CHAPTER

3

Dedicated Followers of Success

A Computational Model of Fashionable Innovation

Michael W. Macy and David Strang

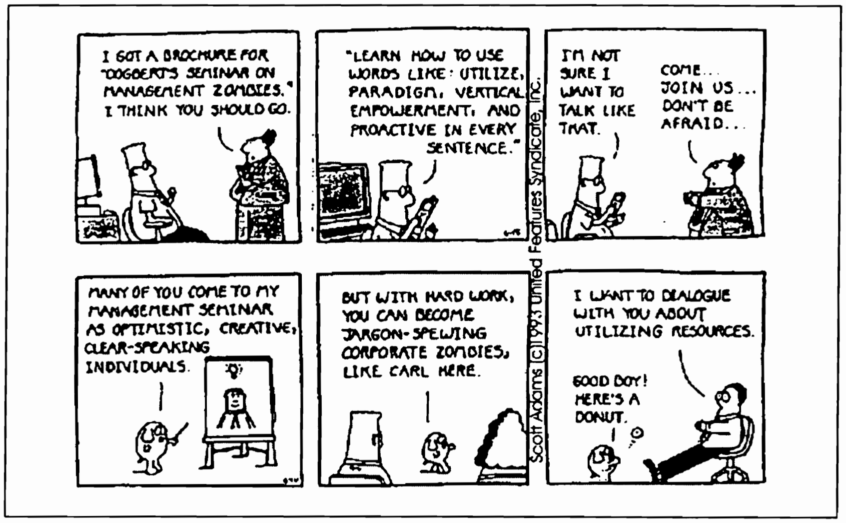
"It will make or break him so he's got to buy the best, 'cause he's a dedicated follower of fashion."

– Ray Davies, "Dedicated Follower of Fashion," 1966

Rationality and Conformity

In *Dogbert's Top-Secret Management Handbook,* Scott Adams (1996) pillories corporate executives as dimwitted lemmings who proselytize recycled innovations to justify their existence. The faddish pattern is familiar: A "hot" innovation is widely heralded in the business press, followed by harsh critiques that lampoon the innovation and the managers taken it by it.

The view of managers as "dedicated followers of fashion" is contradicted by evidence that most are skeptical, rational, and paid high salaries to "get it right." But if managers are in fact dedicated followers of *success*, how do we explain the mercurial pattern? We use an agent-based computational model to suggest a possible answer: The explosive growth and sudden collapse of fashionable innovations may actually be due, not to lemminglike conformity, but to the opposite problem—a preoccupation with performance.



From *Dogbert's Top-Secret Management Handbook*, 1996, by Scott Adams.
Courtesy, United Media.

The Problem of Fashionable Innovation

The variety of managerial innovations that have gained widespread attention but little staying power in the business community is indeed extraordinary. A short list might include: "management by objectives," "T groups," "sensitivity training," "work redesign," "work simplification," "job enrichment," quality of work life," "zero defects," "quality circles," "Scanlon plans," "pay for performance," "matrix management," "autonomous work groups," "management by policy," "total quality management," "reengineering," and "value chain integration."

Quality circles provide a case in point (Cole 1989, Abrahamson and Fairchild 1997, Strang 1997). The Japanese "quality control circle" developed in the early 1960s as a way to expand already comprehensive quality control efforts down to the shop-floor. Although quality circles appeared in a handful of American firms during the next two decades, the movement really took off at the beginning of the 1980s, led by a flurry of high-profile articles in the business media touting their responsibility for Japanese success. Articles like "Made in Japan-Quality Control Circles" emphasized the remarkable transformation in Japanese products. Much discussion also rehearsed the benefits of quality circles in celebrated exemplars like Lockheed (cost savings of $3 million, tenfold defect reduction, 6:1 ROI, and ninety percent employee satisfaction). From their base in high-tech manufacturing, quality circles spread to industries like health care, financial services, and even education and government. By mid-decade, Lawler and Mohrman (1985) reported that virtually all of the Fortune 500 had given quality circles a try.

Yet by mid-decade, the bloom was clearly off the rose. While most empirical research suggested reasonably high levels of quality circle effectiveness (Barrick and Alexander 1987), both firms and consultants moved on to new approaches, like total quality management and autonomous work groups. The "quality circle" all but disappeared from management discourse. How could highly paid managers, trained in the best business schools, and under enormous pressure to "get it right," be so gullible and mercurial?

Previous efforts to solve this riddle can be divided into two approaches, which we will label "performance models" and "conformance models." "Performance models" assume that some innovations are better than others, that some actors are more innovative than others, and that better ideas tend to be more popular (and vice versa[1]). The optimizing mechanism in these models is either ecological selection or rational choice. Organizational ecologists (Hannan and Carroll 1992) can explain cascades as a response to selection pressures given competition for limited resources. "Best practice" comes to dominate the population (due to selection pressures that favor a positive balance sheet), while popular practices are more likely to outperform those that are innovative (due to selection pressures that favor legitimacy).

Game and rational choice theorists also posit a link between popularity and performance that reconciles optimizing behavior and what Banerjee (1992) calls "herd dynamics." Managers economize on information costs by relying on the observed actions of others as indicators of private information. "Herd dynamics" are also generated by Hedström's (1998) "social proof" decision model. Like Banerjee, Hedström assumes that innovations differ in utility and decision-makers differ in self-reliance. A few self-reliant managers will identify "best practice," based on rational evaluation. The majority then follows suit by imitating observed behavior. Assuming constant utilities, the population eventually converges on the best practice without the need for every member of the population to rediscover the wheel. These models reconcile "herd dynamics" and rationality by postulating that innovations affect outcomes and that imitation is a cost-effective way for the herd to find the best practices.

"Conformance models," in contrast, assume that models are imitated because they are popular, even if they are inferior in performance and costly to implement. A variety of mechanisms giving force to popularity are noted. Institutional accounts point to taken-for-granted understandings (Meyer and Rowan 1977, DiMaggio and Powell 1983, Fligstein 1990). In more choice-theoretic approaches, emergent norms are seen as coordinating mutual expectations of behavior that reduce uncertainty (Brinton and Nee 1998). Network arguments emphasize how close contacts make a new practice understandable (Davis 1991). Krackhardt (chapter 8) models intraorganizational diffusion of controversial innovations whose value is "socially" rather than "rationally" determined. And in social psychological versions, conformance is generated by emotional and cognitive needs such as the affirmation of "social identity" (Tajfel 1982) or reduction of cognitive dissonance (Festinger 1957). Dissonance theories predict stronger attraction to proposals that are costly than to those that require little investment. Such uncritical and lemminglike behavior is captured by Adams' *Top-Secret Management Handbook*, in which management fads are portrayed as a parody of rational calculation.

While both "performance" and "conformance" models provide important insights, they are not well posed to explain a fad, where popularity first skyrockets and then plummets. Neither approach has trouble with the up-phase, understanding it either as the spread of a better mousetrap or as a self-reinforcing dynamic where popularity feeds upon itself. But both approaches have trouble with the down-phase. If innovations are adopted because they are effective, why are they dropped? And if popular practices are highly legitimate or taken-for-granted, how do managers get the courage or the insight to abandon them?

Because they tend to imply similar dynamics, even arguments combining performance and conformance have trouble accounting for faddish cycles. For example, Abrahamson and Rosenkopf (1993) suggest bandwagon models incorporating both perceptions of the effectiveness of an innovation and sensitivity to the adoptions of others. But this framework suggests that diffusion processes equilibrate once the choice-contagion feedback is played out.[2]

We propose an alternative line of analysis, extending the argument developed in Strang and Macy (1999). Let us imagine that *firms never adopt an innovation simply because it is popular,* and that managers are highly intelligent and reluctant to innovate without compelling evidence that they will earn a sufficient return on their investment. Moreover, like the "conformance models," let us assume that *popular* innovations *are no better than any other.* Management fads are nevertheless possible, we contend, caused not by conformity but by the very preoccupation with performance that is supposed to prevent lemminglike behavior. We hypothesize that *an obsession with performance leads to a focus on winners, not losers, by managers as well as the business press.* This in turn generates a self-reinforcing bias in access to confirming over disconfirming information.

This bias generates behavior that resembles "pluralistic ignorance," an unstable equilibrium in which individuals suppress private beliefs that contradict public knowledge and in so doing reinforce the credibility of a self-organized illusion (Miller and McFarland 1987, see also Korte 1972 and Taylor 1982). Pluralistic ignorance reflects a type of cognitive error—epistemic bias—in which people attribute a private source to public knowledge. Individuals tend to assume that others have privately confirmed the public knowledge that the focal individual is adopting in the absence of private confirmation (or even in spite of private disconfirmation). The response in turn reinforces an equivalent illusion in the minds of others. Epistemic bias implies conformist behavior—a tendency to accept as true that which appears to be popularly believed.

The prototypical illustration of epistemic bias is Andersen's story of "The Emperor's New Clothes," in which the emperor is tricked by swindlers (consultants?) into purchasing fictitious garments that would make him the most admired man in the kingdom (Andersen 1994). Each citizen assumes that her inability to see the clothes is her own deficiency, given that everyone else is cheering with admiration. By joining the chorus, each citizen in turn reinforces the evidence that sustains the illusion. Each individual has ready access to public evidence that supports the illusion but can know only their own individual doubts and not those of others. Each overlooks the possibility that others are suppressing contradictory private beliefs for precisely the same reason, namely, that most people are publicly confirming the belief. In Andersen's story, the illusion is shattered when a child, attending only to her direct observation, comments that the foolish old man is naked.

This form of pluralistic ignorance explains conformity with unpopular social norms. Kuran (1995) and Bicchieri and Fukui (1998) have formally modeled information cascades generated by individual reluctance to reveal private doubts about the validity of public knowledge. "Once an individual acts only on the information obtained from others' actions, his decision conveys no truthful information about his private information or preferences. Because the conformity of individuals in a cascade has no informational value, cascades are fragile and could be upset by the arrival of new (truthful) public information" (1998, p. 14). The problem they model centers on the widespread but mistaken assumption that others support an emergent consensus on the basis of an independent assessment, when, in fact, everyone suppresses private assessments that contradict what they believe to be independent assessments by others. Curiously, each individual knows that they personally ignored (or failed to make) an independent assessment of the validity of the information, yet they assume that they are the only one to be influenced by others.

Confirmation Bias

We believe a similar self-reinforcing illusion can also result from a different cognitive error, a bias towards confirmational evidence in the attribution of success. Confirmation bias does not imply conformity and is consistent with the assumption that managers adopt innovations based on readily available information about their relative merits. Paradoxically, confirmation bias can be attributed in part to a preoccupation with performance and success, a "rational" orientation that is usually regarded as the antithesis of "mindless conformity." We will show that cost-benefit calculations, when distorted by a focus on winners rather than losers, can generate cascades of "fashionable innovation" that resemble fadlike conformance.

Confirmation bias has been demonstrated in the "Wason selection task," a laboratory experiment in which subjects are shown four cards, "A," "B," "2," and "3"(Gilovich 1993: 33). Subjects are asked to test the claim that *"all cards with a vowel on one side have an even number on the other"* and are asked to turn over the minimum number of cards needed to test this hypothesis. Readers are invited to try this experiment at home. Intuitively, which cards would you want to inspect?

In laboratory experiments, subjects correctly turn over the "A" in order to confirm the presence of an even number. However, they then turn over the "2" to confirm the presence of a vowel, even though this does not contribute any useful information. Even if the "2" has a consonant on the other side, this does not contradict the hypothesis that all cards with a vowel on one side have an even number on the other. Subjects generally fail to turn over the "3," even though a vowel on the other side would disconfirm the hypothesis.

This logical error is compounded by asymmetry in access to confirming and disconfirming information. This asymmetry is often the case when the information is about success and failure. It is generally much easier to obtain information about individuals who succeed than about those who fail. Success is then attributed to associated traits even if identical traits might be present among those who fail. If these traits are then adopted, the odds are increased that the next "winner," even if randomly selected, will confirm the "value" of the traits to which success has been falsely attributed. Thus, confirmation bias at the individual level is structurally reinforced. Bias in access to confirmatory evidence may also shape managers' "first impressions" which in turn produce overconfidence and a neglect of disconfirmatory tests. (For a study of first impressions in confirmation bias see Rabin and Schrag 1999.)

Consider the use of "bells and whistles" on slot machines. Most bettors at the casino are losing, but the cacophony on the winning machines creates the impression that many others are winning. The losers assume that they are atypical and continue to play (and lose). The more people who play, the more frequent the alarms, and the more credible the illusion of widespread success.

An important element in the slot machine example is the bias in access to information about winners and losers. It is easy to know about the winners—indeed, one cannot help but hear the celebration—but knowledge of the losers tends to remain private.

We hypothesize that "confirmation bias" can generate managerial fads. To develop a strong test of this argument, we adopt a strict view of the intended rationality of managers and executives. We assume that managers neither know nor care what most people think or do. Indeed, they prefer to be different so long as "different" means being alone at the top. Managers have no interest in adopting innovations simply because they are popular, and they base their decisions strictly on evaluations of the cost-effectiveness of competing proposals. Thus, in contrast to Bicchieri and Fukui, we do not assume that "individuals assess others' reasons for action as different from their own" (1998, p. 15). Nor do we assume that individuals suppress private doubts out of concern for acceptance or fear of ostracism. They are concerned not with "fitting in" or"going along" but only with competitive success.

Of course, this is a partial view that misses important features of management and the institutional landscape of contemporary business. Management centrally involves the interpretation and creation of social meaning, and managers are thus necessarily alert to what others do and say (Eccles and Nohria 1992). And decisions are often made by teams, which may be more vulnerable to epistemic bias due to a diffusion of responsibility. (For an overview of differences in cognitive bias between individuals and groups, see Kerr, McCoun, and Kramer 1996).

We should note, however, that an atomistic, efficiency-centered model of managerial decision-making resonates with the way the business community understands and portrays itself. Managers seldom view their actions as efforts to join bandwagons, fit in,or stick with the herd. They celebrate instead innovative and experimental tactics, the drive to beat the competition, and a single-minded focus on results.

What these managers may not realize is that this preoccupation with performance also makes them vulnerable to adoption cascades, even when the fad has little or no intrinsic merit. If the manager's concern is to find out what works and what does not, the manager's problem is that she has much greater access to knowledge about the former than the latter. Evaluations of competing alternatives depend on perceptions that may be distorted by behavioral responses to bias in the access to information needed for those evaluations. Managers cannot directly assess how well innovations work, but they can try to learn from the coincidence of innovative ideas and subsequent outcomes. This involves making decisions on the basis of outcomes experienced by other adopters, and on their own experience once they adopt.

The fundamental problem is that managers (like bettors) are much more likely to hear about innovations when innovators are successful than when they fail, for three reasons:

1. Firms have very limited information about the strategies and performance of other firms. Given that constraint, confirmation bias directs learning toward "success stories"—innovations that were adopted by top-performing firms.
2. Like slot machines, firms are more likely to announce their successes than their failures.
3. Finally, actors seeking to promote a new practice (managers inside the firm, outside consulting agencies, and the business press covering a "hot" innovation), are more likely to disclose it if the firm's performance was outstanding rather than disappointing. That is, promoters can be expected to announce the winner's "secret" if the firm succeeds, but to keep quiet if the firm stumbles.

In sum, managers' interests in *acquiring* and *promoting* the innovations used by top performers combine to create a strong bias in access to knowledge needed for evaluation. As a result, firms have greater difficulty and less interest in obtaining public information about innovations that fail.

Moreover, public knowledge of "best practice" is confirmed when others observe these approaches used by top-performing firms. This knowledge then leads more firms to adopt the innovation. The higher the proportion of the population that adopts the leading innovation, the greater the odds that the top performing firm will be using it, even if the firm's success had nothing to do with it but was entirely due to chance. The problem is that in acting on this public knowledge, each manager inadvertently reinforces the process by which this information is confirmed.[3]

In addition, managers find it difficult to know if their own unfavorable experience with an innovation reflects intrinsic problems or their own mistakes. Managers may downplay private information that an innovation is ineffective because they observe it being used by many different firms that are all highly successful. The weight of the evidence suggests to them that their private failings may reflect improper implementation, a need for more time, or temporary bad luck. Like losers at the slots, managers may assume they are unlike the others, suppress disclosure of the poor results they experience, and behave so as to perpetuate the illusion that the innovation is working.

We wanted to know if this self-reinforcing bias in access to information could induce empirically plausible cascades of diffusion and abandonment of worthless innovations, without any assumption of conformist behavior on the part of managers. We hypothesize that confirmation bias will generate information cascades that resemble the fadlike behaviors produced by conformity. If so, then the willingness of managers to invest in innovations with little or no intrinsic value need not mean that they are "dedicated followers of fashion." We suspect the problem may be quite the opposite—that they are dedicated followers of *success*.

More precisely, we hypothesize a fragile equilibrium, characterized by an unpredictable "bandwagon" followed by sudden and unexpected collapse. This fragility stems from the tension between accurate perceptions of individual failure and biased perceptions of the success of others. We hypothesize that fadlike cascades of innovation are possible even when the innovations are all but worthless.

We further ask whether the appearance of fadlike behavior is robust across a range of assumptions about the cost-effectiveness of competing approaches (expanding on Strang and Macy 1999, where such costs were not considered). As the cost of an innovation increases, we assume that managers become more reluctant to abandon the investment without convincing evidence that it has failed. This chapter examines the net impact of these opposing tendencies on cascades, and how this interacts with the intrinsic effectiveness of innovations.

To test these predictions, we developed an agent-based computational diffusion model with free parameters for the cost and effectiveness of innovation. We then conducted a series of experiments in which we manipulated these parameters and observed the distribution of innovations across a population of firms in an artificial world.

Simulation Model and Experimental Design

Our decision model entails behavioral assumptions that are surely controversial. Our strategy in making these assumptions was not to be "realistic" but to create a conservative test of our theory that fadlike behavior can be caused by a preoccupation with performance, even when innovations are worthless or nearly so. We therefore assume away any consideration of popularity, not because we think managers are immune to such influences but because we want to see if fadlike cascades will persist even if such immunity should obtain.

Instead, we assume that managers are performance oriented and skeptical of hype. We also assume that the obsession with performance leads to a focus on winners rather than losers, both by managers and the business press.

We model a population of 100 firms that are competing with one another to attain valued economic outcomes, such as revenues, profits, return on investment, dividends, stock appreciation, growth in market share, etc. We abstract these outcomes as a single score that indicates the firm's relative performance in each round. This score depends on three factors: the firm's market position, choice of innovation, and luck. Formally, outcome O for firm i at time t is given by:

$$O_{it} = \alpha K_i + \beta S_{it} + (1 - \alpha - \beta)e_{it} \quad (1)$$

where K is market position, S is the performance score of innovation s, e is noise, and α and β are weights that determine the relative importance of K, S, and e in the determination of performance. All parameters lie within the interval [0,1] inclusive. K, S, *and* e are randomly distributed over i, while α and β are constants, such that $\alpha + \beta \leq 1$.[4] Although the distribution of market position is fixed for the duration of the experiment, innovation and luck can change from round to round. For simplicity, we assume that innovations are implemented without any time lag and that the duration of t is sufficient to implement a new practice and experience its effects on earnings.

Market position refers to a firm's comparative advantage prior to the firm's decision to innovate. These advantages reflect fortuitous structural or contextual factors that are exogenous to our model and cannot be affected by a firm's decisions.

We assume that there are initially as many possible innovations as there are firms.[5] After that, each firm must choose an innovation each round, where these vary in effectiveness. We begin with the assumption that innovations are entirely worthless ($\beta = 0$), such that outcomes depend entirely on market position and luck, with the latter two factors weighted equally. We then increase the relative importance of innovations ($0 < \beta < 0.5$), while holding market position constant ($\alpha = 0.5$),[6] such that firms increasingly "make their own luck." At maximum effectiveness ($\beta = 0.5$), performance is determined equally by market position and management choice, and luck plays no part.

Choice of Innovation

The decision to innovate is made in two steps: search and adoption. First, the firm decides whether to abandon its current practice and search for an alternative. If the firm decides to retain the current practice, the decision process is over. However, if the firm abandons current practice, then it moves on to the second step, the adoption decision.

The search decision. Firms base the decision to abandon current practice on their direct experience. The more success they experience with current practice, the less likely they will be to abandon it. The experience of other firms is irrelevant; we assume that when managers have personal knowledge about performance, they do not take into account the experience of others.[7]

We assume firms evaluate current practice based on outcomes (O_i in equation 1). If performance improves, the firm is pleased. If performance declines, the firm is disappointed. Performance is measured as the mean outcome obtained while using innovation s during the previous M rounds, where M is the limit of memory. We assume that M is a population parameter that is constant across firms.[8] Thus, mean performance is based on the outcomes that i recalls with s, aggregated over available memory.

$$\overline{O}_{ist} = \frac{\sum_{m=t-M}^{t} o_{ism}}{\sum_{m=t-M}^{t} B_{ism}} \tag{2}$$

- O_{ist} is i's mean outcome using s at time t, $0 \leq$ - $O_{ist} \leq 1$. B_{ism} is a Boolean operator that tests for firm i's use of s at moment m recalled from memory, where $B = [0,1]$, m designates a recalled point in time, and M is the number of events that firms can retain in memory, where $M < t$ by definition (memory cannot go back prior to time $t = 1$).

We then get i's evaluation of s relative to i's experience with the previous innovation p:

$$E_{ist} = \frac{(\overline{O}_{ipt} - \overline{O}_{ist} + 1)}{2} \tag{3}$$

such that $E_{ist} = 0$ if s is optimal relative to p and $E_{ist} = 1$ if s is maximally disappointing relative to p.

Further, we assume that firms are highly skeptical, that is, they assume that an innovation's performance may be a fluke. To test for a fluke, skeptics weight mean performance by the number of observations on which the evaluation is based. Let C_{ist} represent i's confidence in s, based on its aggregate experience with s:

$$C_{ist} = \frac{\sum_{m=t-M}^{t} B_{ism}}{M} \tag{4}$$

The greater i's experience with s, the greater i's confidence in its evaluation.

In addition, we assume that the cost (δ) of implementing innovations affects firms' inertia, or reluctance to innovate, where $0 \leq \delta \leq 1$.

We experimented with two different assumptions about the effects of cost on search behavior. The first specification assumes "satisficing" behavior, such that firms stop searching when they find a good solution, even if it costs nothing to continue searching. That is, firms "exploit" rather than "explore" (March 1991).

The probability $\Pr(D_{ist})$ that i will drop s and search for an alternative is then a function of i's evaluation of s (from equation 3), weighted by i's confidence in this evaluation (given the length of i's experience with s) and i's reluctance to change (given the cost of implementation):

$$\Pr(D_{ist}) = E_{ist}(1-\delta)C_{ist} \tag{5}$$

where D is a Boolean operator for the binary decision ($D_{ist} = [0,1]$).

Equation 5 means that firms tend to satisfice. If i is highly satisfied with s ($E_{ist} \approx 0$), i is unlikely to search, regardless of i's confidence in the evaluation or the cost of innovation. If i is highly dissatisfied with s ($E_{ist} \approx 1$) and highly confident in the evaluation ($C_{ist} \approx 1$), i is likely to search, but only if the cost (δ) is relative low. If the cost is high, i is highly reluctant to drop s, even if s is performing badly. And if i lacks confidence in its negative evaluation of s, i postpones the search until more experience with s is accumulated.

We also experimented with an alternative specification that relaxes the assumption that firms satisfice. The functional form makes the effect of cost additive rather than multiplicative:

$$\Pr(D_{ist}) = \big((1-\omega)E_{ist} - \omega\delta + \omega\big)C_{ist} \tag{6}$$

where ω is a free parameter for the relative weight given to performance and cost in the decision whether to search. (We experimented with $\omega = 0.5$ and $\omega = 0.33$ and found little difference; the reported results use the latter value.)

With this specification, if i is highly dissatisfied with s ($E_{ist} \approx 1$), i is likely to search only if i is also confident in that evaluation and the cost is very low. This is equivalent to equation 5. Conversely, if i is highly satisfied with s ($E_{ist} \approx 0$), i is unlikely to search if the cost is very high, regardless of i's confidence in the evaluation. Put differently, confidence is needed to *change* practices but not to *retain* them. Having invested in an innovation, skeptical firms require cumulative disappointment over several periods before concluding that an investment was ill advised. This is also equivalent to equation 5.

However, if the cost is very low, i is tempted to search (with probability approaching ω) even when i is satisfied with outcomes experienced under s. In other words, compared to equation 5, equation 6 assumes that i has a residual tendency to "explore" when exploration is costless. This is a highly unrealistic assumption, given what we know about organizational inertia (Hannan and Freeman 1984), but it affords a more conservative test of the formation of cascades, given that firms are now relatively more likely to jump off the "bandwagon." We therefore report results using equation 6 and only note that, as expected, cascades were slightly more robust with the equation 5 specification, but the differences were too small to warrant elaboration.

The adoption decision. If a firm chooses to drop s ($D_{ist} = 1$), it then looks around for "best practice." The "best practice" in our model is simply the innovation being used by the top performing firm.[9] This selection rule is dictated by constraints on access to information and a cognitive bias toward confirmatory evidence. We assume that firms have very limited information about the strategies and performance of other firms. Given this constraint, a cognitive bias toward confirmatory evidence makes firms most interested in learning which innovations were used by "winning firms," that is, those that outperform their competitors. In addition, we assume that firms are more likely to announce their successes than their failures, and actors seeking to promote an innovation (managers inside the firm or outside consulting agencies), are more likely to advertise if the firm's performance was outstanding than if it stumbled. That is, promoters can be expected to announce the winner's "secret" if the firm succeeds, but to keep quiet if the firm loses. These two factors, interest in *acquiring* and *promoting* the innovations used by top performers, combine to create a strong bias in access to knowledge used in the evaluation of innovations. For simplicity, we assume that all firms know the innovation used by the top performer in each round, and that they store this information for later recall.

We assume that firms are skeptical about "secrets of success" used by other firms. Before adopting a seemingly successful innovation, skeptical managers first reject three alternative possibilities:

1. That an innovation's success was a random fluctuation (a "lucky break").
2. That a top firm's performance was due to some other attribute besides the use of the innovation.
3. That the innovation may work for others but not for firms such as itself.

To test for a fluke, firms check to see how frequently an innovation is associated with a given outcome. They do not attach much weight to a single "win." But suppose an innovation is identified as "best practice" two times in succession; now a few dissatisfied managers may begin to take the association between use of the innovation and a firm's success more seriously. And if an innovation looks like a "winner" three times in a row, it may soon become the talk of the town.

More precisely, let B_{wm} be a Boolean operator ($B_{wm} = [0,1]$) that tests whether a winning innovation w was used by the top-performing firm at moment m recalled from memory. Agents then search their memory to find the relative frequency f_{wt} at which w has repeated as a winner (i.e., $B_{wm} = 1$) during M periods just prior to t:

$$f_{wt} = \frac{\sum_{m=t-M}^{t} B_{wm}}{M} \tag{7}$$

To test for a spurious association, managers check to see how many different firms F have had success using w. The more firms that have used w with success, the higher the probability that w, and not some other attribute of the firm, is responsible for the firm's success. The calculation is the same as equation 7 except here B_{Fwm} tests whether firm F used w and F is a nonredundant member of the set of winning firms all using this same innovation. If the firm is not a member of the set (because the firm did not use w) or is a *redundant* member (because that same firm won more than once using w) then $B_{Fwm} = 0$, otherwise $B_{Fwm} = 1$.

$$f_{Ft} = \frac{\sum_{m=t-M}^{t} B_{Fwm}}{M} \tag{8}$$

Finally, to check for the generality of the association, firms discount innovations that they have previously tried without success. The more recently it was tried, the lower i's confidence in w:

$$C_{iwt} = 1 - \frac{B_{iwm}}{t-m} \tag{9}$$

where B_{iwm} tests for prior use of w at time m, the most recent time at which i recalls using w (by definition, $m = 0$ if w is not recalled). Equation 9 means that i discounts w's value if i recalls using w recently.

The probability that i adopts w ($A_{it} = 1$) is then given by the results of these three tests:

$$\Pr(A_{it}) = C_{iwt} f_{wt} f_{Ft} \tag{10}$$

where f_{Ft} is the number of firms that won with w in the M time periods just prior to t, as a proportion of the number of winners stored in memory, and f_{wt} is the number of times w has repeated as winner (regardless of the identity of the firm), as a proportion of M. If $A_{it} = 1$, then i adopts w. If $A_{it} = 0$, then i innovates without regard to what other firms are doing (in natural settings, this might correspond to internal research and development). We assume that such choices are random over the population of possibilities.[10]

Suppose that relative performance has nothing to do with either innovation or market position and is entirely a matter of luck. With an initially uniform distribution of innovations across firms, it is very unlikely that randomly selected winners will cluster around particular innovations. However, it is only a matter of time until, by chance, an innovation wins twice. Once that happens, there may be sufficient new adoptions such that the innovation then wins a third time, perhaps unleashing a cascade that can be expected to sweep through the population.

This simple dynamic is complicated by a second consideration. Firms are assumed to evaluate innovations based not only on the "vicarious experience" of successful others, but also on their own individual experience. Given that popular innovations may be entirely worthless, initial enthusiasm can be expected to decline over time for most firms. In short, individual experience can be expected to drive the distribution of innovations toward uniformity (all have equal probability of adoption), while vicarious experience can be expected to periodically drive the distribution toward conformity. We use computer simulation to explore how these two tendencies interact in populations that differ in the cost and effectiveness of innovations.

Simulation Results

We used four measures of outcomes. Two of these measure the amplitude and periodicity of fads. Amplitude measures the proportion of the population using the leading (or most popular) innovation and ranges from $1/N$ (every firm uses a different innovation) to 1 (all firms use the same innovation). Periodicity measures the average duration of a fad and the rate at which cascades collapse. Duration ranges from 1 (the leading innovation changes at every iteration) to 1000 (the leading innovation never changes over the 1000 iterations of the experiment.) We assume that fadlike behavior is most prominent when the leading innovation increases rapidly in amplitude (a cascade), remains stable for a period of time that exceeds thatof the cascade, and then collapses. Thus, fadlike periodicity involves neither very low and nor very high duration.

We also have two performance measures (see equation 1): average outcome scores ($\bar{O}$) over the population of firms, and the performance score (S_l) of the leading innovation relative to the best-performing innovation, ranging from 0 (the leading innovation is the worst-performing innovation) to 1 (the leading innovation is the best-performing).

Figures 1–4 track these four measures over 1000 iterations for a population of 100 firms, under four combinations of cost ($\delta = 0$, $\delta = 1$) and

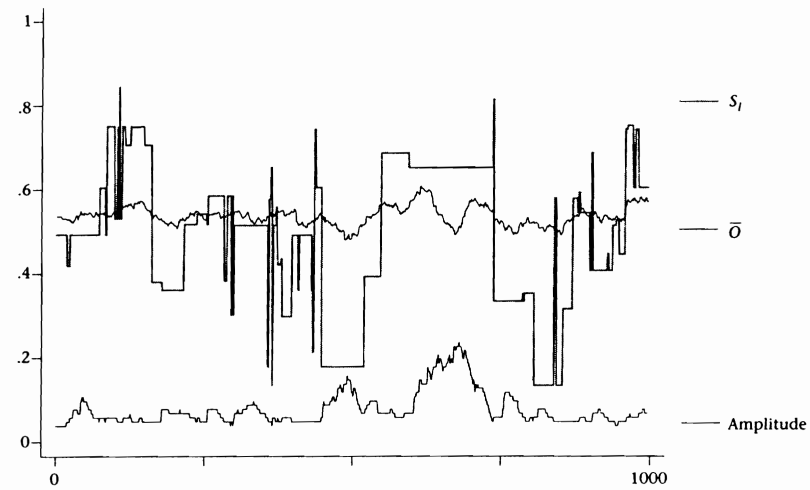


Figure 1. Amplitude and periodicity of cascades with worthless and costly innovations.
($N = 100$, $\alpha - 0.5$, $\beta = 0$, $\delta = 1.0$)

effectiveness ($\beta = 0$, $\beta = 0.05$) of innovations. We assume that 50 percent of performance is always due to a firm's market position ($\alpha = 0.5$), with luck counting for 50 percent (when $\beta = 0$) or 45 percent (when $\beta = 0.05$).

Figure 1 shows that cascades are possible even with innovations that are completely worthless yet costly to implement. The most prominent cascade begins about iteration 600 and spreads to about 20 percent of the population before collapsing by iteration 700. Not surprisingly, the spread of this innovation has no effect on average performance, which remains at the value expected by random chance ($\bar{O} = 0.5$). The changes in leading innovation are indicated by sharp changes in the value of *SI* (the performance score associated with leading innovation *I*). These changes are tracked by the gray pen-color and show very rapid turnover in leading innovations, with an average duration of 15 iterations. Clearly, while fads can occur, they are not robust when innovations are costly and have no intrinsic value.

Figure 2 shows that reducing the cost of innovation does not make fads more robust. On the contrary, there is even more rapid turnover of innovations, and no cascade reaches more than 15 percent of the population.

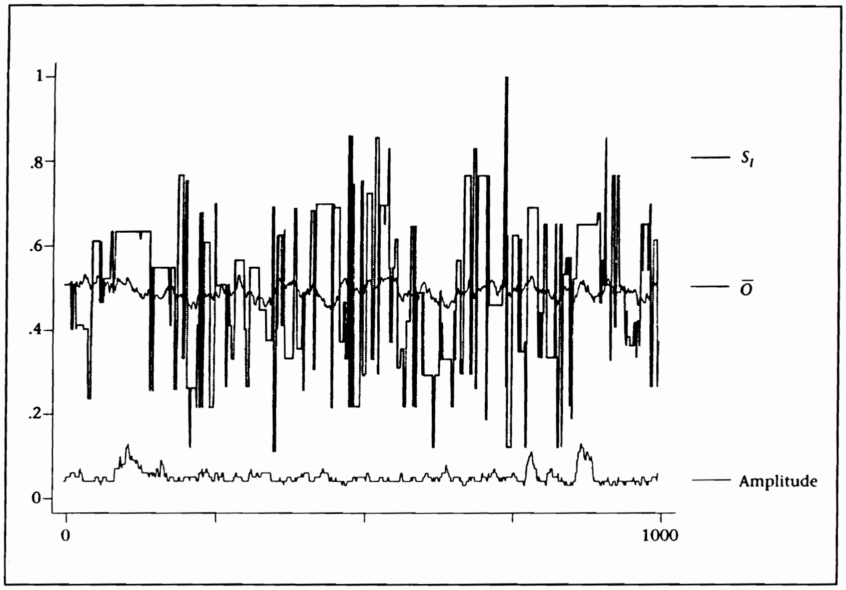


Figure 2. Amplitude and periodicity of cascades with worthless and costless innovations.
($N = 100$, $\alpha - 0.5$, $\beta = 0$, $\delta = 0$)

We now slightly increase the effect of innovations on firm performance, from $\beta = 0$ to $\beta = 0.05$, with cost at the maximum ($\delta = 1$). Figure 3 presents a striking result: cascades become highly robust, with one innovation reaching nearly half the population and lasting nearly 200 iterations. During this period, managers identified one of the best innovations (SI = 0.85), but with $\beta = 0.05$, the marginal effect on $\bar{O}$ is quite modest, as indicated by performance scores reported in figure 3.

Figure 4 shows that fads are much more prominent when innovations are relatively expensive (compared to figure 3). When innovations are costless ($\delta = 0$), there is little effect of an increase in β from 0 to 0.05. Paradoxically, the less costly the innovation, the greater must be the effectiveness for cascades to develop. This reflects our assumption that managers have some nonzero probability of searching, even when performance is strong, if it costs them little or nothing to try (equation 6).

Figures 5–8 show in greater detail how cost and effectiveness interact in shaping fadlike behavior. To study the interactions, we crossed these two factors using a 6 by 11 factorial design: six levels of effectiveness (β = [0...0.5] in increments of 0.1) and 11 levels of cost (λ = [0...1] in increments of 0.1).

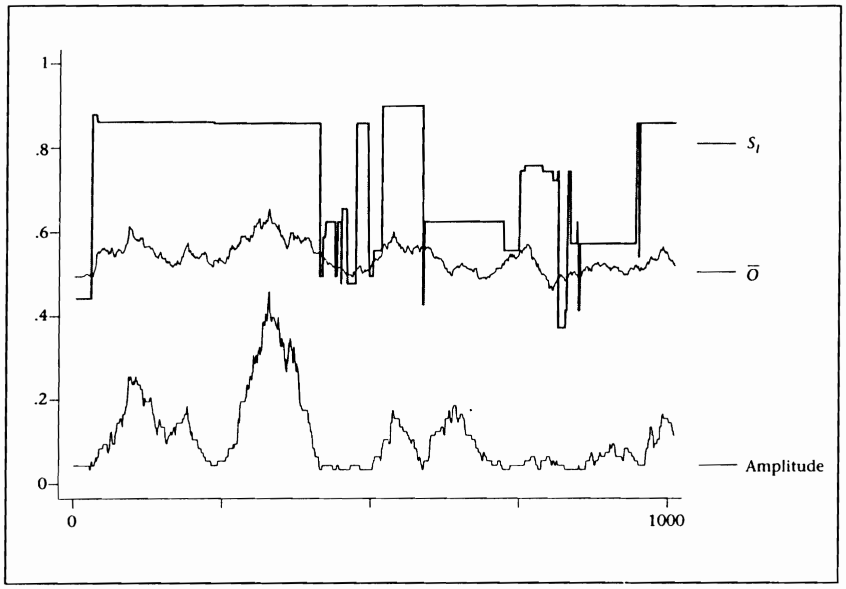


Figure 3. Amplitude and periodicity of cascades with almost worthless and costly innovations.
($N = 100$, $\alpha - 0.5$, $\beta = 0.05$, $\delta = 1.0$)

Figure 5 reports the amplitude of information cascades over the range of cost-effectiveness. Note that effectiveness can also be read as luck; outcomes are determined 50 percent by luck when innovations are entirely worthless and luck plays no part when innovation and market position count equally (50 percent). Cascades reach maximum amplitude (a 45 percent rate of adoption) when innovations and luck count about equally and innovations are maximally costly to implement. Of course, if innovations are entirely worthless, there is nothing to jump-start the cascade. Fad amplitude increases with the effectiveness of innovations, although the latter effect is nonmonotonic, with maximum amplitude at about 0.3 effectiveness. At this level, outcomes are determined 50 percent by market position, 30 percent by innovation, and 20 percent by luck. As the value of innovation pushes luck into the background, cascades become rapidly dampened, especially when innovations are maximally costly to implement. These results suggest that randomness is as important for generating cascades as the intrinsic merit of the innovation. Combined with high effectiveness, randomness allows new (lucky) firms to emerge as winners using the same winning innovation. When luck no longer plays a role, the same market leader always wins, and skeptics cannot rule out the possibility

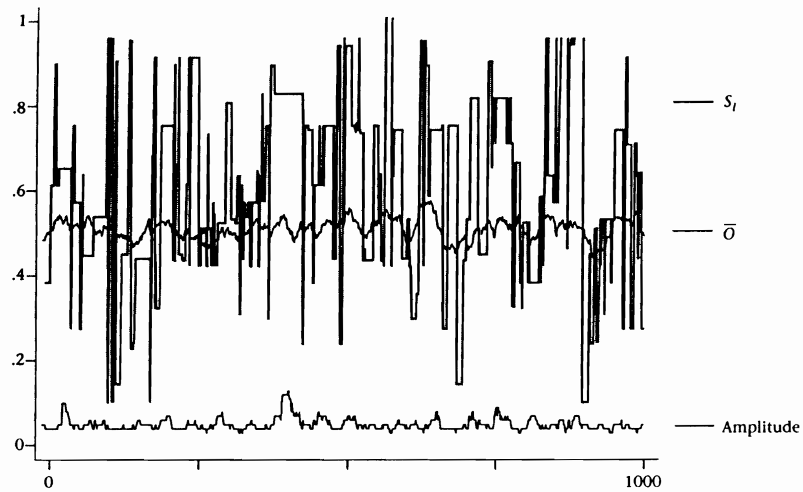


Figure 4. Amplitude and periodicity of cascades with almost worthless and costless innovations.
(N = 100, α – 0.5, β = 0.05, δ = 0)

that the performance of the innovation was a fluke (from equation 8). This skepticism becomes especially important as the cost of implementation increases inertial tendencies. High cost amplifies the dampening effects of skepticism. Yet high cost has a positive effect on amplitude, holding randomness constant. Reluctance to abandon "sunk costs" keeps firms from jumping off the bandwagon.

Figure 6 shows how the cost and effectiveness of innovations affect the duration of cascades measured in iterations. As might be expected, duration is highly correlated with amplitude (r = 0.7). Unstable cascades do not have time to spread, and those that do not spread are less durable. However, overly stable cascades are not fadlike either. Fads occur when no single innovation is hegemonic. Figure 6 is important because it shows that fadlike behavior occurs over most of the parameter space. So long as innovations are neither worthless nor costless, cascades will be sufficiently durable to allow wide diffusion, but not so durable that cascades cannot also crash. Durability ranges from about 5 iterations at zero cost and zero effectiveness, to 280 iterations at maximum cost and 30 percent effectiveness. The average duration over the entire parameter space is about 80 iterations.

Figure 7 reports changes in average firm performance as the cost and

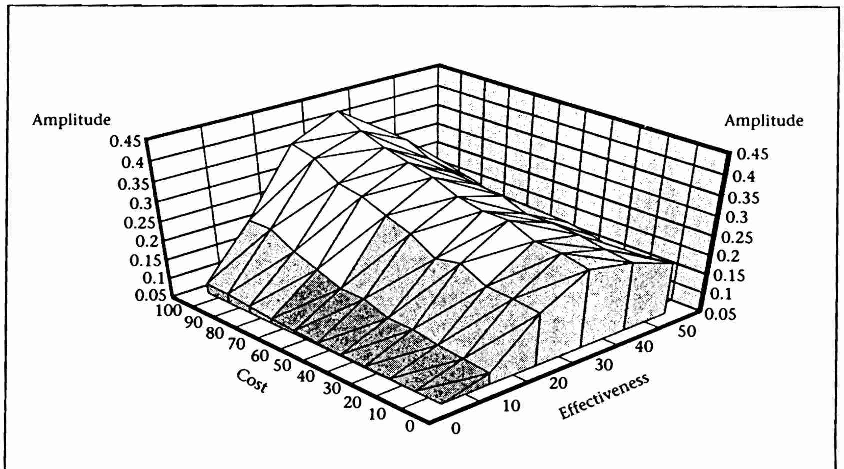


Figure 5. Cost, effectiveness, and amplitude of leading innovations.

effectiveness of innovations increases. In effect, this measures the ability of managers to identify "better mousetraps." Not surprisingly, this ability is absent when innovations are worthless, regardless of cost. However, as innovations become more effective, overall population performance improves, but only up to a point. For β >0.3, overall performance actually declines slightly as effectiveness increases, especially if innovation is costly. There are two possible explanations: (1) leading firms may not be using the best innovations or (2) other firms may not be copying the leaders. Figure 8 confirms the latter explanation. At β =0.3, the leading firms are using "best practice" almost all the time, with little or no improvement above the 30 percent mark. The problem is that other firms become less likely to copy these leaders, especially as the cost of innovation increases, as already apparent from figure 5. As innovation (and not luck) becomes a strong determinant of performance, the same market leaders always seem to come out on top. The lack of turnover in the winner's circle makes it harder to persuade skeptics that the choice of innovation (and not market position) is what really matters.

Discussion

Diffusion models of the adoption of innovations tend to focus on two contradictory explanations. "Performance models" assume that popular innovations are adopted and retained because they confer legitima-

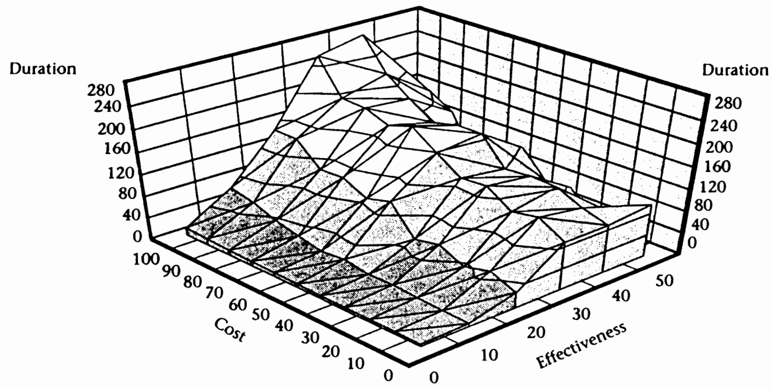
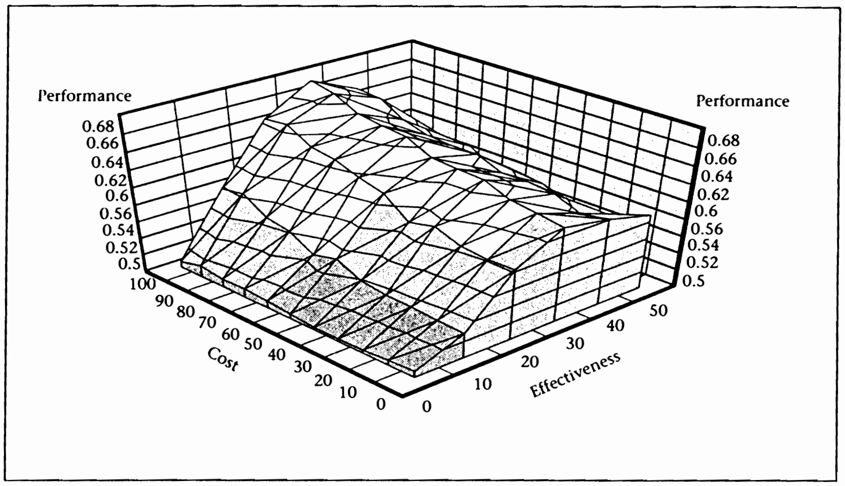


Figure 6. Cost, effectiveness, and duration of leading innovations.

cy, provide efficient solutions to common problems, or economize on information or communication costs. In contrast, "conformance models" assume that popularity builds upon itself, even if popular innovations are inferior in performance.

In this chapter we have proposed an alternative framework that treats adoption and abandonment decisions as intendedly rational but shaped by limited and biased information. Like the performance models, ours assumes that managers are paid to find "best practice," not to follow the pack; lemmings are not likely to find bonus money in their end-of-the-year pay envelopes. At the same time, like conformance models, ours assumes that managers make systematic and self-reinforcing cognitive errors. However, the error we investigate is not about the source of knowledge but about its confirmation. Epistemic bias entails the assumption that public knowledge has been repeatedly and privately evaluated, overlooking the possibility that others likewise assume away the need for independent evaluation. Confirmation bias centers on a tendency to stop testing an hypothesis once it has been confirmed. If all winners use innovation *w*, then success can be attributed to *w*, without testing whether losers also use that same innovation. This cognitive tendency is reinforced by structural asymmetry in access to knowledge about the innovations used by winners and losers. Decision-makers are provided with a restricted and biased set of examples: they know their own experience and that of their most successful competitors, but not the experience of the many also-rans. When their own experience contradicts the externally observed pat-



Figures 7. Cost, effectiveness, and firm performance.

tern, managers may attribute failure to situational factors rather than to the innovation.

This simple model builds on the intuitive idea that managers "follow the leader" (Haveman 1993), based on the assumption that an innovation used by several leading firms is likely to be worthwhile.[11] Admittedly, our model abstracts away important elements in managerial decision-making, including risk-aversion (Kahneman and Tversky 1990), and important departures from rational decision-making (Gilovich 1993, Russo and Schoemaker 1990). For example, we do not consider the managerial problem of identifying "best practices" internally within the organization (O'Dell et. al 1998). However, this simplification provides a stronger test of our claim that fadlike cascades can be generated without the assumption of conformist behavior.

We also abstract away from structural complexities. We assume population homogeneity: innovations are modeled as making the same contribution to each adopter's outcomes, rather than allowing the utility of innovations to vary across firms. We also assume managers never test whether identical innovations might also be used by nonwinners. More reasonably, all managers have some disconfirming knowledge, and managers vary in their awareness of the logical necessity for this evidence. It seems best, however, to begin by investigating simple forms of the model before moving to a wider set of complications.

The computational experiments establish an important existence proof: a self-reinforcing bias in access to information is sufficient to generate managerial fads, even among managers who are rational,

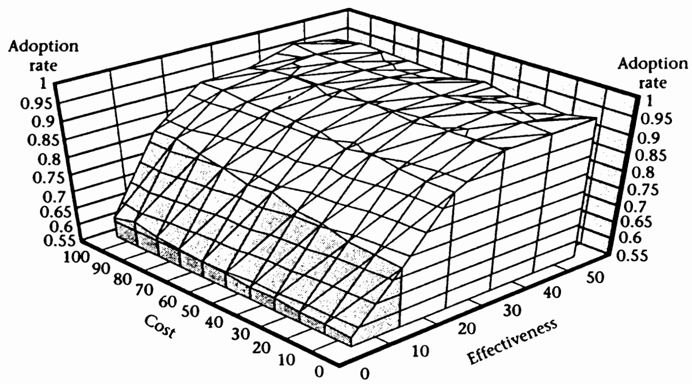


Figure 8. Cost, effectiveness, and adoption of "best practice."

skeptical, and nonconformist, even when innovations are worthless, and especially when innovations are costly to implement. Further, we investigate how the cost and effectiveness of innovation shapes the amplitude and duration of fads and the performance of firms. Convergence occurs when leading innovations are worthless as well as when they are worthwhile. For example, as illustrated in figure 1, even when innovations are costly yet have no impact on randomly generated outcomes, a cascade appeared that spread to about 20 percent of firms who converged on a single innovation.

Examination of the costs of implementation suggests a paradoxical result. Where innovations are worthless, higher costs make for greater stability as firms are slow to discard ideas in which they are heavily invested. But where innovations can be moderately effective, high costs increase rather than decrease the amplitude of fads. Here, slower turnover permits intrinsically effective innovations to demonstrate what they can do. Only where innovations are highly effective do costs serve to dampen cascades.

We recognize that an "existence proof" is just that. Computational models are numerical, not deductive, and unlike mathematical models, they cannot be used to form generalizations (Holland 1995, p. 100). However, they are more suitable to the study of nonlinear processes such as the explosive growth and unpredictable collapse of cascades. In complex systems, very simple rules of interaction can produce highly complex global patterns. Although this principle has been mainly applied to physical and biological systems, Simon draws out the striking implication for the study of human social interaction—the possibility that the complexity of social life is a global, not a local property (Simon 1998, p. 53). "Human beings," he contends, "viewed as behaving systems, are quite simple." We follow rules, in the form of norms, conventions, protocols, moral and social habits, and heuristics. The idea is that complexity emerges in our interactions with each other, not in the rules that govern much of our everyday behavior. Although the rules may be quite simple, they can produce global patterns that "may not be at all obvious" and are very difficult to understand (Axelrod 1997, p. 5). "The apparent complexity of our behavior," Simon concludes, "is largely a reflection of the complexity of the environment." In just this sense, we would argue that the global pattern of boom and bust is an emergent property of agent *interaction*, and not a reflection of a mercurial property of the agents themselves.

Even so, we emphasize that the stylized model of behavior presented here slights the rich texture of managerial decision-making. Our aim in this study is exploration, not prediction. The exploration of artificial worlds "can contribute to knowledge about unforeseen and unknown effects" and "possible alternatives to a performance observed in nature" (Conte and Gilbert 1995, p. 4). Agent-based modeling "does not necessarily aim to provide an accurate representation of a particular empirical application. Instead, the goal of agent-based modeling is to enrich our understanding of fundamental processes that may appear in a variety of applications" (Axelrod 1997: 25). It is our hope that these results might be applied in precisely that spirit.

Notes

1. The reverse relation obtains in the diffusion of technological innovations. These tend to entail convex payback functions that save most of the rewards for the late adopters. For example, email and telephones have little benefit if only a few people use them, cars get bogged down in mud-holes if too few people drive to warrant the expense of paved roads, and so on.

2. For reviews of the literature on diffusion and vicarious learning, see Strang and Soule (1998) and Miner and Raghavan (1999).

3. For the cognitive difficulty that decision-makers have incorporating feedback effects into predictions of future outcomes, see Paich and Sterman 1993.

4. Although outcomes are also affected by the cost of innovation, but we simplify the model by assuming that costs are identical across innovations, such that relative performance is affected only by their effectiveness. We focus instead on how costs affect firms' readiness to abandon practices whose effectiveness is uncertain. Thus, we manipulated cost as a treatment condition affecting all firms equally (see below). To test for robustness, we also allowed cost to vary across innovations (orthogonal to effectiveness) and obtained similar results.

5. We experimented with ratios greater and less than 1 and found no important effects.

6. In another study (Strang and Macy 1999), we manipulate market position and managerial skepticism, while holding constant the cost of implemention. We also study the effects of organizational inertia, an important factor in ecological models (Hannan and Freeman 1984).

7. We experimented with an alternative specification in which the firm evaluated current practice by comparing it not only to its experience with previous practice but also to the performance of other firms, with equal weight given to both assessments. This meant that firms with below-average performance were more likely to search. Incorporating this "sideways comparison" alongside the "backward-looking" comparison (from equation 2) made no difference in the results, so we present the results using the simpler specification.

8. We tested robustness by allowing M to vary and found no effects so long as $M > 5$.

9. Our model is thus related to Mezias and Lant's (1994) formalization of mimetic search, but we assume that imitators follow the high performer while Mezias and Lant assume imitators follow the largest firm (where size is in part a function of past success). One might also take geographic/product-based similarity to exemplary firms into account, but we leave issues of network structure for future investigations.

10. Alternatively, i might pick the innovation used by the next-best performer, rather than choosing randomly. However, random selection is a more conservative specification in that it refreshes the heterogeneity of the distribution of innovations in the population. If we can get fads to emerge with this specification, we can expect fads to emerge by having firms pick innovations from a group of top performers, a much smaller subset of the population of firms.

11. By "leader," Haveman means an early entrant into a new market whom imitators then follow. In contrast, we refer to a market leader or top-performing firm whose "winning" innovation others copy.