradicts our assumption. Thus, Part (ii) of Lemma 3 must hold and Lemma 2 follows for Case I.

Case II: $l^i > 2p(2-2p)$.

If Individual $i+1$ decides not to converse with his predecessor, then $l^{i+1} \in S$ by Lemma 1.

If Individual $i+1$ decides to converse with his predecessor, then individual i's action is not fully revealing and $l^{i-1} = l^i < (2p)^2$. Suppose that Individual $i+1$ accepts project Y. This implies that he has observed two HH signals. In this case, l^{i+1} lies in

$$\left(\frac{2-2p}{2p}, \frac{1}{2p},\right)$$

and he is in set S. Individual $i+2$ will not converse with his predecessor and $l^{i+2} \in S$. If Individual $i+1$ takes project Z, this implies that there could be at most one H signal among the signals he observed. If Individual $i+2$ does not converse with his predecessor, then

$$l^{i+1} \in \left(\frac{3[2p(2-2p)]}{(2-2p)^2 + 2(2p)(2-2p)}, \frac{3(2p)^2}{(2-2p)^2 + 2(2p)(2-2p)}\right) \in S,$$

and $l^{i+2} \in S$ by Lemma 1. In this scenario, Lemma 2 holds.

If Individual $i+2$ decides to converse with his predecessor, this means that the signals Individual $i+2$ observes may break the cascade/herd. The most favorable signals for project Y are HHL. Thus, for the conversation to be valuable, we must have

$$l^i > (2p)^2(2-2p).$$

If Individual $i+2$'s action is fully revealing, then Individual $i+2$ must have observed a combination of HHL signals. This means that

$$1 > l^{i+2} = \frac{l^{i-1}}{(2p)^2(2-2p)} = \frac{l^i}{(2p)^2(2-2p)} > \frac{1}{2p}.$$

Thus,

$$l^{i+2} \in \left(\frac{1}{2p}, 2p\right) \in S.$$

Moreover, since Individual $i+2$'s action fully reveals his information, Individual l^{i+3} will not converse with his predecessor and $l^{i+3} \in S$ by Lemma 1. In this scenario, Lemma 2 holds.

Finally, if Individual $i+2$ still follows the cascade/herd, then the most favorable possible signal combination for project Y of individuals i to $i+2$ is HLL. Individual $i+3$ will not converse with his predecessor, because even if he learns of 2 H signals and 2 L signals, the private likelihood ratio of Individual $i+3$ will still be larger than

$$\frac{1}{(2p)^2(2-2p)^2 l^i} > \frac{1}{2p(2-2p)} > 1,$$

so he will follow the cascade/herd. Thus Individual $i + 3$ will not converse with his predecessor and

$$l^{i+2} \in \left(\frac{(2p)^2}{(2-2p)(0.5+p)}, \frac{(2p)^2}{(2-2p)(0.5+p)} \right) \in S.$$

By Lemma 1, $l^{i+3} \in S$. Thus Lemma 2 must also hold for Case II. This completes the proof of Lemma 2. Following Lemma 2 and the proof of Lemma 2, the public likelihood ratio lies in S which is inside the interval of the likelihood ratio of three L signals and its inverse. The welfare loss is finite and bounded below by a positive number as the likelihood ratio is finite and bounded below by a positive finite number. This completes the proof of Result 3. $\|$

References

Anderson, Lisa, and Charles Holt, "Information Cascades in the Laboratory," *American Economic Review*, 1997, 87, 847-862.

Avery, Christopher and Peter Zemsky, "Multi-Dimensional Uncertainty and Herd Behavior in Financial Markets," *American Economic Review*, 88(4), (1998):724-48.

Banerjee, Abhijit. "A Simple Model of Herd Behavior." *Quarterly Journal of Economics*, 107, August (1992):797-818.

Banerjee, Abhijit and Fudenberg, Drew. "Word-of-Mouth Communication and Social Learning." Harvard University Working Paper, (1999).

Bernardo, Tony and Welch, Ivo. "On the Evolution of Overconfidence and Entrepreneurs." *Journal of Economics and Management Strategy*, Fall 2001.

Bikhchandani, Sushil; Hirshleifer, David; and Welch, Ivo. "A Theory of Fads, Fashion, Custom and Cultural Change as Informational Cascades." *Journal of Political Economy*, 100(5), October (1992):992-1026.

—"Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades," *Journal of Economic Perspectives*, 12(3), (1998):151-170.

Bolton, Patrick and Harris, Christopher, "Strategic Experimentation." *Econometrica*, 67(2), March (1999): 349-74.

Bose, Subir, Orosel, Gerhard and Lise Vesterlund, "Optimal Pricing and Endogenous Herding," Iowa State University, 2001

Brunnermeier, Markus, *Asset Pricing under Asymmetric Information: Bubbles, Crashes, Technical Analysis, and Herding*, (2001), Oxford University Press, Oxford.

Caplin, Andrew and Leahy, John. "Sectoral Shocks, Learning and Aggregate Fluctuations." *Review of Economic Studies*, 60, (1993):777-794.

Celen, Bogachan and Kariv, Shachar, "A Laboratory Experiment: Distinguishing Informational Cascades from Herd Behavior," New York University, October, 2001.

Marco Cipriani and Antonio Guarino, 2001, "Herd behavior and contagion in financial markets," New York University Economics Department, Working Paper.

Chamley, Cristophe and Gale, Douglas. "Informational Revelation and Strategic Delay in Irreversible Decisions," *Econometrica*, 62(5), September (1994):1065–1085.

Chung, Kai Lai. *Elementary Probability Theory with Stochastic Processes*, New York, NY: Springer-Verlag, 1974.

Chalkley, Martin and Lee, In Ho, 1998, "Learning and Asymmetric Business Cycles," *Review of Economics Dynamics*, 1(3), 623-645.

Dasgupta, Amil, "Social Learning with Payoff Complementarities," May, 2000, Kellogg School, Northwestern University.

Ellison, Glenn and Fudenberg, Drew (1993), "Rules of Thumb for Social Learning," *Journal of Political Economy*, 101, 612–643.

Ellison, Glenn and Fudenberg, Drew (1995), "Word-of-Mouth Communication and Social Learning," *Quarterly Journal of Economics*, 110(1), 93-125.

Herr, Paul M., Kardes, Frank R., and Kim, John. "Effects of Word-of-Mouth and Product-Attribute Information on Persuasion: An Accessibility-Diagnosticity Perspective," *Journal of Consumer Research*, vol. 17, March (1991):454-62.

Hirshleifer, Jack. "The Private and Social Value of Information and the Reward to Inventive Activity," *American Economic Review*, 61(4), September (1971):561-574.

Hong, Harrison, Kubik, Jeffrey and Stein, Jeremy, "Social Interactions and Stock Market Participation," Stanford Business School Working Paper, February (2001).

Kelly, Morgan and O'Grada, Cormac O., "Market Contagion: Evidence from the Panics of 1854 and 1857," *American Economic Review*, 90(5), (2000):1110-1124.

Khanna, Naveen and Slezak, Steve L. "The Effect of Organizational Form on Information Aggregation and Project Choice: The Problem of Informational Cascades." *Journal of Economics and Management Strategy*, 9(1), Spring (2000):115-56.

Lee, In Ho, "Market Crashes and Information Avalanches," *Review of Economic Studies* 65(4), (1998):741-59.

Neeman, Z. and Orosel, G.O., 1999, March, "Herding and the winner's curse in markets with sequential bids," *Journal of Economic Theory*, 85, 1, 91-121.

Ottaviani, Marco, "Monopoly Pricing with Social Learning," London Business School, 1999.

— and Sorensen, Peter, "Information Aggregation in Debate: Who Should Speak First?" *Journal of Public Economics*, 81:(3), (2001):393-421

Rothschild, Michael. "A Two-Armed Bandit Theory of Market Pricing," *Journal of Economic Theory*, 9 October (1974):185-202).

Shiller, Robert J., *Irrational Exuberance*, Princeton University Press, 2000.

Shiller, Robert J. and Pound, John. "Survey Evidence on the Diffusion of Interest and Information Among Investors," *Journal of Economic Behavior and Organization*, 12 (1989):47-66.

Teoh, Siew Hong, "Information Disclosure and Voluntary Contributions to Public Goods." *Rand Journal of Economics*, 28(3), Autumn (1997):385-406.

Vives, Xavier, How Fast Do Rational Agents Learn, *Review of Economic Studies*, 60, (1993):329–347.

Welch, Ivo. "Sequential Sales, Learning and Cascades." *Journal of Finance* 47 (1992).

Zhang, Jianbo, "Strategic Delay and the Onset of Investment Cascades," *Rand Journal of Economics*, 1997, 28, 188-205.