Essays in Learning and the Revelation of Private Information

by

Robert Bruce Noah

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in The University of Michigan
1998

Doctoral Committee:

Professor Ted Bergstrom (chair)
Professor David Hirshleifer
Professor Carl Simon
Professor Tayfun Sönmez

UMI Number: 9909956

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

Table of Contents

List of Figures

Figure

List of Tables

Table

Chapter 1

Reputation Effects, Sequential Cascades and Speculative Investment

1.1 Introduction

Informational cascades have previously been studied as singular events. Bikhchandani, Hirshleifer and Welch [4] define an informational cascade as an event where individuals, in sequence, choose to copy the play of an earlier player without regard to any private information they might have. Models of such an event place players in a queue, one following the other, in some pre-determined ordering that is known to all. Later players observe the actions of earlier players, but not necessarily the private information that these earlier players may possess. When the precision of information that a player infers from the play of earlier players exceeds that of his own information, the player rationally chooses to act according to the inferred information. Such action typically induces later players to do the same, resulting in the informational cascade. The obvious problem with informational cascades is that a great deal of private information may not be revealed in the process of player actions.

It is reasonable to expect cascades to occur in a series of similarly structured events. Moreover, the particular structure of the individual events will often have the property that the ordering of players is similar across time. In particular, the *leaders* or early players are very often the same individuals in each event. For example, a market analyst advises clients of the value of not just one IPO, but of a number of IPOs over time. An industry leading firm may invest in a series of new technologies, where each investment decision affects the actions of other firms. For each IPO or new investment, a cascade among followers may occur.

The primary theoretical innovation of this paper is to apply the observation that leaders will tend to be involved in numerous cascades to a model where the reputations of leader's can be open to manipulation. This paper views cascades as components of a series of events and not as singular, isolated occurrences. In each event, players, in sequence, choose whether to purchase a project with an uncertain payoff. Player action space will thus be discrete and, as a result, cascades will eventually occur. The precision of a player's private information will vary among players. In particular, players can have different probabilities of committing type 1 and type 2 errors in project valuation. This asymmetry in each player's ability to identify project quality coupled with well informed leaders will result in an opportunity for project producers to price projects in order to manipulate the reputations of leaders in subsequent events. IPO underpricing is presented as an example of this behavior, where underpricing favors the reputations of leaders more desirable to investment bankers likely to face the same leaders during subsequent new issues. In addition to the IPO market, the model of leader reputations and sequential cascades is applied to capital investment decisions and cyclical production patterns.

Bikhchandani, Hirshleifer and Welch [4] and Banerjee [1] introduced informational cascades (or herd behavior) as a concept that can explain localized conformity. Bikhchandani, Hirshleifer and Welch also demonstrated the instability of this conformity with respect to new information. Lee [14] examines the convergence of cascades and shows that the number of actions available to a player is critical to the existence of cascades. Generally, a larger set of best-response actions available to players results in greater information revelation, with a continuum of actions resulting in full information revelation. Zhang [20] studies the ordering of players within cascades and show that, in general, players with the most precise information choose to move first. Gul and Lundholm [9] also endogenize player ordering and find that the actions of players tend to cluster together, where the players with the most extreme information choose to act first. Chamley and Gale [6] examine strategic delay in a model of investment and, in some cases, find informational cascades. Informational cascades have been applied to the market for initial public offerings by Welch [18]. A primary result is that evidence of IPO underpricing can be explained by issuers who under-price in an attempt to avoid cascades where players choose not to purchase the new offering.

The remainder of this paper is structured as follows. In the next section, a model is presented where a player is offered a project at a fixed price with an uncertain payoff. Section 1.3 solves the model for a single project offering and derives the probability that a player is correct in his choice to accept or reject the project. Leaders, their impact on followers and leader reputations are discussed in section 1.4. Sections 1.5 and 1.6 discuss single period project pricing as well as project pricing over time. Section 1.7 presents applications and section 1.8 concludes the paper.

1.2 The Model

Players can either accept or reject a project. If a player rejects the project, he receives a payoff of 0. If a player accepts the project, he pays a fixed price P and receives a stochastic payoff $\tilde{V}$. A given project's payoff is not a function of the number of players that accept the project nor does this payoff vary from one player to another. All players that accept the project will pay the same price P and receive the same payoff realization V, though this payoff is unknown when players choose to accept or reject the project. Players are risk neutral wealth maximizers and will therefore choose to accept the project if the expected project payoff exceeds the price or $\mathrm{E}[\tilde{V}] > P$. Following Rock [15], Welch [18] and others, I presume players to have limited wealth such that each player can purchase at most one unit of the project.

There are two distinct types of projects. *Low quality* projects have a payoff that is distributed under the probability distribution function $\mathrm{f}_L(V)$ while *high quality* projects have a payoff that is distributed under the probability distribution function $\mathrm{f}_H(V)$. For notational convenience, $\overline{V}_H$ is the expectation of $\tilde{V}$ for a high quality project and $\overline{V}_L$ is the expected value of $\tilde{V}$ when the project is of low quality. The function $f_H(V)$ will be defined to first order stochastically dominate $f_L(V)$. As a result, $\overline{V}_H$ is greater than or equal to $\overline{V}_L$ and, at any project price P, a player would always (weakly) prefer a high-quality project over a low-quality project. The project cumulative distribution functions are defined as $F_L(V)$ (for low quality projects) and $F_H(V)$ (for high quality projects):

$$F_L(V) = \int_{-\infty}^{V} f_L(V)\, dV \qquad\qquad F_H(V) = \int_{-\infty}^{V} f_H(V)\, dV.$$

As $f_H(V)$ is defined to first order stochastically dominate $f_L(V)$, $F_L(V)$ is greater than or equal to $F_H(V)$ for all V.

For notational simplicity, θ_H will signify the event that the project is of high quality while θ_L will be the event where the project is of low quality. All players have a prior expectation that the project is equally likely to be of high or low quality. Hence, $\Pr(\theta_H) = \Pr(\theta_L) = 1/2$ and, prior to the acquisition of any information, a player expects a project's payoff to be distributed under the probability distribution function $[f_H(V) + f_L(V)]/2$ with an expected value of $\overline{V} = (\overline{V}_H + \overline{V}_L)/2$. Note that a project's payoff is truly uncertain. No amount of information can perfectly identify the value of $\tilde{V}$. At best, a perfectly informed player will be able to know with certainty that a project is of high or low quality.

Each player receives a private signal S that is either high (H) or low (L). The signal is related to the project payoff as follows:

$$\Pr(S = H|\theta_H) = \alpha \qquad \Pr(S = L|\theta_L) = \beta \qquad \{\alpha, \beta\} \in [1/2, 1].$$

When the project is of high quality (the event θ_H), the player receives a high signal with probability α. When the project is of low quality (the event θ_L), the player receives a low signal with probability β. The values α and β will be referred to as the *precisions* of a player's signal. The values of these measures of signal precision may vary among players. Higher value of α and β (values closer to 1) represent higher signal precision and a greater player ability to identify the proper distribution of $\tilde{V}$. The dichotomy of player signal precision can be viewed in terms of type 1 and type 2 errors. Under the hypothesis that a project is of high quality, a type 1 error occurs when a player incorrectly identifies a project as low quality when it is in fact a high quality project. The probability of this type 1 error is $1-\alpha$. Similarly, under the same hypothesis of a high project quality, the probability of type 2 error (receiving a signal of high project quality when project quality is actually low) is $1-\beta$.

1.3 Model Solution

Consider a single player facing the choice of whether to accept or reject a project based upon his private signal as his sole source of information. Given the private signal, S, the player's posterior distribution of $\tilde{V}$ is (via Bayes' Theorem):

$$\begin{aligned}
\Pr(\theta_H|S=H) &= \Pr(S=H|\theta_H)\Pr(\theta_H)/\Pr(S=H) \\
&= (\alpha)(1/2)/[(\alpha)(1/2)+(1-\beta)(1/2)] \\
&= \frac{\alpha}{\alpha+1-\beta},
\end{aligned}$$

and, similarly:

$$\Pr(\theta_H|S=L) = \frac{1-\alpha}{1-\alpha+\beta}.$$

The posterior expectations of $\tilde{V}$ are then:

$$\begin{aligned} \mathrm{E}[\tilde{V}|S=H] &= \Pr(\theta_H|S=H)\,\mathrm{E}[\tilde{V}|\theta_H] + \Pr(\theta_L|S=H)\,\mathrm{E}[\tilde{V}|\theta_L] \\ &= \left(\frac{\alpha}{\alpha+1-\beta}\right)\overline{V}_H + \left(1-\frac{\alpha}{\alpha+1-\beta}\right)\overline{V}_L \\ &= \frac{\alpha\overline{V}_H + (1-\beta)\overline{V}_L}{\alpha+1-\beta}. \end{aligned} \tag{1.1}$$

Likewise:

$$\mathrm{E}[\tilde{V}|S=L] = \frac{(1-\alpha)\overline{V}_H + \beta\overline{V}_L}{1-\alpha+\beta}. \tag{1.2}$$

After receiving his private signal, each player will have one of two possible expectations of the project payoff $\tilde{V}$. These expectations can be ordered.

$$\overline{V}_L \leq \mathrm{E}[\tilde{V}|S=L] \leq \overline{V} \leq \mathrm{E}[\tilde{V}|S=H] \leq \overline{V}_H. \tag{1.3}$$

If the project price $P \leq \mathrm{E}[V|S=L]$, then the player will choose to purchase the project, regardless of his signal.[1] If P is in the range $(\mathrm{E}[\tilde{V}|S=L], \mathrm{E}[\tilde{V}|S=H]]$, then the player will accept the project only when he receives a high signal. When $P \geq \mathrm{E}[\tilde{V}|S=H]$, the player will never accept the project. Theorem 1.1 describes the relationship between posterior expectations of $\tilde{V}$ and total signal precision.

Theorem 1.1. *Higher values of player signal precisions α and β result in higher posterior expectations of project value upon receipt of a high signal and lower posterior expectations of project value upon receipt of a low signal.*

Proof. The values of $\mathrm{E}[\tilde{V}|S=H]$ and $\mathrm{E}[\tilde{V}|S=L]$ from (1.1) and (1.2) are simply differentiated with respect to the signal precisions. □

[1]When indifferent between project acceptance and rejection, players choose to accept the project as a tie-breaking mechanism. This does not affect the results.

Players with high signal precisions will have a larger range of prices over which they will respond to their private signal. At one extreme, a player with signal precisions $\alpha = \beta = 1/2$ will not derive any information from his private signal and will thus accept a project with a price less than or equal to $\overline{V}$ and reject a project with a price greater than $\overline{V}$. At the other extreme, a player with signal precisions $\alpha = \beta = 1$ will know with certainty the true distribution of $\tilde{V}$ and will accept a project with price less than or equal to $\overline{V}_L$, reject a project with price greater than $\overline{V}_H$ and react to his signal when the price is on $(\overline{V}_L, \overline{V}_H]$.

1.3.1 Rationally Aggressive and Tentative Players

Whatever the price of a project, a player will be more likely to accept the project if he receives a high signal.[2] The probability of receiving a high signal is $(\alpha + 1 - \beta)/2$ and is therefore increasing in the difference between α and β. When a player has a relatively high value of α and a relatively low value of β, he will be more likely to receive a high signal, while relatively high values of β and low values of α result is a greater probability of receiving a low signal.

A player is defined to be *aggressive* if he is more likely to receive a high signal than a low signal and thus have signal precisions $\alpha > \beta$. As discussed in section 1.2, high values of α can be interpreted as lower probability of committing a type 1 error in project evaluation. Similarly, low values of β represent a higher probability of committing a type 2 error in project evaluation.[3] Aggressive players are thus more likely to err by identifying a low quality project as a project of high quality than they are to err by identifying a high quality project as a project of low quality.

Tentative players are the opposite of aggressive players and are more likely to receive low signals than high signals. Tentative players therefore have signal precisions such that

[2]This presumes that the private signal is relevant. That is, price is not so low that a player always accepts the project or so high that a player always rejects the project. Of course, a player is always *weakly* more likely to accept any project, given a high signal.

[3]This is under the hypothesis that the project is of high quality.

$\alpha < \beta$. Tentative players are more likely to err by identifying a high quality project as a project of low quality than to err by identifying a low quality project as a project of high quality.

Aggressive and tentative players are rational, and thus realize the asymmetric nature of their signal realizations. As a result, they adjust their posterior distributions of $\tilde{V}$.

Theorem 1.2. *With equal total signal precision $\alpha + \beta$, for a given signal, aggressive players will have lower posterior expectations of project value than tentative players.*

Proof. Consider two players with equal total signal precision $2\rho + \varepsilon$ where ε is positive. The aggressive player will have precisions $\alpha = \rho + \varepsilon, \beta = \rho$ and the tentative player will have precisions $\alpha = \rho, \beta = \rho + \varepsilon$. When both players receive a high signal, the aggressive player will have a lower posterior expectation of $\tilde{V}$ than the tentative player when, via (1.1):

$$\frac{(\rho + \varepsilon)\overline{V}_H + (1 - \rho)\overline{V}_L}{1 + \varepsilon} < \frac{\rho\overline{V}_H + (1 - \rho - \varepsilon)\overline{V}_L}{1 - \varepsilon}$$

$$\frac{\varepsilon(2\rho - 1 + \varepsilon)}{(1 + \varepsilon)(1 - \varepsilon)}(\overline{V}_L - \overline{V}_H) < 0. \qquad (1.4)$$

As ε is positive and $\rho \geq 1/2$, $\varepsilon(2\rho - 1 + \varepsilon)$ must be positive. Since $\rho + \varepsilon$ must be on $[1/2, 1]$, ε can be $1/2$ at most, thus $(1 + \varepsilon)(1 - \varepsilon)$ is positive. Finally, $\overline{V}_L - \overline{V}_H$ is always negative. The relationship in (1.4) therefore always holds. The solution when both players receive a low signal is much like this high signal case and is therefore omitted. □

The nature of rational aggression in the model can be described as a greater probability of receiving a signal of high project quality, coupled with a resulting pessimism over project value. As an example, consider a perfectly aggressive player, with $\alpha = 1$ and $\beta = 1/2$. When this player receives a low signal, he can be certain that the project is of low quality as he always receives a high signal for high quality projects. Therefore, for a low signal, the perfectly aggressive player has the lowest possible posterior valuation of the project, $\overline{V}_L$. The perfectly aggressive player knows that if the project is of low quality, he has an even chance of receiving a high signal, while he always receives a high signal for high quality projects. Thus, upon receiving a high signal, the perfectly aggressive player does not value the project at the highest possible level of $\overline{V}_H$, but at the lower level of $2/3\overline{V}_H + 1/3\overline{V}_L$. A perfectly tentative player (with $\alpha = 1/2$ and $\beta = 1$) has low and high signal valuations of $2/3\overline{V}_L + 1/3\overline{V}_H$ and $\overline{V}_H$, respectively. In both signal cases, the aggressive player has lower posterior valuations of project quality. Yet, when the project is priced between $2/3\overline{V}_L + 1/3\overline{V}_H$ and $1/3\overline{V}_L + 2/3\overline{V}_H$, both players will respond to their signals and the aggressive player will be more likely to receive a high signal and accept the project.

1.3.2 The Probability of a Correct Choice

The objective of a risk neutral, wealth maximizing player is to accept projects that will succeed and reject projects that will fail. Success is defined as the event that a project's payoff exceeds the project's price. For a given project price P, the prior probability of project success is $[1 - F_H(P) + 1 - F_L(P)]/2$. This is simply the probability that the payoff will exceed P under the prior belief that the project quality is equally likely to be high or low. Clearly, projects with a lower price are more likely to succeed than are those with a higher price.

If players had no private information, then the act of project selection would be very simple — all projects with prices less than or equal to $\overline{V}$ should be purchased, while all other projects should be rejected. Upon the realization of a private signal, however, a player develops a posterior distribution of the project payoff. In the manner described previously, this posterior distribution will result in players at times purchasing projects with prices exceeding $\overline{V}$ and, at other times, rejecting projects with prices that are less than $\overline{V}$. When a player accepts a project that eventually succeeds, or rejects a project the eventually fails, the player can be said to have acted correctly.

A player can be correct in his action in two ways. He can either reject a project that eventually fails or accept a project that eventually succeeds. If a high signal results in acceptance and a low signal results in rejection, the player is correct with the following probability:

$$\Pr(S = H)\Pr(\tilde{V} > P|S = H) + \Pr(S = L)\Pr(\tilde{V} < P|S = L)$$

$$= \frac{\alpha + 1 - \beta}{2}\left(\frac{\alpha}{\alpha + 1 - \beta}[1 - \mathrm{F}_H(P)] + \frac{1 - \beta}{\alpha + 1 - \beta}[1 - \mathrm{F}_L(P)]\right)$$

$$+ \frac{1 - \alpha + \beta}{2}\left(\frac{1 - \alpha}{1 - \alpha + \beta}\mathrm{F}_H(P) + \frac{\beta}{1 - \alpha + \beta}\mathrm{F}_L(P)\right)$$

$$= \alpha[1/2 - \mathrm{F}_H(P)] + \beta[\mathrm{F}_L(P) - 1/2] + \frac{1 - [\mathrm{F}_L(P) - \mathrm{F}_H(P)]}{2}. \quad (1.5)$$

A perfectly informed player with $\alpha = \beta = 1$ is correct with probability $1/2 + [\mathrm{F}_L(P) - \mathrm{F}_H(P)]/2$ which is greater than 1/2. In addition to the expression in (1.5), the following lemma will prove to be useful.

Lemma 1.1. When $\mathrm{f}_L(V)$ and $\mathrm{f}_H(V)$ are each symmetric distribution functions, for all prices such that a private signal can possibly affect a player's action, $\mathrm{F}_L(P) > 1/2$ and $\mathrm{F}_H(P) < 1/2$.

For univariate symmetric distributions, half of the distribution's mass is on either side of it's mean. Thus, when $\mathrm{f}_H(V)$ and $\mathrm{f}_L(V)$ are symmetric, $\mathrm{F}_H(\overline{V}_H) = 1/2$ and $\mathrm{F}_L(\overline{V}_L) = 1/2$. From (1.3), regardless of his signal precisions, a player will never have a posterior expectation of $\tilde{V}$ less than $\overline{V}_L$ or greater than $\overline{V}_H$. Hence, for a player's private signal to be relevant to his actions, the project price must be on the interval $(\overline{V}_L, \overline{V}_H)$. As a result, $\mathrm{F}_L(P) > 1/2$ and $\mathrm{F}_H(P) < 1/2$.

Theorem 1.3. *When $\mathrm{f}_L(V)$ and $\mathrm{f}_H(V)$ are symmetric and project price is such that a player accepts the project given a high signal and rejects the project given a low signal, an increase in either or both signal precisions will increase a player's probability of being correct.*

Proof. This is a direct result of (1.5) and Lemma 1.1. □

Theorem 1.4. *For a project price such that a player accepts the project given a high* *signal and rejects the project given a low signal, a uniform increase in signal precisions will increase a player's probability of being correct.*

Proof. In (1.5) when $\alpha \to \alpha + \varepsilon, \beta \to \beta + \varepsilon$ with $\varepsilon > 0$, the change is $\varepsilon[\mathrm{F}_L(P) - \mathrm{F}_H(P)]$ which is positive. □

Theorems 1.3 and 1.4 imply, not surprisingly, that greater signal precision generally increases a player's chances of making a correct choice in project evaluation. As an example, when project payoffs are normally distributed, any increase in precisions will result in an increase in the probability of a correct choice.

A more interesting issue is how a player's probability of making a correct choice changes when his signal precisions become more aggressive or tentative. Consider two players with the same total signal precision where the aggressive player has signal precisions $\alpha = \rho + \varepsilon, \beta = \rho$ and the tentative player has signal precisions $\alpha = \rho, \beta = \rho + \varepsilon$ where ε is positive. From (1.5), the probability that the aggressive player is correct *minus* the probability that the tentative player is correct is:

$$\varepsilon[1 - \mathrm{F}_H(P) - \mathrm{F}_L(P)] \tag{1.6}$$

When the expression in (1.6) is positive, the aggressive player is more likely to be correct. When (1.6) is negative, the tentative player is more likely to be correct.

Theorem 1.5. *For a project price such that a player accepts the project given a high signal and rejects the project given a low signal, an aggressive player is more likely to be correct than a tentative player (with the same total signal precision) when the prior probability of project success is greater than* 1/2. *The tentative player is more likely to be correct when the prior probability of project success is less than* 1/2.

Proof. The expression in (1.6) is positive when $\mathrm{F}_L(P) + \mathrm{F}_H(P) < 1$ or the prior probability of project success $(1 - [\mathrm{F}_L(P) + \mathrm{F}_H(P)]/2)$ is greater then 1/2. □

This leads us to the primary result of this section.

Theorem 1.6. *For a project price such that a player accepts the project given a high signal and rejects the project given a low signal, the difference between the probability that an aggressive player is correct and the probability that a tentative player (with the same total signal precision) is correct is decreasing in project price.*

Proof. The expression in (1.6) is decreasing in project price as $\mathrm{F}_H(P)$ and $\mathrm{F}_L(P)$ are increasing in project price. □

For any two players with equal total signal precision, as project price rises, the probability that the aggressive player is correct falls relative to the probability the tentative player is correct. In later sections, when the discussion turns to the reputations of market leaders, the impact of project price on the probability that a particular leader is correct will be a primary determinant of leadership structure.

1.4 Leaders

The previous section solved the model for a single player. In this section, a project is offered to a series of players, with one player following another. Players can observe the actions of earlier players, but can not observe the signals that these earlier players receive. The leader is the first player and is referred to as player $\mathcal{L}$. Followers are indexed by their order of play and are referred to as players $\mathcal{F}_1 \dots \mathcal{F}_n$ where n is the total number of followers. Play proceeds in the order $\mathcal{L}, \mathcal{F}_1, \dots, \mathcal{F}_n$.

1.4.1 A Relatively Well-Informed Leader

The following discussion focuses on the impact of a leader on the amount of information that followers can be expected to receive. The measure of the amount of information available to a particular follower will be the probability that the follower is *correct* in his choice to accept or reject the project. Equation (1.5) translates signal precision to the probability that a player is correct in project selection. In order to simplify the material that follows, it will be assumed that no translation from total precision to the probability of a correct choice is necessary. Two assumptions are sufficient to achieve this simplification. (1) The project is priced at its prior expected value of $\overline{V}$. As a result, a player will accept the project if he believes the project is more likely than not to be of high quality. When players are indifferent, they randomize evenly over accepting and rejecting the project.[4] (2) High and low quality project payoffs are disjoint. Thus, at $P = \overline{V}$, $F_L(P) = 1$ and $F_H(P) = 0$. The probability that a particular player is correct is then just the player's total signal precision or $(\alpha + \beta)/2$.

To examine the effect that the leader has on the outcome of a single event, consider the case where the leader has an (accurate) reputation for being relatively well informed with a precision of $\alpha = \beta = \rho + \varepsilon$ where ε is positive. All other players have a precision of $\alpha = \beta = \rho$. The leader will observe his private signal $S_{\mathcal{L}}$ and will choose to accept the project if he receives a signal of H and will reject with a signal of L. The first follower, $\mathcal{F}_1$, will be able to infer $S_{\mathcal{L}}$ by observing the leader's action. With the common knowledge that the leader has superior precision, $\mathcal{F}_1$ will ignore his own signal and simply follow the leader. All other followers will do the same and an informational cascade will occur. Only the information of the leader's signal affects play and the probability that any player is correct is the same as that of the leader, $(\alpha + \beta)/2$ or, in this case, $\rho + \varepsilon$.

Are the followers, in this case, better off than they would have been without a relatively well-informed leader? Given a leader with the common precision of ρ, the first follower will certainly follow the leader if his signal matches the signal he infers from the leader's action. If not, then the first follower will have observed two conflicting signals with equal precision and will assess the probability of a high quality project at 1/2. As discussed earlier, in this case the follower will accept the project with probability 1/2 and reject the project with probability 1/2. The first follower will then be correct in his action when both signals ($S_{\mathcal{L}}$ and S_1) are correct (probability ρ^2) or half of the time when the signals conflict ($\rho(1-\rho)/2 + (1-\rho)\rho/2$). The sum of these probabilities is simply ρ. The first follower, $\mathcal{F}_1$, clearly benefits from the presence of the relatively well informed leader as his chances of being correct fall from $\rho + \varepsilon$ to ρ without this leader.

[4] This assumption is innocuous and merely results in a symmetrical solution.

When the actions of the leader and first follower are the same, the second follower will follow $\mathcal{L}$ and F_1 and ignore his signal S_2. In this case a cascade will occur and proceed to all subsequent players. If the leader and first follower's actions conflict, then $\mathcal{F}_2$ will act according to his own signal. The probability that the second follower will follow his own signal is equal to the probability that no cascade occurred. This is true when the signals of the two earlier players conflict and the first follower chooses not to follow the leader. As above, this probability is $\rho(1-\rho)$.

A correct cascade can occur in two ways. Either both $\mathcal{L}$ and $\mathcal{F}_1$ had the same, correct, signal (probability ρ^2) or the leader had a correct signal and the first follower had an incorrect signal with the first follower choosing the correct action with probability 1/2. The probability of this occurring is $\rho(1-\rho)/2$. When the second follower follows a correct cascade he is correct with probability equal to one. When he follows his own signal, he is correct with probability ρ. The probability the second follower is correct in his action is then:

$$[\rho^2 + \rho(1-\rho)/2] + [\rho(1-\rho)\rho] = \frac{3\rho^2 - 2\rho^3 + \rho}{2} \geq p. \tag{1.7}$$

For this second follower to be better off with the relatively well informed leader, then $\rho + \varepsilon$ must be larger than the value in (1.7) or:

$$\varepsilon > \frac{3\rho^2 - 2\rho^3 - \rho}{2} \geq 0. \tag{1.8}$$

The expression in (1.8) is zero when ρ is uninformative and equal to 1/2. In this case, a leader with any amount of increased precision benefits player $\mathcal{F}_2$. When ρ is perfectly informative (and equal to 1), the expression in (1.8) is zero again, as player $\mathcal{F}_2$ can do no better than with a perfectly informed leader. Finally, the expression for the minimum value of ε in (1.8) is maximized when ρ is $(3+\sqrt{3})/6 \approx 0.7887$ which produces $\varepsilon > 0.0482$. For player $\mathcal{F}_2$, a relatively well informed leader is acceptable, regardless of the precision of the followers' signals, if this leader has a signal precision at least 4.82% greater than the followers.

Proceeding in this manner, player $\mathcal{F}_i$ is correct if a correct cascade has occurred or if no cascade has occurred and he uses his own information and happens to be correct. *Presuming no cascade has yet occurred, for two successive players, a cascade will occur exactly as modeled above for players $\mathcal{L}$ and $\mathcal{F}_1$.* The probability of any cascade is then $1-\rho(1-\rho)$. The probability of a correct cascade is, as above, $\rho^2+\rho(1-\rho)/2$. Hence, if a cascade has occurred, the probability that it is the correct cascade is $(\rho^2+\rho)/(2[1-\rho+\rho^2])$. After $2m$ players, a cascade will *not* have occurred with probability $[\rho(1-\rho)]^m$.[5] Combining these calculations, after m pairs of players, the probability that the player $\mathcal{F}_{2m}$ is correct in his action is the sum of his chances of following a correct cascade and his chances of using his own information multiplied by his precision. For player $\mathcal{F}_{2m+1}$ to benefit from the presence of the relatively well informed leader, the following must then be true:

$$\rho+\varepsilon > \frac{(\rho^2+\rho)[1-(\rho-\rho^2)^m]}{2(1-\rho+\rho^2)} + \rho(\rho-\rho^2)^m$$

$$\varepsilon > \frac{(3\rho^2-2\rho^3-\rho)[1-(\rho-\rho^2)^m]}{2(1-\rho+\rho^2)} \geq 0. \tag{1.9}$$

Since $\rho-\rho^2<1$, the term $(\rho-\rho^2)^m$ is decreasing in m which means that ε in (1.9) is maximized as m approaches infinity.

$$\varepsilon_{m\to\infty} > \frac{3\rho^2-2\rho^3-\rho}{2(1-\rho+\rho^2)} \geq 0. \tag{1.10}$$

As for player $\mathcal{F}_2$ above, the expression in (1.10) is zero when ρ is uninformative ($\rho=1/2$) and perfectly informative ($\rho=1$). The expression in (1.10) is maximized with respect to ρ when $\rho \approx 0.7700$ resulting in a value of $\varepsilon > 0.0582$. Table 1.1 illustrates some minimal values of ε for various values of ρ and m. A couple of interesting points emerge. Primarily, while it is true that a relatively well informed leader can be harmful to some

[5]This follows a similar derivation in Bikhchandani, Hirshleifer and Welch [4].

	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=10$	$m=500$
$\rho=0.55$	0.0124	0.0154	0.0162	0.0164	0.0164	0.0164	0.0164
$\rho=0.65$	0.0341	0.0419	0.0437	0.0441	0.0441	0.0442	0.0442
$\rho=0.75$	0.0469	0.0557	0.0573	0.0576	0.0577	0.0577	0.0577
$\rho=0.85$	0.0446	0.0503	0.0510	0.0511	0.0511	0.0511	0.0511
$\rho=0.95$	0.0214	0.0224	0.0224	0.0224	0.0224	0.0224	0.0224

Table 1.1: Minimal values of ε from equation (1.9).

followers by initiating an early cascade and preventing information aggregation, it is also the case that this is true for only marginally well informed leaders. With a population of followers with the most unfortunate precisions with respect to this problem, a leader need only have a precision that is less than 6 percent more accurate than the followers. Generally, then, for leaders that are even moderately better informed than their followers, all followers will benefit.

In the same manner as in the above derivation for a relatively well-informed leader, it can be shown that a relatively poorly-informed leader can benefit some followers (when compared to a commonly informed leader). In light of the earlier discussion, this should not be too surprising. The general result is that well informed leaders inhibit the process of information aggregation by providing information of such quality as to cause an eventual follower to ignore his private signal and instigate a cascade. The point of this section is to demonstrate, that while a well informed leader can harm most followers by inhibiting information aggregation, the additional information that such a leader provides need not be terribly precise (relative to the followers) in order to overcome this problem.

1.4.2 Multiple Leaders

Rather than a single leader, consider a model with N leaders $\mathcal{L}_1 \ldots \mathcal{L}_N$ with each leader $\mathcal{L}_j$ observing a private signal $S_{\mathcal{L}_j}$. All leaders move simultaneously.[6] Even when these leaders have no informational advantage over the followers, followers can benefit from

[6]When leaders do not move simultaneously, the resulting process is just like that with a single leader with the remaining leaders acting as the first few followers.

the presence of multiple leaders. With two leaders, the first follower, $\mathcal{F}_1$, will be able to infer the signals $S_{\mathcal{L}_1}$ and $S_{\mathcal{L}_2}$. Thus, each follower will benefit from the presence of multiple leaders whether these leaders possess higher quality signals or not.

Leaders may have differing precisions. When the leaders choose actions based upon their signals, followers will be able to infer the signals and will incorporate this information into their decision. When a given player faces a set of signals with differing precisions, any number of outcomes can arise. For example, consider one leader with signal precision $\alpha = \beta = \rho + \delta$ and two other leaders with precision $\alpha = \beta = \rho$. If the first leader's signal can be inferred to be L and the other two leaders' precisions can be inferred to be H, then an observer can conclude the higher precision leader to be correct if his precision exceeds the probability that the other two leaders are correct. This is true when $\rho + \delta > \rho^2/(\rho^2 + (1-\rho)^2)$. When ρ is 3/4, the higher precision leader would need a precision of 9/10 or greater to be more precise than the concurring signals of the other two leaders. Generally, leaders with greater precision will be more likely to influence followers, though, as in this example, a concurrence of less precise leaders can overcome the more precise leader's influence.

The possibility does then exist that multiple leaders may not influence play in the same manner as would a single leader with the same precision. For example, in the earlier example, a single leader with precisions $\alpha = \beta = \rho + \varepsilon$ with followers that all have signal precision $\alpha = \beta = \rho$ always causes a cascade to occur. Two leaders with precision $\rho + \varepsilon$ and the same followers will not impart any information on the first follower when their signals do not concur. With multiple leaders, the proper prior expectation is that more information will be revealed and therefore will generally benefit the population of followers. However, for some sets of signal realizations, it may be the case the leaders fail to benefit the followers as much as would a single leader. In any event, prior to play, a leader with greater signal precision can be expected to have more impact on the play of followers than a leader with a less precise signal.

1.4.3 Leader Reputation and Selection

Each player prefers the player immediately ahead of him to be as well informed as possible, hence, the first follower always wants the most informed leader(s) that he can find. Later followers may disagree with this as they benefit from the information provided by $\mathcal{F}_1$ when he is not able to rely on a well informed leader and is forced to act upon his own information – and necessarily reveal that information. The same is true for player $\mathcal{F}_2$ and all of the players that follow him. Even though almost all players may prefer a leader with signal precision less than that desired by player $\mathcal{F}_1$, player $\mathcal{F}_1$ will follow the most informed leader available as will each successive player in the queue. This need not result in a selected leader that is globally inferior, however. As was shown earlier, a leader may need only have precision modestly, but not trivially, higher than that of his followers to be preferred by all to a less informed leader.

It is likely, that in many circumstances, the precision of a leader's signal may not be public information. Followers will then rely upon their prior expectation of a leader's signal precision. This prior will be updated over time as a series of projects (and possible cascades) occurs. Implicit in the model is the notion that projects eventually succeed or fail. As this success or failure affects all who accept the project, it will be publicly observable and the reputations of leaders will be updated accordingly.

It can be argued that this uncertainty over a leader's signal precision can be extended to all players in the game. However, due to the informational cascades, followers in most cases are not able to infer the signal of all (or even a few) of the followers that precede them. Posterior expectations of the signal precisions of other followers will not necessarily approach the true values over time and will often be identical to the prior expectations. The result will then be that the most accurate posteriors will be associated with leader signal precisions and the least accurate will be associate with the signal precisions of other followers. In the limiting and likely case that leaders are sufficiently well informed to create cascades instantly by their actions, no information regarding the signal precision of followers will ever be revealed as no follower's signal can ever be inferred.

Presuming followers to be a fairly homogeneous group, not well informed of each others' signal precision, the ordering of followers is not very important for the results derived in this paper. If there are a few distinctly well informed followers that tend to play early in the queue and these followers are just as well or more informed than the leader(s), play will be affected by their presence at the beginning of the queue. If these followers are, in some events, absent or at a later queue position, then play will be affected. These followers may, due to their relatively high signal quality, break cascades and impart more information to later players. In such cases, however, other followers will have a clear incentive to postpone their own play until these well informed followers have played. The well informed followers will have less incentive to wait. The result will be well informed followers positioned at the beginning of the queue. In such cases, these well informed followers can be modeled as leaders. An argument similar to this is presented in Bikhchandani, Hirshleifer and Welch [4]. In a model of strategic delay in cascades, Zhang [20] finds a unique equilibrium where players with the most precise signals always move first, resulting in instant cascades.

In the remainder of the paper it will be presumed that leaders' signals are generally more precise than the signals of followers. As a result, followers will copy leaders and will always ignore their private information.

1.5 Project Pricing

In some circumstances, the price of the project may be chosen by a project producer. In such cases, the task of a project producer is to select project price in order to induce a sufficient number of market leaders to purchase the project and, as a result, cause the remainder of the market (followers) to purchase as well. Effectively, then, the project producer is selling the project to the leaders, with emphasis on selling to leaders that will be most influential to followers. As a result, leaders can be viewed as proxy for the market of all potential purchasers of the project, with each leader's weight in this proxy market an increasing function of his reputation among followers. To induce a highly reputed leader to purchase the project may be equally satisfactory to a producer as to induce many less reputed leaders to accept.

Consider a project producer that faces a market consisting of N homogeneous players that move simultaneously. The producer will choose a price that maximizes the price multiplied by the probability that each player accepts the project combined with the producer's reservation value for the project multiplied by the probability of project rejection.[7] The optimal price is then:

$$P^* = \arg\max_P \Pi NP + (1 - \Pi)NP_0,$$

where P^* is the optimal price, Π is the probability that the player will accept the project and P_0 is the producer's reservation value. When all players have the same precisions α and β, the project producer will either price at the low signal valuation of $\mathrm{E}[\tilde{V}|S = L]$ or at the high valuation of $\mathrm{E}[\tilde{V}|S = H]$. To price below the low valuation would gain the producer nothing as, at the low valuation, all players will accept the project. To price above the low valuation, yet below the high valuation, would also be pointless, as only players with high signals will accept the project, in which case, the high valuation point is the optimal price. Finally, a price above the high valuation point would cause all players to reject the project.

A player will receive a high signal with probability α for high quality projects and probability $1-\beta$ for low quality projects. As projects are evenly distributed over high and low quality, the probability that a player receives a high signal is then $(\alpha-\beta+1)/2 = \Pi$.

---

[7] I assume costless production without loss of generality.

The project producer will price at the high level when:

$$\Pi\,\mathrm{E}[V|S=H] + (1-\Pi)P_0 > \mathrm{E}[V|S=L]$$

$$\frac{\alpha\overline{V}_H + (1-\beta)\overline{V}_L}{2} + \frac{\beta-\alpha+1}{2}P_0 > \frac{(1-\alpha)\overline{V}_H + \beta\overline{V}_L}{1-\alpha+\beta}. \tag{1.11}$$

For notational convenience, the firm's reservation value will be written as a linear combination of low and high quality project values.

$$P_0 = \phi\overline{V}_H + (1-\phi)\overline{V}_L.$$

A value of zero for ϕ would represent a firm that values unsold projects at the expected value of low quality projects ($\overline{V}_L$). A value of one for ϕ represents a reservation value of $\overline{V}_H$. Of course, a firm may have reservation values outside of the $[\overline{V}_L, \overline{V}_H]$ range. In particular, reservation values below $\overline{V}_L$ would not be unreasonable. As P_0 is linear in ϕ, values of ϕ below zero or greater than one are perfectly acceptable in this formulation of P_0.

When players have symmetrical signal precisions such that $\alpha = \beta = \rho$, the relationship in (1.11) is such that project producers select the high price when $\rho > (2-\phi)/3$. Thus, for sufficiently well informed markets (sufficiently large values of ρ), the project producer will price at the high level. Also, a larger reservation value (ϕ) for unsold projects will cause a project producer to be more likely to price at the high level.

Proposition 1.1. *The equilibrium project price is greater than the prior expected value of the project for well informed markets and less than the prior expected project value for poorly informed markets.*

Consider a market of perfectly aggressive players with $\alpha = 1$ and $\beta = 1/2$. When these values are substituted into (1.11), a project producer will price at the high level when $\phi > -2$. When facing a market of perfectly tentative players with $\alpha = 1/2$ and $\beta = 1$, a project producer will price at the high level if $\phi > 1/9$.

Proposition 1.2. *A project producer is more likely to price at the high level when facing a relatively aggressive market than when facing a relatively tentative market.*

1.6 Pricing Sequential Offerings

Well informed leaders possess a great deal of power over relatively less informed followers. Any agent that wishes to control the action of these followers will find a clear point of leverage in manipulating their leaders. There are two general ways to do this. The first is simply to convince the leaders to change their action in some way (via a side payment or other incentive). The second is indirectly to affect leader selection (and reputation) by selecting project price.

The task of a project producer is to price the project such that a sufficient number of influential leaders will choose to accept the project and thus induce the market of followers to accept the project as well. Clearly, then, a project producer prefers to face a market where influential leaders are more likely to accept the offered project. A project producer prefers influential leaders to be *aggressive*. Aggressive leaders are more likely to receive high signals, and thus, for leaders with precisions such that signals are relevant, are more likely to accept a given project than are tentative leaders. Consider what happens when a project producer faces a tentative leader. The producer can either price above the tentative leader's low-signal valuation and the tentative leader will accept the project only if he receives a high signal, or, the producer can price at the low-signal valuation and the tentative leader will accept for certain. Since the tentative leader will receive a high signal with probability less than 1/2, the project producer must price at the low level or, more likely than not, the tentative leader will reject the project. When facing an influential aggressive leader, the project producer can price above the aggressive leader's low-signal valuation and with probability greater than 1/2, the aggressive leader will receive a high signal and accept the project.

From Theorem 1.6, as project price falls, aggressive leaders are increasingly more likely to be correct in project selection relative to tentative leaders. Imagine a project producer with an optimal one-time project offering price. If this producer anticipates facing the same market some time in the future, he has a marginal incentive to lower his one-time offering price and marginally favor the reputations of more aggressive market leaders. As a result, this producer will face a more favorable market in the future. In a market of more than one project producer, this marginal incentive to decrease price will be a function of the probability that the project producer will be able to extract the benefits of a more aggressive market in the future. From the point of view of project producers, an aggressive leadership structure is a public good. Thus, the usual concept of under-provision will apply.

Proposition 1.3. *A multiple-project producer has an incentive to price below the one-time offering price in order to create a more aggressive market leadership for future projects. This incentive will be greater for markets with fewer project producers or for project producers that offer greater numbers of projects, relative to rival producers, over time.*

From Proposition 1.2, when facing a leadership structure that is already fairly aggressive, a project producer will be more likely to have a high one-time offer price. Similarly, a tentative leadership structure is more likely to favor a low one-time offer price. Finally, when the leadership structure is tentative, and a project producer expects a more aggressive leadership structure in the future, he has an incentive to postpone the project offering.

Proposition 1.4. *More projects will be offered during periods of relatively aggressive market leadership and project prices will be higher than during periods of relatively tentative market leadership.*

Evaluating a given project's success or failure (upon realization of project payoff, V) will take some amount of time. When this time lag is longer than the frequency of new project appearances, leader reputations will tend to be "sticky" in response to recently offered projects.

Proposition 1.5. *When the time to the realization of project value exceeds the time between new project offerings, multiple projects can be offered before the success or failure of any one of the projects affects leader reputations.*

A result of Proposition 1.5 is that there can exist extended periods of aggressive and tentative market behavior which is unaffected by the price and quality of recent projects. This can conceivably lead to cyclical market behavior. An aggressive market will tend to stay aggressive even when in the midst of a series of higher-priced projects. When the success of this series of projects can be evaluated, leader reputations will respond to favor more tentative leaders. These tentative leaders will tend to inhibit the production of new projects as well as induce producers to offer lower-priced projects. This will result in greater reputations for aggressive leaders when this series of projects is evaluated. Leadership structure can also change as new leaders enter the market or existing leaders exit. With the possibility of entry and exit alone, market leadership cannot be expected to be constant over time.

1.7 Applications

1.7.1 Initial Public Offerings

The model of leader manipulation closely mirrors observable phenomena in the market for initial public offerings. Underpricing of new issues and hot issue markets are consistent with the model's predictions. In the language of the model, project producers can be thought of as investment bankers or other agents representing issuing firms. Market leaders can be viewed as market analysts, industry analysts or investors that tend to be well informed, such as institutional investors or fund managers. The project payoff represents the short-run return on a new issue. The project price is the offering price.

It is well known that IPOs are generally under-priced. This is reflected in short-run positive price changes following offering periods. Ibbotson and Ritter [12] cite evidence of underpricing of new issues in the U.S. market at a rate of about 15%. The model in this paper offers a new explanation of IPO underpricing. While an issuer has no incentive to manipulate market leadership by underpricing[8], the issuer's agent (the investment banker) will be facing the market again when it offers another issuer's IPO. One reason to under-price an offering is to create a market of more aggressive leaders for future offerings.

Empirical evidence surrounding underpricing is consistent with the model in this paper. Ibbotson and Ritter cite greater levels of underpricing for smaller offerings. This can be explained by less informed leaders for such offerings[9] as well as fewer investment bankers populating niche IPO markets. When fewer investment bankers are present (or a few investment bankers handle most IPOs), the incentive to under-price in order to influence future market leadership is greater. As smaller IPOs tend to be offered to such markets, one would expect them to be more severely under-priced.

The concept of hot issue markets (markets when IPOs are most severely under-priced and thus "hot" buys for investors) was observed by Ibbotson and Jaffe [11] in their 1975 paper. A simple explanation of hot issue markets is that these markets reflect periods when market leadership is relatively tentative. An empirical implication is that hot-issue markets should see relatively fewer offerings over time. IPO volume should then be greater following a hot-issue market (as aggressive leader reputations rise). As leaders become more aggressive in response to a hot issue market, IPO volume should increase and underpricing should become less severe.

This explanation of hot-issue markets is not dependent upon whether investment bankers choose to price issues in order to manipulate market leaders. Such pricing may occur, but in any case, some price will be chosen (for whatever reason), and, as a result, leader reputations will be affected. Thus, for example, a period of relatively severe underpricing will favor a more aggressive market leadership in the future, regardless of the reason for this underpricing. One implication of this explanation of hot-issue markets

[8]This is not entirely the case. A new issuer may plan a seasoned offering at a future date. Welch [17] shows that issuers may prefer to under-price in order to improve the chances of success of a subsequent offering.

[9]Beatty and Ritter [2] offer a similar explanation.

is that subsets of the over-all market may experience hot-issue markets at different times. Thus, for example, the market for high-technology issues (with specialized analysts) may experience a hot-issue market at the same time that the market for high-fashion clothing issues (with it's own analysts) is cooling off.

1.7.2 Capital Investment

Upon the realization of new technologies or expectations of a change in demand, firms may have an incentive to modify their production techniques or otherwise alter the structure of the firm in response. In industries where there are leading firms (for example, larger or "cutting edge" firms), these firms will develop reputations as they respond to changes in technology and demand. In the language of the model, the project represents the choice of whether to accept a new technology or otherwise modify firm structure. The price is the relative cost of this change when compared to the potential benefit. During periods when demand is particularly high or technological improvement relatively cheap, the relative cost of undertaking a change in firm structure will be low. This will favor the reputations of more aggressive firms.

It is likely that a change in firm structure or technology will take some amount of time to be judged as a successful or failed project. When market conditions change and the benefit of technological improvement or other structural alteration falls relative to the cost, the high reputations of aggressive firms will impede the ability of the industry to respond to changing conditions. Likewise, when market conditions improve following a period of low demand or stagnant technological development, tentative firms will have developed higher reputations and will therefore slow the industry's response to new opportunities. An indicator of behavior such as this would be industries that are slow to react to new market opportunities, yet seem to over-react once industry changes begin.

As an example, a mining industry may experience a positive demand shock for a particular ore. At the same time, an aggressive industry leader happens to find a new, high quality source of this ore. It will take some time before other mining firms are able to verify that the new source is of high quality. When the firms are able to confirm the quality of the new find, they note the success of the aggressive industry leader. Later, when this leader purchases land in a new area, other mining firms purchase land in the same area, following the previously successful leader. They do this even as the demand for ore falls back to its previous level, causing the purchase of new land for mining to be less profitable. Eventually, firms note that neither they, nor the previously successful leader is doing as well as expected. The firms then begin paying more attention to industry leaders that recently chose not to open new mines. As a result, when another positive demand shock arises, the industry is slow to respond with increased exploration and new mines.

1.7.3 Cyclical Production

Much like the case for IPOs, firms in an industry may choose to price a new product at a low level on order to create an aggressive market for future products. As an example, consider the market for computer software. A software publisher may offer a new program at a low price in order to increase the reputations of aggressive market reviewers. A software reviewer for a computer industry magazine may review the new program and recommend the program to readers. If the program is priced at a low level, it will be more likely that purchasers will view the product as a success and thus increase the reputation of the reviewer. When the software company offers a new version of the software, the aggressive market reviewers will have greater control over the purchasing decisions of readers.

The same story can be told for just about any product market where leaders exist. Examples include the market for clothing fashions (where leaders are models and celebrities) and the market for automobiles (where consumer guidebooks lead purchasers to particular manufacturers). It need not be the case that firms set price in order to manipulate market leaders. Whatever the price, either aggressive or tentative leader reputations will be favored. For example, a monopolist that prices low to drive out a competitor not only remains a monopolist, but receive a bonus of a more aggressive market leadership structure for new products.

When the time lag in determining product quality exceeds the rate of new product offerings, reputations of leaders will be unresponsive to new products and cyclical patterns of aggressive and tentative leadership structure may emerge. For example, it may take a few years to determine the true quality and value of a new automobile. The auto industry could then face multi-year periods of aggressive market leadership (where influential consumer guides and industry analysts proclaim new autos generally to be well worth the purchase price) followed by periods of tentative market leadership (where industry analysts urge buyers to wait for next years' models). The result can be periods of low prices with few sales followed by higher prices with more sales. Interestingly, these periods need not be correlated across industries. Thus, one industry can experience a sales boom at the same time that another industry experience low sales volume.

1.8 Conclusion

Often, the same set of individuals or firms face sequences of distinct decisions and observe the past decisions of others. In this paper, informational cascades are modeled as elements of such a sequence of decision problems, with one potential cascade following another. Each event is the offering of a project with uncertain payoff. Players observe private signals of project quality and choose whether to accept the project or not based upon their private signals and information that can be inferred from the actions of earlier players. Leaders are the earliest players in each event. If multiple leaders exist, they all take action simultaneously. As a result, all leaders' signals can be inferred by the followers.

Each player has an incentive to observe all earlier players and extract as much information as possible prior to taking action. There is therefore an incentive for early followers to choose the most informed leaders. Leaders that are better informed than their followers will initiate instant cascades, where each follower follows the leader without regard to their own signal. It is possible that, by preventing information aggregation, this instant cascade effect may be harmful to later followers. It is shown that leaders that are sufficiently more informed than the followers will impart enough information on the followers to overcome the costs of less information aggregation. Furthermore, these sufficiently better informed leaders may only need a few percentage points greater signal precision over the followers. Multiple leaders can provide a great deal of information to followers and a more informed leader will have a greater impact on follower actions than will a less informed leader.

The precision of a given leader's information may not be public knowledge. In this event, as the leader's signal can always be inferred and projects will eventually succeed or fail, followers will be able to update their prior expectations of a leader's precision. This process is, however, open to manipulation by agents who offer the projects. If potential leaders vary in their ability to identify various project types, it is possible for project producers to influence a leader's reputation. This influence can affect the success of future projects. The model of leader reputations and sequential cascades is applied to a number of speculative investment decisions and associated aggregate patterns. These include underpricing and hot issue markets in IPOs, capital investment decisions and cyclical production patterns.

Chapter 2

Endogenous Local Interaction and Convergence in Learning Models

2.1 Introduction

This paper endogenizes opponent selection in a model of learning in a coordination game. Players are allowed to search for an opponent either among a global population consisting of all players or among a local subset of this population. In coordination games (with multiple Nash equilibria), the task of learning models is to refine further the set of Nash equilibria. This paper shows that the ability of players to (optionally) search for an opponent among subsets of the population leads to greater stability for the Pareto efficient equilibrium than would be the case when players must seek opponents either globally or locally.

Kandori, Mailath and Rob [13] and Young [19] both find that in multiple-Nash equilibria coordination games, when players play a strategy of myopic best response with possibility of mutation, the risk-dominant equilibrium is favored as the predictor of long-run play. The risk dominant equilibrium need not be efficient. Thus, simple learning models are shown to produce Pareto inferior outcomes in some circumstances. Ellison [7] examines rates of convergence for such models and demonstrates that under very general conditions, convergence rates can be so slow as to make these models poor predictors of equilibrium play over the time horizons typically of interest to economists. In particular, Ellison showed that existence of local interaction among players crucially affects convergence and, when players always choose opponents among local subsets of the population, convergence to the risk-dominant equilibrium is much quicker than when opponents are selected among the entire population.

Learning models, and, in this case, models of bounded rationality are of interest to economists to the extent that they both approximate expected behavior and do so in a time frame that is economically significant. To assert that one equilibrium is more stabie than another places heavy emphasis on the meaning of stability. If local play affects convergence, it is then important to understand not just how players may play the game in question, but also with whom they choose to play. In some cases it may be sufficient merely to assert that players will choose opponents (or partners) within a local neighborhood or that they may seek an opponent without regard to location. In the model that follows, this is not the case. There is no exogenously imposed strategy for opponent selection. Player strategy space is two-dimensional over opponent selection method and play in the game. The primary innovation of this paper is to study local interaction as an endogenous component of learning models.

Players are located at fixed positions and, in each time period, they choose (1) to seek an opponent either locally or globally and (2) the game strategy to play against the partner they find. The presumption of fixed location can be viewed in a couple of ways. First, and most simply, players may literally not be able to move from one location to another. This is particularly true when the metric of space is not literal distance, but involves another measure, such as education or wealth. But, when location is measured in the literal sense, the presumption of fixed locality is not necessarily unreasonable. While the model addresses players playing a single game, it may well be the case that players play a large set of diverse games over the course of a single time period. While it may be optimal to change location for the sake of a particular game, there will in fact be an optimal location that is almost certainly sub-optimal for any particular game. In this sense, players can easily be imagined as occupying a fixed location. In addition, the game itself may be played in a central location where the metric of location for the players is, for example, familiarity with other players. Examples of such an arrangement are security

		Column Player A	Column Player B
Row Player	A	7,7	8,0
	B	0,8	9,9

Figure 2.1: An example coordination game.

exchanges, trade meetings or farmer's markets. Finally, in models of biological evolution, local play is an important factor in speciation and, in this case, literal location is quite relevant. In cases where location is not fixed, Ely [8] has found equilibrium convergence rates quite different from those in the fixed-location models, with results that strongly favor Pareto efficient equilibria.

The other component of the game, selection of play, also deserves some justification. In games of coordination (which are the focus of this paper), players benefit from coordinating on their choice of play. In this case, why not just talk the matter over and each agree to make the same play? If the game is truly a coordination game, there is no incentive for either player to deviate. This line of reasoning is troubling, but observe the coordination game in Figure 2.1. The Row Player, while preferring coordination, always prefers the Column Player to play B. In this case, the Row Player will always promise to play B as well (to induce the column player to respond in kind), regardless of his actual choice of play. The Column Player has the same incentive structure (the game is symmetrical). Hence, when the two players meet, no meaningful signaling is possible. There is no point in either promising to play B, since both expect the other to say just that. Hence, there exist even coordination games where no credible signaling is possible. Of course, other games possess this same property for much more straight-forward reasons.

2.2 The Model

Players P indexed by i are located on a space such that there exists a boolean measure $L(P_i, P_j)$ with the property that $L(P_i, P_j) = L(P_j, P_i)$. If $L(P_i, P_j)$ is true, then players

	A	B
A	a, a	c, d
B	d, c	b, b

Figure 2.2: Coordination Game Payoff Matrix

P_i and P_j are said to be *local* to one another. The players are matched and play a coordination game with the payoffs as in Figure 2.2. Conditions for the game to be a coordination game are $a > d$ and $b > c$. Further, the Nash equilibrium (A, A) will be arbitrarily defined as the ***risk-dominant*** equilibrium. That is, under the assumption that one's opponent is randomizing evenly over strategies A and B, the response with the higher expected payoff is A. Hence, $a + c > b + d$. In each period of the game, the following events occur:

1. Each player, P_i, independently selects a strategy, s_i, consisting of two components:

    - Game strategy: A or B.
    - Location strategy: G (global) or L (local).

    This produces the strategy set $S = \{AG, AL, BG, BL\}$ with $s_i \in S$.

2. Players seek opponents either locally (if they chose strategy component L) or globally (if they chose strategy component G).

3. Players play their game strategy component (A or B) and receive payoffs based upon their play and that of their opponent.

Each player chooses his strategy s_i myopically. That is, when choosing a strategy for the next period, a player chooses the strategy that is the optimal response to the strategies of all other players in the previous period under the implicit assumption that other players will continue to use their previous strategy in the next round of play. In this sense, this is a model consisting of boundedly rational players where each player can be said to learn, over time, as he chooses his strategy to best respond to the previous strategies of his opponents.

At some point, matching will become important. If a player chooses local play, then he presumably does so under the impression that this will yield a different distribution of opponents than under global play. But, if all of the members of the player's local group themselves choose global play, then there is no clear reason to expect a local match for the original player. Similar logic applies to players choosing global play in the face of a population of largely local players. This issue can, however, easily be avoided by clearly stating that the relevant probabilities are such that they indicate the distribution of the players that a local or global search may yield and not necessarily the distribution of the players in the local or global group.

Finally, players "mutate" away from this myopic best response with some small probability. Following Ellison [7], I'll denote this probability by 2ε, where players choose the optimal myopic best response with probability $1-2\varepsilon$. With probability ε players choose, with equal probability, AL or AG. Similarly, with probability ε, players randomize over BL and BG.

2.3 Model Solution

In order to characterize each player's solution to this myopic best response problem, consider a player that is known to play A, regardless of whether he plays locally (L) or globally (G)[1]. Let P_{AL} indicate the probability that an opponent, when chosen with the local strategy L, played strategy A in the previous period with similar notation for P_{BL},

---

[1] $s_i \in \{AL, AG\}$.

	$a > c, b < d$	$a > c, b > d$	$a < c, b > d$
$P_{AL} > P_{AG}$	$s_i \in \{AL, BL\}$	$s_i \in \{AL, BG\}$	$s_i \in \{AG, BG\}$
$P_{AL} < P_{AG}$	$s_i \in \{AG, BG\}$	$s_i \in \{AG, BL\}$	$s_i \in \{AL, BL\}$

Table 2.1: Strategy Subsets in the Coordination Game

P_{AG} and P_{BG}[2]. The player will then choose strategy AL over AG if and only if:

$$E[AL] > E[AG]$$

$$aP_{AL} + cP_{BL} > aP_{AG} + cP_{BG}$$

$$(a-c)P_{AL} > (a-c)P_{AG}$$

And, similarly:

$$E[BL] > E[BG]$$

$$(b-d)P_{AL} < (b-d)P_{AG}$$

To find s_i, one now needs simply to consider the sign of $(a-c)$, the sign of $(b-d)$ and the larger of P_{AL} and P_{AG}. Based upon the parameters of the coordination game, there are three general cases, each with two sub-cases[3] as in Table 2.1. The three cases described below divide the coordination game space into areas with similar equilibrium player behavior. In particular, this separation concentrates on a comparison between the risk-dominant equilibrium (A, A) and the other Nash equilibrium (B, B). Imagine a player facing some population of potential opponents in the upcoming coordination game, with equal probability of being matched with any player in the population. This player will then choose strategy A over B if, and only if, $aq + c(1-q) > dq + b(1-q)$, where q is the probability of being matched to a player that will play A. One way to

[2]By definition, $P_{BG} = 1 - P_{AG}$ and $P_{BL} = 1 - P_{AL}$.

[3]The case of $a < c, b < d$ is ruled out by the coordination game requirement of $a > d, b > c$ which directly leads to the fact that $a + b$ must exceed $c + d$, contrary to this case.

write this is to note the threshold value of q for which the player will switch from strategy B to A. This threshold, q^* is $(b-c)/(a-c+b-d)$. When q exceeds q^* the player will choose strategy A. To say that q^* is less than one-half is just another way of saying that A is risk-dominant. Put another way, if one is facing an opponent that randomizes evenly over $\{A, B\}$, the best response is to play the risk-dominant strategy.

2.3.1 Case 1: Significantly Efficient Risk Dominance

In this case $a > c$ and $b < d$. Not only is (A, A) the risk-dominant equilibrium but it is also always the efficient equilibrium as well. With $b < d$ and the coordination game requirement of $a > d$, the payoff to coordinating on (A, A) exceeds the payoff to a coordination on (B, B). But there is more happening in this case than just the certain efficiency of the risk-dominant equilibrium. In order to see this, consider a player that has, for whatever reason, chosen to play either A or B in the game to come. This player can now choose to play either globally or locally, in effect, choose between two groups of potential opponents, one a subset of the other. In this case, the choice is easy: always choose the group with a higher probability of producing an opponent that will play A. Whatever the players choice of play, with $a > c$ and $d > b$, the payoff is higher when the opponent plays A.

From Table 2.1, this intuition coincides with two possible strategy subsets, $\{AL, BL\}$ and $\{AG, BG\}$. If P_{AL} exceeds P_{AG}, the proper strategy subset is $\{AL, BL\}$. The optimal strategy is to choose the higher expected value. In particular, choose AL if:

$$E[AL] > E[BL]$$

$$aP_{AL} + c(1 - P_{AL}) > dP_{AL} + b(1 - P_{AL})$$

$$P_{AL} > q^*$$

And, similarly, if P_{AL} is less than P_{AG}, choose AG over BG if P_{AG} is greater than q^*. This produces the optimal strategy space diagram in Figure 2.3.

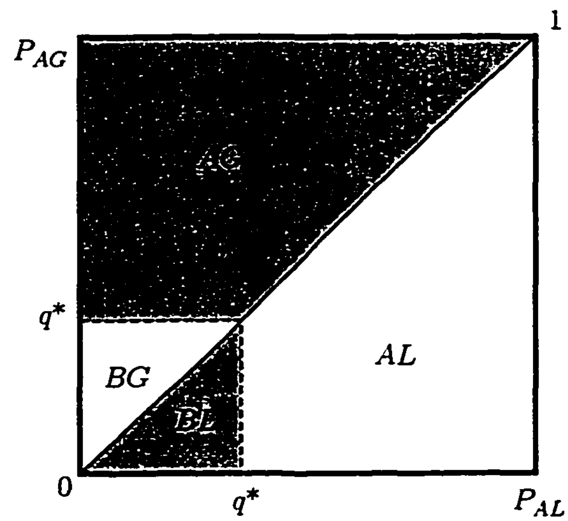


Figure 2.3: Optimal Strategy Space, Case 1: Significantly Efficient Risk Dominance ($a > c$ and $b < d$).

A player will play A whenever either his local group, the global group or both yield potential opponents that play A with a probability in excess of q^*. Compare this to the case where the player can only play globally. If a lower proportion than q^* of the population is expected play B, then the optimal response is B. But, with the option of local play, a player now requires his local group to play B in proportion less than q^* as well. Hence, in this case, the *option* of local play yields faster convergence to the risk-dominant equilibrium than one would expect under a system of only global play or only local play.

Proposition 2.1. *In a coordination game with payoffs as in Figure 2.2, (A, A) risk-dominant ($a + c > b + d$), and $a > c$, $b < d$, a myopic best response strategy with the option of local play will result in faster convergence to the (A, A) equilibrium than would be the case under the same game with either forced local play or forced global play.*

2.3.2 Case 2: Significantly Inefficient Risk Dominance

Here $a < c$ and $b > d$. While the previous case shows that the option of playing locally can only increase the rate of convergence to the risk-dominant equilibrium, this case has quite the opposite effect. The risk-dominant equilibrium is still (A, A), but, with $a < c$ and, by the coordination game requirement $b > c$, the payoff to the (B, B) equilibrium always exceeds the risk-dominant payoff. Completely opposite to the previous case, players now prefer to select opponents among groups that are most likely to yield opponents playing B, since, regardless of one's own strategy, the payoff is higher when faced with an opponent playing B.

From Table 2.1, the two possible strategy subsets are $\{AL, BL\}$ and $\{AG, BG\}$. If P_{AL} exceeds P_{AG}, the proper strategy subset is $\{AG, BG\}$.[4] The optimal strategy is to choose the higher expected value. In particular, choose AG if:

$$E[AG] > E[BG]$$

$$aP_{AG} + c(1 - P_{AG}) > dP_{AG} + b(1 - P_{AG})$$

$$P_{AG} > q^*$$

And, similarly, if P_{AL} is less than P_{AG}, choose AL over BL if P_{AL} is greater than q^*. This produces the optimal strategy space diagram in Figure 2.4.

A player will play B whenever either either his local group or the global group can be expected to produce B players with probability greater than $(1-q^*)$. As before, compare this to the case where players can only seek opponents among the global group. With only one group from which to choose opponents, players simply play A if the probability of getting an opponent that will play A is greater than q^* — otherwise, play B. With the option of playing locally, the chances of a player choosing to play A can only fall. Now, to play A, not only does the player require the global group to produce "A" players at a proportion greater than q^*, but he also require the same of the local group. Otherwise, the optimal play is B. The over-all result is precisely the opposite of the previous case. Convergence to the risk-dominant equilibrium, if it occurs, proceeds at a slower rate if players are allowed the *option* of local play.

---

[4]Note that this is the opposite of the previous case where the strategy subset was $\{AL, BL\}$.

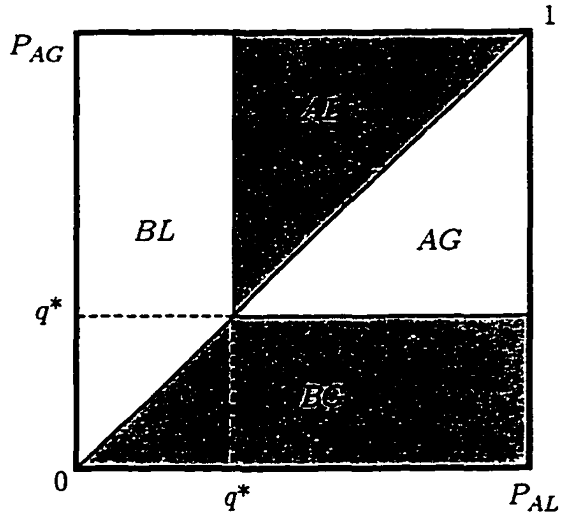


Figure 2.4: Optimal Strategy Space, Case 2: Significantly Inefficient Risk Dominance ($a < c$ and $b > d$).

Proposition 2.2. *In a coordination game with payoffs as in Figure 2.2,* (A, A) *risk-dominant* $(a + c > b + d)$*, and* $a < c$*,* $b > d$*, a myopic best response strategy with the option of local play will result in slower convergence to the* (A, A) *equilibrium than would be the case under the same game with either forced local play or forced global play.*

In this case, with players seeking a "B" dominated group of opponents and (B, B) the efficient Nash equilibrium, local play acts to impede convergence to the risk-dominant equilibrium. The next, and final case of the three, is intermediate between the two thus far presented. Hence, at one end of the spectrum is the option of local play increasing convergence to the risk-dominant equilibrium and at the other is optional local play doing just the opposite. But, in both cases, the option of local play acts to guide play away from the inefficient equilibrium and toward the efficient.

2.3.3 Case 3: The Intermediate Case

Here $a > c$ and $b > d$. As an intermediary between the two previous cases, this case is least interesting in its predictions, but more interesting in the choice of optimal strategy. Here, neither equilibrium can universally be declared efficient and the general result is that as one equilibrium begins to dominate the other in terms of efficiency, the more

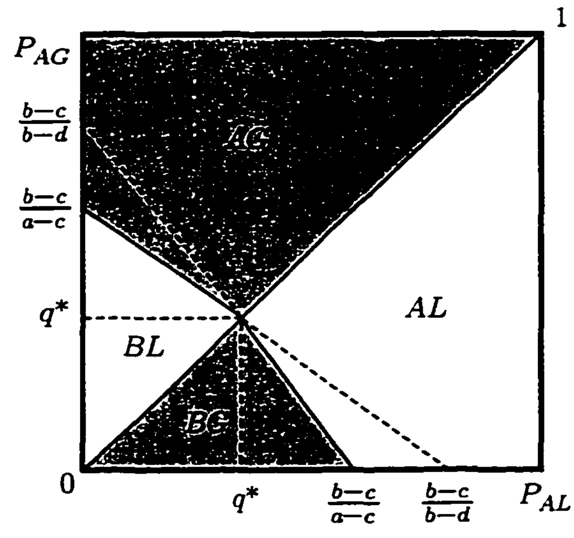


Figure 2.5: Optimal Strategy Space, Case 3: The Intermediate Case ($a > c$ and $b > d$).

the dynamics in the previous cases begin to favor the more efficient equilibrium. With $P_{AL} > P_{AG}$, Table 2.1 indicates that the proper strategy subset is $\{BG, AL\}$. As always, the optimal strategy has the higher expected value:

$$E[AL] > E[BG]$$

$$aP_{AL} + c(1 - P_{AL}) > b(1 - P_{AG}) + dP_{AG}$$

$$(a - c)P_{AL} + (b - d)P_{AG} > (b - c)$$

$$P_{AG} > \frac{b - c}{b - d} - \frac{a - c}{b - d}P_{AL}$$

This produces an indifference line that goes through (q^*, q^*) in (P_{AL}, P_{AG}) space. Similarly, when $P_{AL} < P_{AG}$, the strategy AG is preferred over BL when $P_{AG} > \frac{b-c}{a-c} - \frac{b-d}{a-c}P_{AL}$. This yields the optimal strategy space diagram in Figure 2.5.

As mentioned previously, the results are somewhat between those in the earlier cases. As b falls toward d, the BG and BL regions in Figure 2.5 shrink and approach those in

		$P_{AL}=0$	$P_{AL}=1/2$	$P_{AL}=1$
Case 1	$P_{AG}>q^*$	A	A	A
	$P_{AG}<q^*$	B	A	A
Case 2	$P_{AG}>q^*$	B	A	A
	$P_{AG}<q^*$	B	B	B

Table 2.2: Optimal Game Strategies in the 2 Neighbor Scenario

Figure 2.3.[5] Similarly, as a falls toward c, the AG and AL regions shrink and approach those in Figure 2.4. It is interesting to note that when the two Nash equilibrium payoffs equal one another ($a = b$), the A and B regions evenly split the intermediate regions that are consumed either all by A in Case 1 or all by B in Case 2.

Proposition 2.3. *In a coordination game with payoffs as in Figure 2.2, (A, A) risk-dominant ($a + c > b + d$), and $a > c$, $b > d$, a myopic best response strategy with the option of local play will result in an indeterminate effect on convergence to the (A, A) equilibrium when compared to the case under the same game without the option of local play.*

2.4 Example: The 2 Neighbor Game

In order to better illustrate the previous concepts, consider the game as described before with players arranged in a circle. Let players be considered local to one another if they are side-by-side. Each player will then have two neighbors, one on the left and one on the right. Using the above results, one can easily derive the optimal player strategies for each of the three cases. As Case 3 is intermediate, only Case 1 and 2 are shown in Table 2.2.

Consider a situation where all players begin by playing B, either locally or globally. In Case 1, as soon as a player's neighbor (either or both) mutates to A, the player will respond by playing A. This will impact both of his neighbors and so on, until in a short time, most of the population will be playing A with only the occasional mutations to B.

[5] That is, $\frac{b-c}{a-c}$ approaches q^*.

The important thing to note in Case 1 is that the only way a player will ever play B is when both neighbors have done so in the previous round and a proportion greater than $1 - q^*$ (greater than at least one-half) of the population is playing B. In this case, most players playing B is very unstable.

In Case 2, with all players beginning by playing B, the only way anyone will play A is if either or both neighbors previously played A and a proportion greater than q^* of the population played A in the previous round. Hence, any single mutation to A will not spread among the population. With N players, what is required is q^*N simultaneous mutations to A — an event that is extremely unlikely as N gets large[6]. However, if this event did occur, and greater than a q^* proportion of the population began playing A, this would also be quite stable. The only way that a player would return to playing B would be if neither neighbor had played A in the previous round. A single B mutation could not spread among the population in this manner. The general result, in this case, is that both equilibria are quite stable, though moving from the A equilibrium to the B equilibrium is more likely (requiring fewer simultaneous mutations) than moving from B to A.

This example illustrates an important feature of Case 2. Though the B equilibrium is strengthened past the point of mere global play alone, it is not far past this point. The only manner in which the B equilibrium is stronger than in the case of simple global play is that if the proportion of local players playing A falls below q^*, then playing B locally is the optimal response, even if the global proportion of A players exceeds q^*. This is, however, an increasingly unlikely event as the number of players in one's local group increases. Hence, while this effect can make transition from a B equilibrium to an A equilibrium less likely, the effect is likely small compared to simple global play. In comparison to simple local play, however, the impact of optional local play in this case remains extreme.

[6]Actually, it would probably take more than q^*N mutations since, in the next round, players that still had both neighbors playing B would respond with B. If enough players were in this situation, the proportion of the population playing A may fall below q^*.

	A	B		A	B		A	B		A	B
A	10,10	2,6		10,10	11,8		10,10	0,2		10,10	7,0
B	6,2	4,4		8,11	12,12		2,0	4,4		0,7	12,12
	Case 1			Case 2			Case 3a			Case 3b	

Figure 2.6: Example Coordination Games ($q^* = 1/3$)

2.5 Simulation Results

For the purposes of simulation, some example game parameters need to be selected. In order to further the point that more than just the relative sizes of the two Nash equilibrium payoffs drive the results, the examples in Figure 2.6 all have the same values of a and q^*, as well as other similarities. A few assumptions need to be made before simulation is possible. First, players will be located about a circle, with a single neighbor on either side comprising one's local group. Players will also calculate P_{AL} and P_{AG} as if a match is assured, whether they choose to play locally or globally.

The results in Table 2.3 indicate simulated wait times for populations with all players initially playing B to converge to 75 percent or more playing A. Each population has 20 members and the mutation rate, ε, is 5 percent. Each value in Table 2.3 is the average wait time (until greater than 75% of the population is playing A) for 1000 separate simulations. There is still a fair amount of noise in the results, however. Note that values for forced local play and forced global play should be uniform across cases as optimal responses depend only upon q^*.

The results fit nicely with the theoretical model. In Case 1, convergence to the risk-dominant (and efficient) equilibrium is very rapid and exceeds the rate produced by forced local play alone. In Case 2, convergence to the risk-dominant (and inefficient) equilibrium takes a great deal of time and is even slower than that under forced global play (though not by much). In Case 3a, with the risk-dominant equilibrium also the efficient equilibrium, convergence is rapid, while in Case 3b, convergence slows considerably when the risk-dominant equilibrium is not efficient.

It is worth noting that to speak of convergence is technically accurate, but somewhat

	Case 1	Case 2	Case 3a	Case 3b
Endogenous Local Play	5.6	31951.5	6.1	2626.5
Forced Local Play	8.4	8.3	8.5	8.3
Forced Global Play	30251.6	28424.8	29649.3	29199.8

Table 2.3: Simulation Results

misleading. A more descriptive term would be something like "inverse failure rate". In none of the cases does one expect or observe a smooth path from one equilibrium to another. Rather, the equilibrium transitions are abrupt and generally occur at a time when a sufficient number of disruptive mutations arrive simultaneously. This abrupt switch from one equilibrium to another is the primary reason why the simulation results indicate such a large amount of noise even after 1000 simulations per data point.

2.6 Local Play as its Own End

To now, focus has remained on the issue of play in the eventual coordination game. The nature of opponent selection — local or global — is, however, an interesting issue in itself. Unfortunately, there is not much to conclude from the model. In all cases, equilibrium has the effect of stabilizing the magnitudes of P_{AL} and P_{AG}. Of course, the proportion of local players playing a particular strategy will have a higher variance over time than will the population as a whole, presuming local groups to be smaller than global. But, on average, one must expect the distribution of local play to match that of global play in equilibrium. This, however, has the unfortunate effect of making any prediction of local versus global strategy impossible. In terms of the optimal strategy diagrams, one can generally expect the (P_{AL}, P_{AG}) pair to spend a great deal of time upon or near the 45 degree line, riding the edge of local and global play.

The story here is really more about group selection (by players) than it is about location. In fact, had local play been referred to as "group 1" and global play as "group 2", all of the discussion of optimal strategy would have done just as well. The general result, of Case 1 is that when presented with a set of pools from which to draw an opponent, choose the pool with the greatest probability of yielding an A player. In Case

2, one rather seeks B players, and, in Case 3, the issue is unresolved. Couched in these terms, the results of the model are strengthened as the number of pools available to a player rises (presuming they are not, by design, perfectly correlated in their distribution of players). Hence, to the extent that one result of the model is to predict favor for the efficient equilibrium, a large number of groups from which to choose opponents can only strengthen this result.

In a larger sense, the model says something about group selection at a level above that of mere locality. Consider two distinct and separate collections of players, each playing a coordination game as described in the model. The first collection of players is, by accident if nothing else, constructed such that local play is very difficult to accomplish, while the second collection of players can easily manage dividing itself into sub-groups of many kinds. Which collection of players, over time, will generally thrive? The second, of course. And if this second collection of players were to come into contact with the first, one would expect the players in the second group to dominate over time. Local play, is, in this sense, like a release valve. When the collective happens upon an equilibrium set of actions that are inefficient, the option of local play allows members to step out of the group as a whole and attempt to play the efficient strategy set with like minded (or constructed) players. But, it is not local play in itself that makes this case, as forced local play is hardly better than forced global play. It is the *option* of local play (or any form of sub-group opponent selection) that allows players to approach an efficient strategy set in the face of a collective that has chosen to do otherwise.

2.7 Conclusion

Models of coordination games with populations of players in fixed space indicate that the risk-dominant Nash equilibrium is, over time, the equilibrium to which play is most likely to converge. The rate of convergence, however, was shown by Ellison [7] to be very slow in cases where players were forced to play in large neighborhoods or among the entire population. Local play, then, could be seen as increasing the rate at which play would converge to the risk-dominant equilibrium. This paper endogenizes local play and examines the impact on equilibrium stability.

When local play is *optional* and not universally imposed, it emerges not as an accelerator to the risk-dominant equilibrium, but rather as a stabilizer of the efficient equilibrium, whether risk-dominant or not. The risk-dominant equilibrium remains as the more stable equilibrium, but in cases where this equilibrium is inefficient, the option of local play results in greater stability for the efficient equilibrium.

The ability of players to choose among various groups of potential opponents can only strengthen these results. This flexibility in opponent selection allows players a greater chance to escape from an inefficient global equilibrium. A general result is that populations with an ability to exhibit local behavior are better able to adapt to environmental changes that result in inefficient, population-wide, behavior.

Chapter 3

Misfits and Social Progress

This chapter is co-authored with Professor David Hirshleifer, School of Business Administration, University of Michigan.

> The game of history is usually played by the best and the worst over the heads of the majority in the middle.
>
> — Eric Hoffer, *The True Believer*

3.1 Introduction

Misfits, strangers, outcasts, and the young are sometimes the triggers of great social change. Jobs and Wozniak succeeded in the garage at an enterprise, the mass-marketing of the personal computer, that IBM with its vast expertise and resources failed to embrace. In the old story it was the little child who cried, "The emperor has no clothes!," oblivious to the mistaken conformity of the adults. Teenagers often flout social conventions. Often immigrants or strangers can trigger change in a community by trying things which would not have been considered by the locals.

Consider a community where the locals have converged on a behavior which, from each individual's point of view, seems best given their information. This includes the information that others are adopting the conventional behavior. It could be argued that misfits, strangers and children should be the least able to effect social change, because they have the least credibility. We argue here that the reason these groups are important innovators is their willingness to try new things despite opposing evidence. The new things they try are usually wrong, but they occasionally lead to improvements for everyone which otherwise would not have been realized. Thus, misfits can be disproportionately beneficial to society while often acting to their own detriment.

We begin with a generalization of the informational cascade model of Bikhchandani, Hirshleifer and Welch (BHW) [4] to allow for the possibility that individuals have no private information (see also Banerjee [1]). In this setting we examine an issue not focused on in previous papers, the level and determinants of the ultimate precision of the public pool of information arrived at in informational cascades. The paper then takes the further departure of introducing heterogenous individuals in the form of misfits, to examine the welfare effects of different kinds of misfits on society (see also Bernardo and Welch [3], who examine the welfare effects of one kind of misfit, overconfident individuals).

We find that the ultimate level of public information is largely independent of the fraction of individuals that are privately informed and that misfits serve a central role in extending social knowledge beyond the level one would expect in their absence. We also examine the conditions under which misfits would be welcomed or ostracized by the other members of society. Particularly, we find that societies with a relatively small number of members or a low proportion of privately informed members are more likely to exhibit hostility to newcomers.

The remainder of the paper is structured as follows. A model with the possibility of uninformed individuals is examined in Section 3.2 followed by the introduction of social misfits in Section 3.3. In Section 3.4 we examine the effects of misfits when individual knowledge is certified by an outside body or partially hidden from individuals themselves. Section 3.5 concludes.

3.2 The Basic Model without Misfits

We begin with a basic model that extends a binary version of the cascades model of BHW to allow for individuals who do not receive information. Individuals act in sequence and each has the opportunity to accept or reject a project with uncertain payoff. Without loss of generality, the project net payoff, V, is known to have either a high value of 1 or a low value of -1. Ex ante, each outcome is equally likely and this is known to all. Individuals are identical risk neutral expected wealth maximizers and all that accept the project receive the same realization of V. All individuals that reject the project receive a certain payoff of zero.

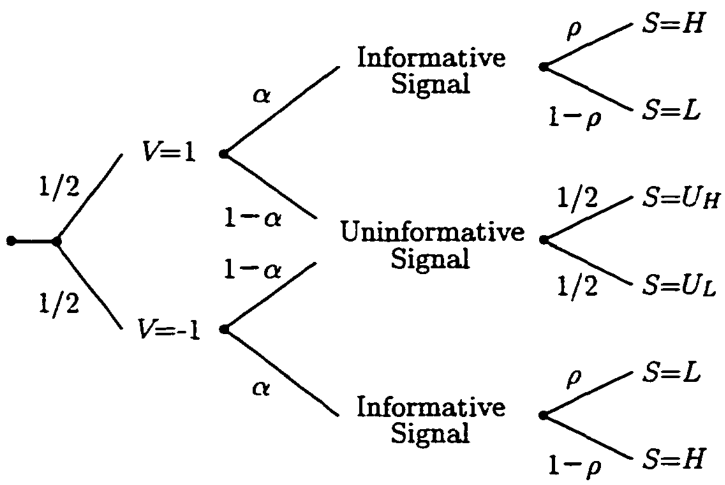
Each individual i receives a signal s_i which can be high (H), low (L) or uninformative (U). The probability α that an individual receives an informative signal ($s_i = H$ or $s_i = L$) is independent of project quality. If an individual does receive an informative signal, the signal is related to the project value as follows:

$$\Pr(s_i = H|V = 1) = \rho = \Pr(s_i = L|V = -1), \qquad 1/2 < \rho < 1.$$

Figure 3.1 illustrates the signal arrival process. The parameters ρ and α are the same for all individuals and will be referred to as the private informative signal *precision* and private informative signal *arrival rate*, respectively. Conditional on project value, the signals are independent across individuals. Individuals act in sequence where later individuals observe the actions, though not the signals, of predecessors. Later individuals are unable to observe whether a predecessor received an informative signal and, if so, whether the signal was high or low.

Ω_i is the set of all public information available to the ith individual. This information in addition to the private signal s_i will contain all available project information relevant to the ith individual at the time of his action. When $\mathrm{E}[V|\Omega \cup \{s\}] > 0$, individuals will always accept and when $\mathrm{E}[V|\Omega \cup \{s\}] < 0$ will reject the project.[1] There

[1] It will be convenient to suppress i subscripts throughout the analysis that follows.



Informative Signal

	H	L
$V=1$	ρ	$1-\rho$
$V=-1$	$1-\rho$	ρ

Uninformative Signal

	U_H	U_L
$V=1$	1/2	1/2
$V=-1$	1/2	1/2

Figure 3.1: Diagram of an individual's signal arrival process. Values along branches and in tables indicate probabilities.

are two situations under which $\mathrm{E}[V|\Omega \cup \{s\}] = 0$ and indifference arises. In cases where an individual's private informative signal (H or L) exactly offsets publicly available information, resulting in indifference, the individual will always follow his private signal. When public information is neutral and an individual receives an uninformative private signal (U), the individual will randomize evenly over accepting and rejecting the project. One could choose almost any convention to resolve individual uncertainty without significantly affecting the results that follow. The conventions chosen here ease the presentation by providing symmetry (when randomizing under an uninformative signal) and minimal signal extraction notation (in the case of following an informative signal under uncertainty).

As a notational convenience, randomizing over an uninformative signal (with otherwise neutral public information) will be explicitly modeled by specifying two types of uninformative signals, U_H and U_L. These signals contain no information about project quality. In cases where an individual would otherwise be indifferent, our convention is that an uninformative signal of U_H results in project acceptance while U_L will result in rejection.

In summary, individuals receive informative (H or L) signals with probability α as described above. Individuals receive uninformative signals with probability $1-\alpha$ and, given an uninformative signal, U_H and U_L are equally likely.

3.2.1 Model Solution

An individual, prior to receiving a private signal and facing a public information set Ω, will be inclined to accept the project if the expected payoff is positive, $\mathrm{E}[V|\Omega] > 0$ or $\Pr(V = 1|\Omega) > 1/2$. The probability that project payoff is high, given the public information set Ω, can be found via Bayes' Theorem:

$$\begin{aligned}\Pr(V=1|\Omega) &= \Pr(\Omega|V=1)\Pr(V=1)/\Pr(\Omega),\\ &= \frac{\Pr(\Omega|V=1)\Pr(V=1)}{\Pr(\Omega|V=1)\Pr(V=1)+\Pr(\Omega|V=-1)\Pr(V=-1)},\\ &= \frac{1}{1+X(\Omega)},\end{aligned}$$

where

$$X(\Omega) \equiv \frac{\Pr(\Omega|V=-1)}{\Pr(\Omega|V=1)}. \tag{3.1}$$

The likelihood ratio function $X(\Omega)$ is non-negative and equal to 1 when Ω is just as likely to have occurred under either project realization. In order to understand the public information set Ω that an individual may face and find an expression for $X(\Omega)$, it is necessary to examine the behavior of individual individuals in and out of an informational cascade. BHW define an *informational cascade* as a situation in which an individual's action choice is not affected by his private signal, regardless of its value. If an individual is in a cascade and if all individuals are identical (apart from signal realizations), all remaining individuals will simply follow the action of the individual that initiated the cascade. An individual who is not in a cascade will follow the balance of public information unless he receives a contradictory (informative) private signal.

Individuals can summarize the information revealed by predecessors as follows. Any predecessors who act in opposition to the action implied by the public information pool at the time of his decision must have received an informative, contradictory private signal. Individuals that act in agreement with the public information pool must have received an uninformative signal or a concurring informative signal. When public information suggests indifference, individual actions will correspond to either informative signals or one of the two uninformative signals, with acceptance suggesting a signal of H or U_H and rejection suggesting L or U_L.

Components of the public information pool Ω will consist of the actions (e.g., project choices) of predecessors, denoted by A for 'accepts' and R for 'rejects'. A subscript of A, R, or U denotes the action implied by the public information pool. When public information (reflected in predecessors' actions) supports project acceptance, an individual action is therefore denoted A_A or R_A. Similarly, an R subscript denotes that public information supports rejection, and a U subscript that public information implies indifference. When a cascade is in progress, the individual's actions will be denoted with a 'c' superscript as either A_A^c or R_R^c.[2]

For example, assume the first 5 individuals reject the project. If the first individual faced no (or indifferent) public information, his action will create a public information set of $\Omega = \{R_U\}$. Public information will now favor rejection and after the second individual rejects, the public information set will be $\Omega = \{R_U, R_R\}$. The rejection of the third individual will result in $\Omega = \{R_U, R_R, R_R\}$. If this is sufficient public information to initiate a rejection cascade then the actions of the next 2 individuals will produce $\Omega = \{R_U, R_R, R_R, R_R^c, R_R^c\}$. Because the order of elements in Ω will prove to be irrelevant it will be convenient to indicate the *number* of each action type in

[2] A_U^c and R_U^c are not possible as there cannot be a cascade when public information is neutral. The actions A_R^c and R_A^c also are not possible at this point as no individual would ever oppose a cascade. Deviations from cascades by 'misfit' individuals will be possible later in the paper.

R_A	A_R	R_U	A_U	R_R	A_A	R^c_R, A^c_A
$\{L\}$	$\{H\}$	$\{L, U_L\}$	$\{H, U_H\}$	$\{L, U_H, U_L\}$	$\{H, U_H, U_L\}$	$\{H, L, U_H, U_L\}$

Table 3.1: Actions and corresponding possible signals.

Ω by $\mathcal{R}$ and $\mathcal{A}$ notation. For this example, an equivalent way to write Ω would be $\Omega = \{\mathcal{R}_U = 1, \mathcal{R}_R = 2, \mathcal{R}^c_R = 2\}$.

For each element of Ω one can identify the set of possible signals that each predecessor must have received. The possible signals that correspond to each relevant element of Ω are in Table 3.1. Calculating $X(\Omega)$ using (3.1) and the above notation for Ω gives:

$$X(\Omega) = \left[\frac{\rho}{1-\rho}\right]^{(\mathcal{R}_A - \mathcal{A}_R)} \left[\frac{\alpha\rho + (1-\alpha)/2}{\alpha(1-\rho) + (1-\alpha)/2}\right]^{(\mathcal{R}_U - \mathcal{A}_U)} \left[\frac{\alpha\rho + (1-\alpha)}{\alpha(1-\rho) + (1-\alpha)}\right]^{(\mathcal{R}_R - \mathcal{A}_A)}. \tag{3.2}$$

Note that actions taken during a cascade (R^c_R and A^c_A) are uninformative to observers (all private signals result in the same action) and therefore are not relevant for the calculation of $X(\Omega)$. In other words, the fraction that would be raised to the power $(\mathcal{R}^c_R - \mathcal{A}^c_A)$ in $X(\Omega)$ above is equal to 1.

Similar to the binary example of BHW, only the net number of like-quality signals matters. For example, knowing that 3 individuals received a low signal and 1 individual a high signal is just as informative as knowing that 10 received low signals and 8 received high signals. The difference ($\mathcal{R}_A - \mathcal{A}_R = 2$ here) is all that affects an individual's decision.

Equation (3.2) reveals that past actions need not be remembered to track future actions. After each individual's action, $X(\Omega)$ is multiplied by the appropriate factor at which point all relevant knowledge of past actions is contained in the scalar value $X(\Omega)$. Individuals only need to know the value of $X(\Omega)$ and update this value after each action. This provides a very convenient way to simulate cascades which is independent of the relative net project payoffs and prior probabilities of high and low realizations. A simulator needs only to update $X(\Omega)$ after each simulated individual acts. Thus memory requirements of simulators (or real individuals) need not grow with the total number of individual predecessors. All relevant knowledge of each past individual action can be summarized by a single scalar value.

3.2.2 When a Cascade Occurs

An individual with public information Ω and private signal s will accept the project when, from (3.1):

$$\frac{1}{1+X(\Omega \cup \{s\})} > 1/2,$$

$$X(\Omega \cup \{s\}) < 1.$$

An adoption cascade occurs when an individual will accept the project even if he receives a low signal. This happens when the public information derived from earlier actions plus the individual's own low signal will not result in project rejection. An individual with a low signal will effectively ignore his private signal and accept the project if:

$$X(\Omega \cup \{L\}) < 1, \quad \text{or}$$

$$X(\Omega) < \frac{1-\rho}{\rho}. \tag{3.3}$$

Similarly, a rejection cascade occurs when

$$X(\Omega) > \frac{\rho}{1-\rho}. \tag{3.4}$$

Since $X(\Omega \cup \{s\})$ is based on the addition of the individual's private signal to the public information set Ω, it can be written as $\frac{\rho}{1-\rho}X(\Omega)$ (when $s = L$) or $\frac{1-\rho}{\rho}X(\Omega)$ (when $s = H$). This can be seen by noting that private signals of H or L are like observing actions A_R or R_A in a predecessor and then simply applying equation (3.2). From equations (3.3) and (3.4), a cascade begins when the current public information, summarized by Ω and mapped to the non-negative real value $X(\Omega)$, is outside of the (necessarily non-empty) interval:

$$\frac{1-\rho}{\rho} \leq X(\Omega) \leq \frac{\rho}{1-\rho}, \qquad \text{(No cascade interval).} \tag{3.5}$$

When $X(\Omega)$ falls below this range, a cascade begins where all individuals accept the project and when $X(\Omega)$ rises above this interval, a cascade will begin where all individuals reject the project. The no cascade interval is notably independent of α implying that the cascade level of public information will not be directly connected to the fraction of the population that is informed.

While equation (3.5) expresses a lower limit on the amount of public information necessary to induce a cascade, there is similarly an upper limit on the precision of public information that can be expected when a cascade begins. When public information is near a cascade boundary, arbitrarily less precise than necessary to induce a cascade, the most informative private action possible that results in a breach of the cascade boundary will determine the upper bound on the precision of public information. The only individual actions that produce information in concurrence with public information and are possible when public information is arbitrarily near a cascade boundary are actions R_R and A_A.[3] Both cases are the same by symmetry. Consider the case of action R_R when public information is arbitrarily near the rejection cascade boundary. From (3.5), $X(\Omega)$ will be equal to $\rho/(1-\rho)$ and the information from action R_R, via (3.2), will produce a value of:

$$X(\Omega) = \left(\frac{\rho}{1-\rho}\right)\left(\frac{\alpha\rho + (1-\alpha)}{\alpha(1-\rho) + (1-\alpha)}\right)$$

This value of $X(\Omega)$ can be translated to a precision by observing that a signal of arbitrary precision ρ' (with a realization of L in this case) produces an X value of $\rho'/(1-\rho')$. We

[3]Actions R_A and A_R will move public information away from the boundary, but not so much that a cascade over the opposite boundary will occur. Actions R_U and A_U are not applicable (none of these actions, R_A, A_R, R_U, or A_U, can ever initiate a cascade).

	Values of ρ				
	55.0%	65.0%	75.0%	85.0%	95.0%
$\alpha = 100\%$	59.9%	77.5%	90.0%	97.0%	99.7%
$\alpha = 70\%$	57.7%	72.0%	83.9%	92.6%	98.2%
$\alpha = 40\%$	56.2%	68.3%	79.4%	89.0%	96.8%
$\alpha = 10\%$	55.3%	65.7%	76.0%	85.9%	95.4%

Table 3.2: Upper bounds of the precision of public information from equation (3.7).

equate this with the above result to find the upper bound on public information precision:

$$\frac{\rho'}{1-\rho'} = \left(\frac{\rho}{1-\rho}\right)\left(\frac{\alpha\rho + (1-\alpha)}{\alpha(1-\rho) + (1-\alpha)}\right)$$

$$\rho' = \frac{\alpha\rho^2 + (1-\alpha)\rho}{\alpha\rho^2 + \alpha(1-\rho)^2 + (1-\alpha)} \tag{3.6}$$

The level of public information during an informational cascade is then bounded above and below by equations (3.5) and (3.6):[4]

$$\text{Cascade Precision}(\Omega) \in \left(\rho, \frac{\alpha\rho^2 + (1-\alpha)\rho}{\alpha\rho^2 + \alpha(1-\rho)^2 + (1-\alpha)}\right] \tag{3.7}$$

These bounds are ***constants of social information aggregation***, the minimum and maximum achievable precision of public information during an informational cascade. The upper bound of social information aggregation can be shown to be increasing in both ρ and α, while, as noted earlier, the lower bound is simply equal to ρ and not at all affected by the probability that a given individual is privately informed.

Proposition 3.1. *The* constants of social information aggregation *are such that an increase in the fraction of informed individuals can increase the precision of public information that is possible, but will not affect the precision that is sufficient to initiate an informational cascade.*

This invariance of the ultimate precision of public information with respect to varying

[4]The lower bound of ρ follows directly from the derivation of (3.5) or by applying the method used to derive ρ' to a boundary of equation (3.5).

a parameter makes it harder to identify policies to improve social outcomes. Of course, it also provides a protection against bad policy choices. For other results on the invariance of the ultimate precision of public information in settings with costly information and limits on observability, see Cao and Hirshleifer [5].

Table 3.2 shows the upper bound of social information aggregation for various values of ρ and α. Notably, large increases in the fraction of informed individuals have relatively little impact on the upper bound of public information precision. This will be further explored later in the paper.

3.2.3 Initial Conformity

Consider a course of actions such that each individual, beginning with the first, acts in a like manner — either all accepting or all rejecting the project. This will require an appropriate series of private signals. At what point will this series of actions result in a cascade (where private signals become irrelevant)?

Cascades where all reject or all accept are mathematically symmetric, so we arbitrarily consider the case where all individuals choose to reject the project. Individual $N+1$ in such a series of individuals will have observed N predecessors who all rejected. $\mathcal{R}_U$ will be equal to 1 and $\mathcal{R}_R$ will equal $N-1$ with all other elements of Ω equal to zero. A rejection cascade will occur when N is sufficiently large. From equations (3.2) and (3.5), the cascade will occur when:

$$\frac{\alpha\rho+(1-\alpha)/2}{\alpha(1-\rho)+(1-\alpha)/2}\left[\frac{\alpha\rho+(1-\alpha)}{\alpha(1-\rho)+(1-\alpha)}\right]^{N-1} > \frac{\rho}{1-\rho}.$$

After some manipulation this yields

$$N > 1+\frac{\log\left(\frac{\rho}{1-\rho}\frac{\alpha(1-\rho)+(1-\alpha)/2}{\alpha\rho+(1-\alpha)/2}\right)}{\log\left(\frac{\alpha\rho+(1-\alpha)}{\alpha(1-\rho)+(1-\alpha)}\right)}. \tag{3.8}$$

Some values for this function are given in Table 3.3.

When $\alpha = 1$ (individuals always receive informative signals), cascades always occur

	Values of ρ				
	55%	65%	75%	85%	95%
$\alpha = 100\%$	2	2	2	2	2
$\alpha = 70\%$	3	3	3	3	4
$\alpha = 40\%$	5	5	5	6	7
$\alpha = 10\%$	20	20	21	24	32

Table 3.3: Minimum values of N from equation (3.8).

after 2 or more initial like-acting individuals. As the probability of receiving an informative signal falls (α falls), an individual will be more likely to follow an informative private signal (when available) that opposes past actions. Also, holding α constant, as informative signal precision rises (ρ rises), individuals are more likely to follow an informative private signal. Thus, lower values of the information arrival rate (α) and higher values of precision (ρ) result in stochastically slower cascade formation. When α is low and ρ is high it is more likely that an individual with an informative signal will be in a position to use it.

An initial series of conformity is not necessarily the fastest path to cascade formation. When individuals conform they reveal less about the value of their private signal than if they act to oppose previous actions (more precisely, to oppose the action suggested by public information). It can be shown that the series of actions A_U-R_A-R_R-R_R will always result in a rejection cascade for non-trivial values of α and ρ.[5]

Result 3.1. *When there is a chance that individuals are uninformed, action contradictory to that suggested by public information reveals more about an individual's private information than does conformity. In consequence, an initial sequence of like actions is not necessarily the quickest path to cascade formation.*

Contrarians are much more noticeable than followers of the herd. But apart from ease of detection, there is good reason to pay attention to contrarians. "He must know something," is a common description of someone that acts to contradict public knowl-

[5] $\alpha \in (0, 1], \rho \in (1/2, 1)$.

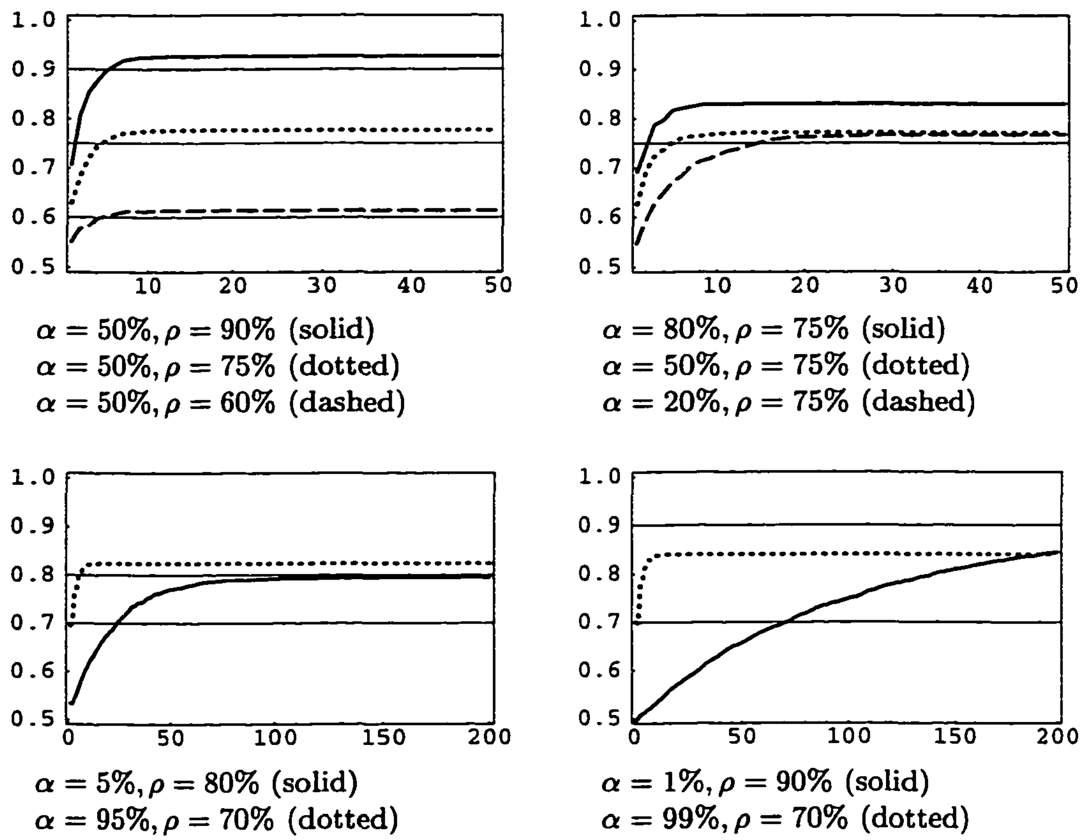


Figure 3.2: Cascade simulations. Values plotted are the simulated probabilities of an individual in a particular queue position making a correct choice. Individual queue position is measured along the horizontal axis. All plotted values are based upon 10,000 simulated cascade events. Horizontal rules indicate precision levels associated with curves in each plot. The slope of each curve indicates the rate of public information arrival, with an informational cascade indicated when the slope reaches zero.

edge. Later, we examine misfits who act in a contrarian fashion but do not necessarily have superior private information.

3.2.4 Cascade Simulations

Figure 3.2 graphically demonstrates the rate of cascade creation for various values of α and ρ. The plots in Figure 3.2 show the simulated probability that an individual in a particular queue position (measured along the horizontal axis) will choose correctly.[6] As the slope of each plotted curve decreases, the rate of public information arrival decreases.

[6]Choose to accept a project with a net payoff of 1 and reject a project with a net payoff of −1.

When a cascade starts the slope of the curves reaches zero.

The rate of public information accumulation depends heavily on α. In the first panel of Figure 3.2, where α is constant for each plotted curve, there is little advantage to being much later than individual 10 or so in the queue as all individuals past this point have roughly the same chance of making a correct choice. In the second panel, where α ranges from 20% to 80%, the value of being later in the queue clearly changes. When $\alpha = 20\%$, individuals as late as number 20 or 25 benefit from additional information revealed by past choices, but when α is relatively high (at 80%) little is gained by waiting past the action of individual 10.

It appears that α is the dominant factor in determining the amount of time (number of individuals) before a cascade begins. This seems reasonable. If individuals seldom receive information, their public actions convey little to others. If an individual does receive an informative private signal, it is less likely that the signal will be dominated by public information and a cascade will take longer to occur.

Based on the first two panels, though α in simulation dominates in determining the *rate* at which public information improves, it has little effect on the eventual *level*. The bottom two panels in Figure 3.2 illustrate the trade-off between α and ρ. These indicate that modest increases in signal precision can offset sizable increases in the likelihood of information arrival.

3.2.5 How Best to Inform?

Earlier individuals clearly benefit more than later individuals from higher rates of private informative signal arrival (α). All individuals benefit from higher values of informative signal precision (ρ). In cases where policy can affect α or ρ, each at some cost, it is not clear which to choose. Should one inform a broad selection of individuals at a cost of not informing any particular individual very well (favor α over ρ) or make certain a few members of the population are very well informed at a cost of not informing most individuals well at all (favor ρ over α)?

On benefits, the plots in Figure 3.2 and the limited ability of an increase in α to impact the terminal level of public information provide a good case for targeting a few members. The cost side seems to reinforce this argument. Broad efforts to inform a population are often very costly compared to targeted efforts to inform small groups. The informational cascade boundary in (3.5) is not affected by α and, as discussed previously, any change in α can affect the ultimate level of public information in only a secondary manner. High levels of α obtained at a cost of low levels of ρ will generally result in quick cascade formation and low ultimate levels of public information.

These results complement earlier cascade literature. It has been an observation that the mere possibility (and in particular environments a near certainty) of cascade occurrence should affect an attempt to control or modify group behavior. In particular, one might do well to focus on key, influential members of a population in order potentially to spur a desired cascade into existence.

3.2.6 Cascade Stability

The stability of an informational cascade depends on the amount of public information on which it is based. This ultimate level of public information, as previously discussed, is largely a function of private informative signal precision ρ and only secondarily of the informative signal arrival rate α. As a result, cascades as modeled thus far are quite fragile and based upon little aggregate public information (in expectation, public information that is just a bit more precise than the information contained in a single informative private signal). There is no significant difference between these results and those of BHW regarding cascade stability. In the next section, where 'misfits' are introduced, the property of cascade stability will change dramatically.

Result 3.2. *Informational cascades remain fragile when only a fraction of the population is informed.*

3.3 The Model with Misfits

What happens to a cascade when an individual breaks from conformity? Such breaks can occur in at least four ways. (1) An individual may be unaware of the actions of predecessors, thus relying on his private signal, whatever it may be; (2) An individual may have private signal precision higher than that of predecessors; (3) An individual may be acting irrationally, without regard to public or private information; and (4) the individual may have different payoffs from others. The section will focus on the first three categories, which will be identified as the cases of newcomers, prophets, and fools.[7]

We will contrast the conventional individuals modeled earlier (*normals*) with others categorized as *misfits*. Newcomers, prophets and fools are types of misfits. Newcomers will be modeled as individuals otherwise identical to normals, with the same values of α and ρ, but for some reason oblivious to any earlier actions. Newcomers may have joined the society later than others, or other barriers could limit their ability to observe predecessors.[8] Internally, newcomers are exactly like normals, they just act without regard to available public information.

Prophets are perfectly informed individuals, $\alpha = 1$ and $\rho = 1$. Prophets will always follow their private signal. Fools, on the other hand, never have an informative private signal and are, like newcomers, unaware of the actions of predecessors. Fools can be seen either as irrational individuals or as rational uninformed newcomers (where $\alpha = 0$ and ρ is irrelevant). Either way, the action of a fool is never informative. We define a cascade, as before, as a situation where an individual chooses to ignore a private signal and follow public information. Thus, a normal may be in a cascade even though a misfit in the same position would not.

Individuals are unable to observe types directly but do know the distribution of types in the population. Later, a program of certification will be considered where types will

---

[7]Regarding (4), see Smith and Sørenson [16].

[8]Viewed another way, newcomers can be seen as over-confident individuals, capable of observing earlier actions but *choosing* to ignore them. See Bernardo and Welch.

be partially revealed.

3.3.1 Model Solution

In contrast with normals, the other categories — newcomers, prophets and fools — may potentially break out of a cascade. Furthermore, an individual who conforms to a cascade may be a newcomer with a possibly informative signal, a prophet, perfectly informed, or merely a fool, with no information at all. Thus, there will be information revealed in conformity, even during a cascade. Every action, whether in a cascade or not, will now affect Ω and thus $X(\Omega)$ and it will become possible for cascades to become more entrenched due to conformity or broken by newcomers, prophets or fools who act in opposition to a cascade.

The solution to a model that includes newcomers, prophets and fools consists entirely of calculating $X(\Omega)$, just as before. In order to do this, it will be necessary to know something about the distribution of individuals. Let $\mathbf{n}$ be the fraction of individuals that are newcomers, $\mathbf{p}$ the fraction that are prophets and $\mathbf{f}$ the fraction that are fools. For convenience, the fraction of normal (regular) individuals will be written as $\mathbf{r} = 1 - \mathbf{n} - \mathbf{p} - \mathbf{f}$.

A newcomer will accept the project with a signal of H or U_H and reject with a signal of L or U_L. Fools can be modeled as uninformed newcomers who accept with a signal of U_H and reject with a signal of U_L. The crucial thing is that newcomers and fools do not react to the actions of predecessors and that fools, in addition, never have private information. Newcomers will always provide information via their actions while fools will not. Normals are modeled just as before and prophets, with perfect private information, are easily modeled below.

Table 3.4 lists the possible signals corresponding to individual actions. Normals always conform to a cascade; thus they can be excluded from possibility when a R_A^c or A_R^c action is observed. There is no private signal that would compel them to oppose a cascade.

The calculation of $X(\Omega)$ that follows answers the question, "what is the probability

Action	Type of Individual Normal (r)	Newcomer (n)	Prophet (p)	Fool (f)
R_A	$\{L\}$	$\{L, U_L\}$	$\{L\}$	$\{U_L\}$
A_R	$\{H\}$	$\{H, U_H\}$	$\{H\}$	$\{U_H\}$
R_U	$\{L, U_L\}$	$\{L, U_L\}$	$\{L\}$	$\{U_L\}$
A_U	$\{H, U_H\}$	$\{H, U_H\}$	$\{H\}$	$\{U_H\}$
R_R	$\{L, U_H, U_L\}$	$\{L, U_L\}$	$\{L\}$	$\{U_L\}$
A_A	$\{H, U_H, U_L\}$	$\{H, U_H\}$	$\{H\}$	$\{U_H\}$
R_R^c	$\{H, L, U_H, U_L\}$	$\{L, U_L\}$	$\{L\}$	$\{U_L\}$
A_A^c	$\{H, L, U_H, U_L\}$	$\{H, U_H\}$	$\{H\}$	$\{U_H\}$
R_A^c	—	$\{L, U_L\}$	$\{L\}$	$\{U_L\}$
A_R^c	—	$\{H, U_H\}$	$\{H\}$	$\{U_H\}$

Table 3.4: Possible individual signals corresponding to elements of Ω.

of a given type of individual (normal, newcomer, prophet, fool) taking a given action (adopt, reject) given a state of the world?" For example, the probability that a fool will take action A_R^c given that $V = 1$ is 1/2. The answer depends on the likelihood of individuals receiving different signals (e.g. the probability that a newcomer will receive a signal that is L or U_L given that $V = 1$). Table 3.4 maps a list of possible signals for each individual type into actions. As before, the calculation of $X(\Omega)$ proceeds from equation (3.1):

$$X(\Omega) = \left[\frac{\mathbf{r}(\alpha\rho) + \mathbf{n}(\alpha\rho + (1-\alpha)/2) + \mathbf{p} + \mathbf{f}/2}{\mathbf{r}(\alpha(1-\rho)) + \mathbf{n}(\alpha(1-\rho) + (1-\alpha)/2) + \mathbf{f}/2}\right]^{(\mathcal{R}_A - \mathcal{A}_R)}$$
$$\left[\frac{(\mathbf{r}+\mathbf{n})(\alpha\rho + (1-\alpha)/2) + \mathbf{p} + \mathbf{f}/2}{(\mathbf{r}+\mathbf{n})(\alpha(1-\rho) + (1-\alpha)/2) + \mathbf{f}/2}\right]^{(\mathcal{R}_U - \mathcal{A}_U)}$$
$$\left[\frac{\mathbf{r}(\alpha\rho + (1-\alpha)) + \mathbf{n}(\alpha\rho + (1-\alpha)/2) + \mathbf{p} + \mathbf{f}/2}{\mathbf{r}(\alpha(1-\rho) + (1-\alpha)) + \mathbf{n}(\alpha(1-\rho) + (1-\alpha)/2) + \mathbf{f}/2}\right]^{(\mathcal{R}_R - \mathcal{A}_A)}$$
$$\left[\frac{\mathbf{r} + \mathbf{n}(\alpha\rho + (1-\alpha)/2) + \mathbf{p} + \mathbf{f}/2}{\mathbf{r} + \mathbf{n}(\alpha(1-\rho) + (1-\alpha)/2) + \mathbf{f}/2}\right]^{(\mathcal{R}_R^c - \mathcal{A}_A^c)}$$
$$\left[\frac{\mathbf{n}(\alpha\rho + (1-\alpha)/2) + \mathbf{p} + \mathbf{f}/2}{\mathbf{n}(\alpha(1-\rho) + (1-\alpha)/2) + \mathbf{f}/2}\right]^{(\mathcal{R}_A^c - \mathcal{A}_R^c)}. \tag{3.9}$$

This expression is equivalent to (3.2) when $\mathbf{n} = \mathbf{p} = \mathbf{f} = 0$ and actions R_A^c and A_R^c are not possible. $X(\emptyset) = 1$ and anything that necessarily moves one of the terms in (3.9) closer to 1 must inhibit the ability of new information to impact $X(\Omega)$. For example, in each of the components of $X(\Omega)$ in equation (3.9), the constant term $f/2$ is in both the numerator and denominator, thus moving each component fraction closer to 1 and reducing the impact of a new information point (individual action) on $X(\Omega)$. This is exactly what one can expect the presence of fools to do. The mere possibility that a particular action may have been chosen by a fool adds noise to inferences drawn from others' actions. In contrast, the fraction of prophets (p) is added only to the numerator in each component of $X(\Omega)$. This moves each component fraction away from 1 and thus increases the implied information content of every individual's action.[9]

3.3.2 Stability and the State of Public Information

As long as there are newcomers or prophets around, each individual's action will always provide some public information. The aggregation of public information never stops and there is no upper bound on the effective precision of public information. Cascades cease to be fragile, quasi-permanent events where all learning stops and become increasingly rigid events where normals ignore their private signals while misfits keep adding to public information. Eventually, the probability that any individual is correct must then almost surely approach 1. Just a small chance that any given individual might be a newcomer, for example, is all it takes.

The impact of public information on periods of cascade will be the same as before. New information will simply be folded into $X(\Omega)$. Early cascade stability is undermined by newcomers and prophets for the obvious reasons. Eventually, though, one can expect things to become very stable. As more and more information gets added to Ω and $X(\Omega)$ approaches an extreme of 0 or infinity, no amount of less-than-certain public information will be able to affect the action of normals. For any given cascade event it is always

[9]It can be shown that each component of $X(\Omega)$ in (3.9) is greater than 1 which means that an increase in the numerator of each component with the denominators unchanged will move each component of $X(\Omega)$ away from 1.

possible that $X(\Omega)$ will never approach these extremes. In expectation, however, the constant stream of new information makes this an almost sure outcome.

Proposition 3.2. *When future acting individuals may be misfits, acting without regard to public information, the precision of public information is no longer limited by a cascade boundary and thus cascades cease to be fragile.*

The possible presence of misfits is what eliminates cascade fragility. If, for example, misfits were present in some of the early queue positions but not later (and this is known to all), then a cascade-inducing level of public information would remain unchanged after the action of the last misfit and thus a cascade in such an environment could be quite fragile. If, in this situation, no cascade is initiated after the action of the last possible misfit, then the situation becomes just as that when no misfits are present, with the same outcome of cascade fragility.

3.3.3 Newcomers

A newcomer is defined as an individual who does not observe the actions of any predecessors. This has several possible interpretations. Newcomers may be recent arrivals, individuals who are poorly placed or even prohibited from observing predecessors, or could be viewed as overconfident individuals who observe predecessors but underweight the value of others' information (see Bernardo and Welch). Whatever the interpretation, newcomers are exactly like normals with the exception that they do not act with regard to public information.[10]

In the $(\mathcal{R}_U - \mathcal{A}_U)$ component of $X(\Omega)$ newcomers are just like normals, which is to say that newcomers and normals react to signals identically when public information suggests indifference. For out-of-cascade reversals (actions R_A and A_R), newcomers dilute the signal when compared to normals. These reversals of public information are

[10] A fraction n of newcomers can be modeled by a total fraction n of prophets and fools where $\mathrm{p} = \mathrm{n}\alpha(2\rho - 1)$ and $\mathrm{f} = \mathrm{n}(\alpha(1-2\rho)+1)$. For example, with $\alpha = 1/2$ and $\rho = 3/4$, a fraction n of newcomers can be modeled as a total fraction n of prophets and fools where there is 1 prophet for every 3 fools. When $\alpha = 1/10$ in this example, newcomers can be modeled as 1 prophet for every 19 fools.

the most revealing action that a normal can take, perfectly signalling the individual's private informative signal.

With newcomers, however, an out of cascade reversal can just be due to a newcomer with a possibly uninformative signal. As a result, the presence of newcomers limits public information gathering in this case. The opposite is true for the remaining cases. During out of cascade conformity (R_R and A_A), newcomers add less noise to the signal's multiplier than normals ($(1-\alpha)/2$ vs. $(1-\alpha)$). In-cascade action clearly favors the presence of newcomers as normal individual actions are not at all informative during cascade conformity (R_R^c and A_A^c) and add pure noise to the multiplier, while cascade reversal is impossible for normals. The presence of newcomers thus produces a mixed result. Out of cascade reversal of public information becomes less informative while all other actions become more informative or are unaffected.

Newcomers can be viewed as normals somehow *prohibited* from viewing the actions of predecessors, though not all individuals would prefer such prohibitions. One line of reasoning would be that everyone would like everyone except themselves to be prohibited from viewing the actions of predecessors and thus be induced to reveal the information in their private signal more clearly. However, when individuals break from a known standard a great deal of information is revealed. The "known" part is what is crucial. When an individual breaks from convention, it is more informative if he did it on purpose. Ideally, from a normal's point of view, everyone that broke with convention would be informed of the convention and everyone that followed convention would have no idea that they were doing it. Then, the maximum possible amount of information would be revealed. Newcomers only get this half right. When they follow convention (action sequences R_R, A_A) they are completely oblivious of the fact and thus transmit just as much information as a normal individual when no convention exists, yet when they happen to break from a convention (actions R_A, A_R), they are just as oblivious and thus don't provide the information that a well informed normal does when acting in the same manner.

Result 3.3. *The inability of newcomers to break from* known *convention results in ac-*

tion that is less than perfectly informative even though newcomers always follow their private signal.

Are newcomers a good thing or not? The answer is yes, most of the time. When a cascade (for normals) is not in progress, newcomers can present both a cost and a benefit. But during a cascade, newcomers are purely beneficial. With newcomers, every action during a cascade *might* contain information and thus, in expectation, does provide information. Breaks from cascade conformity, best of all, are a sure sign of a newcomer (or prophet or fool) with potential information to reveal.

Proposition 3.3. *During cascades, newcomers present an unambiguous benefit to public information when compared to normals, contrasting with the potential cost of newcomers during out-of-cascade actions.*

This contrasts with Bernardo and Welch who find entrepreneurs (equivalent to newcomers here) to be always beneficial to non-entrepreneurs. In the Bernardo and Welch model, all normal individuals (non-entrepreneurs) receive informative private signals. The signal arrival rate (α) in the model presented here reveals that newcomers can present a cost to early acting normal individuals when not all normal individuals receive an informative private signal ($\alpha < 1$).

Returning to the language of prohibitions, if newcomers are generally a good thing and one can view a newcomer as a normal prohibited from observing the actions of predecessors, then aren't such prohibitions a good thing as well? For whom? Certainly not for the prohibited individuals. But in situations where new individuals can be invited to join a group (and current individuals, the normals, set the rules), some limitations on how these new individuals can view the actions of others would appear likely.

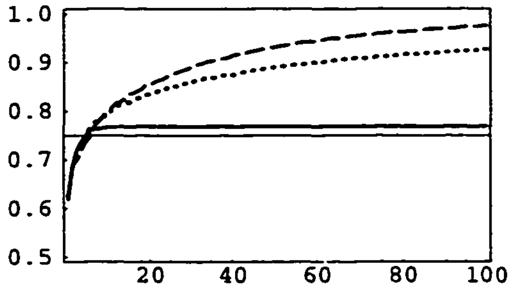
Result 3.4. *To the extent that newcomers are beneficial to normals, there is an incentive for some individuals to regulate the degree to which others can view earlier actions and thus create newcomers.*

The interpretation of newcomers as overconfident normals illustrates how such individuals can be beneficial. To the extent that newcomers are desired by normals so then will overconfident normals be desired as well.

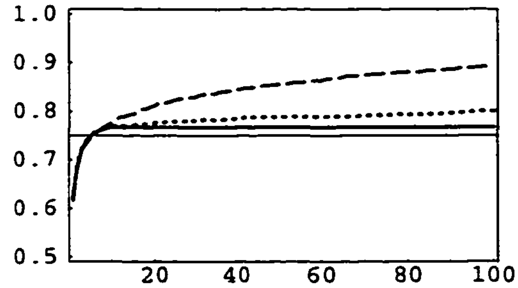
3.3.4 Simulations With Misfits

Figure 3.3 contains summary plots of thousands of simulated cascade events, like those in Figure 3.2. The plotted probabilities are for *normal* individuals. This is done by simulating cascades as before, with the addition that at each queue position a hypothetical normal is introduced. The question is then, "What would a simulated normal have done at this position had he acted?" The plot therefore shows how a normal in the given population would expect to perform if he were assigned a particular queue position. As a result, these plots are directly comparable to those in Figure 3.2. The first two panels of Figure 3.3 illustrate clearly the impact of newcomers.

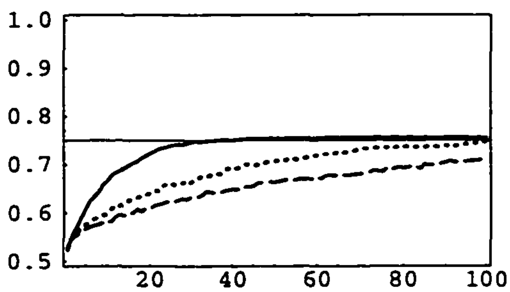
The two middle panels illustrate the relationship between the level of α and the harm that newcomers bring to early normals. Recall that the problem with newcomers (as seen by normals) is that they introduce noise into the most informative (R_A and A_R) actions. This level of noise is revealed in equation (3.9) to be entirely a function of α. In the $(\mathcal{R}_A - \mathcal{A}_R)$ component of this expression for $X(\Omega)$, the difference between the fraction $\mathbf{r}$ of normals and the fraction $\mathbf{n}$ of newcomers is the (effective) noise factor $(1-\alpha)/2$. This factor, appearing in both the numerator and denominator for newcomers drives out information content. Large values of α then lead to greater harm due to newcomers during the R_A and A_R actions, which will naturally tend to occur early in the event, when cascade periods are less likely. Later, when in-cascade actions (R^c_R, A^c_A, R^c_A and A^c_R) are more likely, the benefit of newcomers overcomes this cost. For lower values of α one can expect marginally higher levels of opposition to newcomers, driven by increased harm that newcomers cause early individuals. For small populations of individuals where few individuals are informed, this could manifest itself in clear hostility to newcomers. When newcomers are seen as normals prohibited from viewing the actions of predecessors, a hostility to newcomers could emerge as internal efforts to increase knowledge of the actions of predecessors. Examples of such behavior could be public posts of the actions of predecessors or required attendance at events where actions takes place.



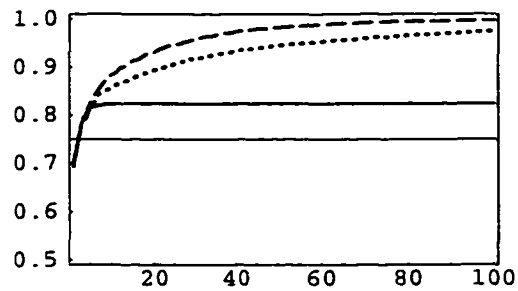
$\mathbf{n} = 0\%$ (solid)
$\mathbf{n} = 20\%$ (dotted)
$\mathbf{n} = 50\%$ (dashed)



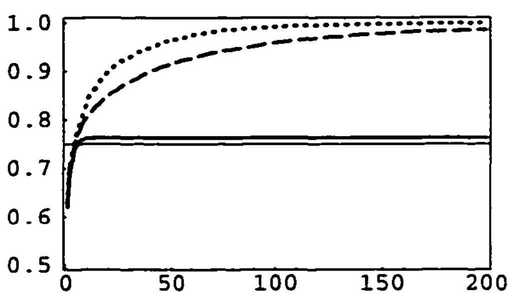
$\mathbf{n} = 0\%$ (solid)
$\mathbf{n} = 1\%$ (dotted)
$\mathbf{n} = 10\%$ (dashed)



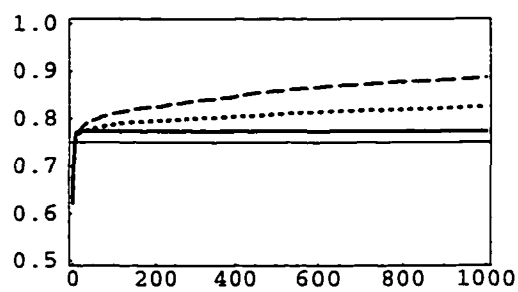
$\alpha = 10\%, \mathbf{n} = 0\%$ (solid)
$\alpha = 10\%, \mathbf{n} = 20\%$ (dotted)
$\alpha = 10\%, \mathbf{n} = 50\%$ (dashed)



$\alpha = 80\%, \mathbf{n} = 0\%$ (solid)
$\alpha = 80\%, \mathbf{n} = 20\%$ (dotted)
$\alpha = 80\%, \mathbf{n} = 50\%$ (dashed)



$\mathbf{p} = 0\%, \mathbf{f} = 0\%$ (solid)
$\mathbf{p} = 12.5\%, \mathbf{f} = 12.5\%$ (dotted)
$\mathbf{p} = 5\%, \mathbf{f} = 5\%$ (dashed)



$\mathbf{p} = 0\%, \mathbf{f} = 0\%$ (solid)
$\mathbf{p} = 0.1\%, \mathbf{f} = 0.9\%$ (dotted)
$\mathbf{p} = 0.5\%, \mathbf{f} = 4.5\%$ (dashed)

Figure 3.3: Cascade simulations with misfits. Unless stated otherwise, $\alpha = 50\%$, $\rho = 75\%$, $\mathbf{n} = \mathbf{p} = \mathbf{f} = 0\%$. Values plotted are the simulated probabilities of a *normal* individual in a particular queue position making a correct choice. Plots are otherwise derived and presented as in Figure 3.2.

Result 3.5. *As the fraction of informed individuals falls, more early-acting individuals will be harmed by newcomers. For small populations where only a small fraction of individuals is privately informed, opposition to newcomers can emerge.*

The lower left panel in Figure 3.3 shows the large benefit normal individuals derive from prophets and the relatively small cost of fools. A small fraction of fools is like a small decrease in α, which, as seen earlier, affects the expected outcome only marginally. In the lower right panel, prophets and fools total to 1% (dotted line) and 5% (dashed line) with fools outnumbering prophets 9 to 1. The benefit of even this small number of prophets (1 out of 1000 individuals for the dotted line) is clear even when the prophets are heavily outnumbered by fools. Recall that individual type is unobservable and that the action of a prophet can never be distinguished from that of a fool.

Result 3.6. *Prophets and fools, while indistinguishable by their actions, can still greatly benefit normals. This is the case even when prophets are a very small fraction of the population and heavily outnumbered by fools.*

3.4 Weakening Private Information

Individuals so far have had two forms of private information. As in other cascade models, individuals privately know the value of their signal, but, in addition, individuals in this paper have private knowledge of whether they have received an informative signal. This section weakens this assumption by removing the individuals' private informational advantage of knowing whether an informative signal was received.

3.4.1 Certification of Knowledgeable Individuals

We examine here the effect of external certification. Suppose that a body can identify individuals who receive informative signals. This certification body then publicly announces the identity of each individual that is informed (receives a signal of H or L).

Without certification, there were the two parameters, ρ and α. Under certification there are two distinct types of individuals, those with $\alpha = 1$ and those with $\alpha = 0$.

The individuals that are certified as informed will collectively be in a setting with ρ as before and $\alpha = 1$. The remaining individuals, those who are uninformed, will be interspersed, will always follow public information and will contribute nothing to public information. Uninformed individuals will simply free-ride on the public information that happens to be available when they act. Everyone else will know they are uninformed and will therefore ignore their actions.

To illustrate the result of certification, imagine that all informed individuals are certified prior to the action of the first individual. Certified individuals are placed in one room and uncertified individuals in another. A one-way glass wall separates the two rooms so that the uncertified individuals can observe the certified. The certified individuals will act in a setting with $\alpha = 1$ and the uncertified individuals will watch and act when it is their turn. There is no need for the certified individuals to observe or even be informed of the existence of the uncertified individuals. One can imagine a counter in each room, indicating the turn of the next individual. If it takes some amount of time for an individual to act (accept or reject the project), the uncertified individuals may slow down the certified individuals, but, other than that, they will have no impact.

Misfits can easily fit into the story. Newcomers will be pre-certified just like everyone else and one can think of certified newcomers as individuals that are in the room of certified individuals, not paying attention until their turn arrives. Importantly, normals can't tell if the newcomers have been paying attention or not (or who they are). Prophets will always be certified, so any individual in the certified room may be a prophet, but will never be a fool. Fools will all be in the uncertified room, not paying attention.

All normals will benefit from certification while newcomers, prophets and fools will not be harmed. Normals without an informative private signal will be harmless (other than possibly slowing the sequence of informed actions) and, like informed normals, will benefit from knowing which predecessors were informed. Informed individuals (some normals, some newcomers and all prophets) will be in a regime modeled earlier with $\alpha = 1$. Uninformed individuals (some normals, some newcomers and all fools) will take their turn and be ignored.

Result 3.7. *Certification of knowledgeable individuals will be welcomed by all with some individuals benefiting and no individuals harmed. Under a program of certification the only possible harm to normals that arises from the presence of misfits is the possible delay that results from the actions of fools.*

Under certification, with α effectively at 1, all normals will welcome newcomers. If newcomers are viewed as normals prohibited from viewing earlier actions, then certification will imply stronger incentives for some normals to try to prohibit others from observing previous actions. Oddly then, certification, a process generally intended to reveal information, can produce an incentive to hide information as some certified individuals try to limit the extent to which other certified individuals can observe predecessors.

Result 3.8. *A process of certification, where individuals with informative private signals are identified to all, will produce an increased incentive for some individuals to prohibit others from observing earlier actions.*

3.4.2 Inability to Identify Informative Private Signals

Certification removes the informational advantage of knowing whether one's private signal is informative by telling everyone else whether it is. Another way to remove this private informational advantage is to hide the knowledge from each individual. What happens if individuals are unable to tell if their private signal is informative?

Individuals can now receive only two possible signals. The signal H_U will indicate what was previously a private signal of H or U_H, while L_U will represent L or U_L. When, for example, an individual receives the private signal H_U, he is unable to tell whether the signal is informative (H) or uninformative (U_H). Each individual still knows his informative signal arrival rate α and precision ρ and all individuals know that no individual can distinguish informative and uninformative signals. Individuals, as a result, have an effectively lower signal precision where every signal must be treated as informative. This signal precision changes from ρ (when an informative signal is received) to $\alpha\rho+(1-\alpha)/2$

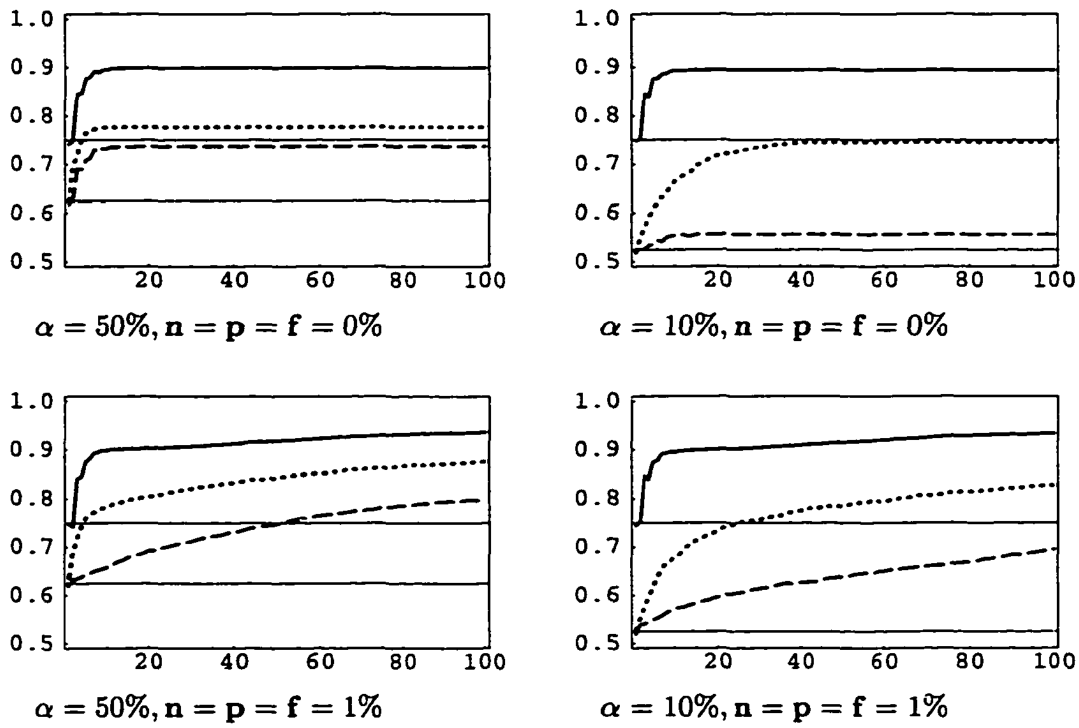


Figure 3.4: Cascade simulations. In all plots the solid line represents the model where all individuals are privately informed ($\alpha = 1$) and know this is so, the dotted line represents the model where a fraction α (indicated above) are informed and can identify an informative private signal. The dashed line represents a model where individuals cannot identify informative private signals as in equation (3.10) with α indicated above. Plots are derived and presented as in Figure 3.3 with $\rho = 75\%$ in all cases.

for every signal:

$$\Pr(s = H_U | V = 1) = \alpha\rho + (1-\alpha)/2 = \Pr(s = L_U | V = -1). \tag{3.10}$$

The parameters describing each individual's signal have changed: $\{\alpha, \rho\} \to \{1, \alpha\rho + (1-\alpha)/2\}$. This will have two impacts. First, the no-cascade range in (3.5) will tighten, resulting in actions taken in a cascade range for more values of $X(\Omega)$, i.e. individuals will cascade based upon less public information. Second, the $X(\Omega)$ function will change. All of the earlier analysis of changes in ρ and α applies. Figure 3.4 illustrates the cost that arises when individuals do not always possess informative private signals or are unable to identify informative private signals. Inability to identify informative private signals is most costly when such signals are relatively rare (α is low). The bottom panels in Figure

3.4 illustrate that this cost to normals remains when misfits are present, though in all cases a small fraction of misfits produces a marked effect on the welfare of normals.

Newcomers and normals are harmed by the inability to identify their informative private signals. Normals are further harmed via the decline in public information Ω. As above, with α effectively at 1, newcomers will be desired by all normals, leading to a greater incentive for some normals to limit the ability of others to observe earlier predecessors.

Result 3.9. *When individuals are unable to know if they are informed, the effective rate of private information arrival is 1 for both newcomers and normals, resulting in no harm to normals arising from the presence of newcomers.*

3.5 Conclusion

In an informational cascade model we introduce individuals with no private information and heterogenous individuals in the form of social misfits. We show that the ultimate precision of public information is largely independent of the fraction of individuals that are privately informed. This level of precision is bounded by constants of social information aggregation. The levels of these constants are bounded below by the precision of an individual's informative private signal and above by both this precision and the fraction of individuals who are privately informed. We also show that cascade fragility is eliminated by the possible presence of misfits and that misfits make important contributions to social learning.

When all individuals are privately informed, newcomers (a type of social misfit) act to the benefit of others. They do so by always revealing their private information. But, when not everyone is privately informed, newcomers can inhibit the growth of public information and thus harm other individuals. In circumstances where a small fraction of individuals is privately informed or where the population size is small, hostility to newcomers can emerge.

Newcomers are like over-confident individuals in that they effectively ignore public information. Newcomers can also be seen as normal individuals *prohibited* from observing the actions of predecessors and thus denied a great deal of what would otherwise be public information. In this sense, the benefits derived from newcomers can produce incentives for prohibitions. To the extent that society benefits from the presence of misfits, society will thus have incentive to create misfits where none or few exist. Prohibiting some from observing the actions of others does this by effectively creating social newcomers.

A social program of certification, where those privately informed are publicly identified, would be welcomed by all. Normal individuals would clearly benefit while misfits would not be harmed. Social norms that act to identify misfits would thus find little constituency or opposition among misfits and clear support among normal individuals.

Bibliography

[1] Abhijit V. Banerjee. A simple model of herd behavior. *The Quarterly Journal of Economics*, 107:797–817, 1992.

[2] Randolph P. Beatty and Jay R. Ritter. Investment banking, reputation, and the underpricing of initial public offerings. *Journal of Financial Economics*, 15:213–232, 1986.

[3] Tony Bernardo and Ivo Welch. On the evolution of overconfidence and entrepreneurs. *UCLA Anderson School Working Paper*, 1997.

[4] Sushil Bikhchandani, David Hirshleifer, and Ivo Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, 100(5):992–1026, 1992.

[5] H. Henry Cao and David Hirshleifer. Conversation, reporting biases and informational cascades. *working paper*, 1998.

[6] Christophe Chamley and Douglas Gale. Information revelation and strategic delay in a model of investment. *Econometrica*, 62:1065–1085, 1994.

[7] Glenn Ellison. Learning, local interaction, and coordination. *Econometrica*, 61:1047–1071, 1993.

[8] Jeffrey Ely. Local conventions. Economics Department, University of California at Berkeley, November 1995.

[9] Farak Gul and Russell Lundholm. Endogenous timing and the clustering of agent's decisions. *Journal of Political Economy*, 103(5):1039–1066, 1995.

[10] Eric Hoffer. *The True Believer.* Harper & Row, 1951.

[11] Roger G. Ibbotson and Jeffrey F. Jaffe. "Hot issue" markets. *The Journal of Finance*, 30(4):1027–1042, 1975.

[12] Roger G. Ibbotson and Jay R. Ritter. Initial public offerings. In R. A. Jarrow, V. Maksimovic, and W. T. Ziemba, editors, *Finance*, volume 9 of *Handbooks in Operations Research and Management Science*, chapter 30, pages 993–1016. North-Holland, 1995.

[13] Michihiro Kandori, George J. Mailath, and Rafael Rob. Learning, mutation, and long run equilibria in games. *Econometrica*, 61:29–56, 1993.

[14] In Ho Lee. On the convergence of informational cascades. *Journal of Economic Theory*, 61:395–411, 1993.

[15] Kevin Rock. Why new issues are underpriced. *Journal of Financial Economics*, 15:187–212, 1986.

[16] Lones Smith and Peter Sørensen. Pathological outcomes of observational learning. *working paper*, 1998.

[17] Ivo Welch. Seasoned offerings, imitation costs, and the underpricing of initial public offerings. *The Journal of Finance*, 44(2):421–449, 1989.

[18] Ivo Welch. Sequential sales, learning, and cascades. *The Journal of Finance*, 47(2):695–732, 1992.

[19] H. Peyton Young. The evolution of conventions. *Econometrica*, 61:57–84, 1993.

[20] Jianbo Zhang. Strategic delay and the onset of investment cascades. *Rand Journal of Economics*, 28(1):188–205, 1997.

IMAGE EVALUATION TEST TARGET (QA-3)

150mm

6"

APPLIED IMAGE . Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989