On the Evolution of Overconfidence and Entrepreneurs

Antonio E. Bernardo

UCLA Anderson Graduate School of Management
Los Angeles, CA 90095
antonio.bernardo@anderson.ucla.edu

Ivo Welch

Yale School of Management and NBER
New Haven, CT 06520-8200
ivo.welch@yale.edu

This paper explains why seemingly irrational overconfident behavior can persist. Information aggregation is poor in groups in which most individuals herd. By ignoring the herd, the actions of overconfident individuals ("entrepreneurs") convey their private information. However, entrepreneurs make mistakes and thus die more frequently. The socially optimal proportion of entrepreneurs trades off the positive information externality against high attrition rates of entrepreneurs, and depends on the size of the group, on the degree of overconfidence, and on the accuracy of individuals' private information. The stationary distribution trades off the fitness of the group against the fitness of overconfident individuals.

> Starting any company is really hard to do, so you can't be so smart that it occurs to you that it can't be done.
>
> —Kathryn Gould, Foundation Capital, Menlo Park, in *GSB Chicago Magazine 21-3, Summer 1999*

1. Introduction

According to DeBondt and Thaler (1995), "Perhaps the most robust finding in the psychology of judgment is that people are overconfident." Such overconfidence induces individuals to undertake ventures

The second author may maintain a Web page at http://welch.som.yale.edu for additional information (if available). We are grateful for comments from Jennifer Aaker, Robert Boyd, Kent Daniel, David Hirshleifer, Jack Hirshleifer, Josh Lerner, Steve Lippman, John Mamer, Bill McKelvey, Terry Odean, Richard Roll, Larry Samuelson, Lars Stole, Avanidhar Subrahmanyam, Karen Stephenson, Richard Thaler, Karen Van Nuys, and workshop participants at the University of Minnesota and Harvard Business School.

Journal of Economics & Management Strategy, Volume 10, Number 3, Fall 2001, 301–330

that more rational individuals might not undertake. For example, overconfidence among economic entrepreneurs has been documented by Cooper et al. (1988). In their sample of 2994 entrepreneurs, 81% believe their chances of success are at least 70%, and 33% believe their chances are a certain 100%. In reality, about 75% of new businesses no longer exist after five years. Busenitz and Barney (1997) compared entrepreneurs' and managers' assessments on a set of real-world questions (e.g., whether cancer or heart disease is the leading cause of death in the United States). Entrepreneurs and managers were about equal in their accuracy, but the level of confidence of entrepreneurs in their own answers was dramatically higher. The question our paper tries to address is if economic principles can offer an explanation for such relatively common overconfident behavior, which has clearly and reproducibly been documented in laboratory settings to be irrational. In addition, while overconfidence and entrepreneurship are important phenomena in themselves, there is another motivation for studying overconfidence: Some recent work in economics and finance (e.g., Delong et al., 1991; Daniel et al., 1998; Odean, 1998) relies on overconfidence as an underlying primitive assumption, often with little theoretical justification as to why such irrational behavior can persist.

Our paper offers a simple explanation for the presence of overconfidence and entrepreneurs. Our main argument is that overconfident entrepreneurs (*independent spirits*, *innovators*, *leaders*, *change agents*, or even *dissidents*) are less likely to imitate their peers and more likely to explore their environment. Entrepreneurial activity can thus provide valuable additional information to their social group.

Our point holds when individual actions can convey valuable private information and when information aggregation within the overall group is otherwise poor. Our specific modeling framework is built on the concept of *informational cascades*, introduced in Welch (1992), Banerjee (1992), and Bikhchandani et al. (1992). In this context, individuals can observe one another and typically end up following the same action, yet information aggregation is poor because rational, nonentrepreneurial individuals who follow "the herd" reveal nothing about their private information. From a social perspective, cascades lead to a suboptimal level of information disclosure, experimentation, and exploration of the environment.

When overconfident, entrepreneurial individuals instead follow their own information, downweighting the information in the herd, their actions in effect *broadcast* their private information to the rest of their group. Unknowingly, overconfident entrepreneurs behave altruistically, making irrational choices that are to their own detriment but help their groups. Because the herd carries relatively little information, this is only mildly individually suboptimal. Still, we would not expect entrepreneurs to reflect very deeply on their actions. Instead, we would expect them to be either socially or biologically "programmed" to overestimate the quality of their own information. Indeed, the presence of such overconfident individuals who act on their own information and who irrationally ignore the actions of other individuals in the group has already been demonstrated in laboratory settings by Anderson and Holt (1996). Our model can easily be calibrated to generate benefits to their group that are larger by a factor of 100 than the cost to the individual.

In Section 2, we identify conditions under which the benefits of entrepreneurship to the group are high and the costs to individual entrepreneurs are low. We then show that when groups compete and inferior groups disappear, groups with some entrepreneurial activity may gain enough of an evolutionary advantage to permit entrepreneurs to survive in equilibrium. Our paper therefore argues that groups with some overconfident individuals have an evolutionary advantage over groups without such individuals. In Section 3, we derive a stationary distribution in which overconfidence persists across generations. This distribution trades off the relative fitness of the group against the relative fitness of the (altruistic) individuals who are overconfident.

Our paper identifies some of the forces important to the relative benefits and costs of being an entrepreneur, and to being a group, culture, society, or firm that fosters or handicaps entrepreneurship. For example, we find that overconfidence/entrepreneurship can be useful if groups are large enough to benefit from the positive information externality, if individuals have low-precision information, and if overconfidence is moderate rather than extreme. There are of course other important aspects to entrepreneurship that are not modeled by our paper, and not every entrepreneur behaves irrationally *ex ante* (e.g., Manove, 1988).

There are surprisingly few papers that explicitly adopt evolutionary selection (e.g., Becker, 1976; Ainslie, 1975; Hirshleifer, 1977; Hirshleifer, 1987; Hirshleifer and Martinez-Coll, 1988; Waldman, 1994; Rogers, 1994; Hirshleifer and Luo, 2001; Wang, 2001), and fewer yet that invoke group selection. As far as we know, models of group selection have appeared only outside economics. Section 3 discusses the history of arguments pro and con group selection. We believe a deviation from the individual optimization paradigm in our context to be necessary:

1. There are many well-documented psychological inference biases that are intrinsically difficult to defend as being in the interest

of the individual. Although some biases can be explained with individual-centric explanations (e.g., Hirshleifer, 1987), explanations for *inference biases* should recognize that they are *inference distortions*, and individual behavior should follow directly from the inference process. This is perhaps best to explain in the context of overconfidence. When cornered, most economists tend to argue that well-documented overconfidence (or other biases) could possibly be directly linked to behavior that could enhance individual survival, for example, an increase in aggressive behavior (see Sec. 4). To defend such an argument, one would have to show (i) an empirical linkage between aggression and overconfidence, (ii) why aggressive behavior is optimal in an environment, and (iii) why it is the distorted inference process that creates aggression. In contrast, our paper's explains "following one's own information" directly.

2. *Homo sapiens* is unusual. We are constantly judging how altruistic our peers are, and we are constantly judged by our peers. The ability to exclude individuals from membership in a society, especially when coupled with our long-term memory, can be a powerful force towards social behavior. It should not be surprising that behavior that enhances group survival can play a role in certain social situations; yet, the fact that humans can behave in a socially valuable manner is often neglected in economics. When altruistic behavior is entertained, the economic literature often simply enters it directly into the utility function. We believe group selection can offer a model-disciplining mechanism about which irrational and near-rational behavior may reasonably enter a utility function and which may not.
3. Groups can evolve mechanisms that are surprisingly effective at overcoming a public goods problem. For example, the costs to being overconfident can be trivial, and individual costs can be orders of magnitudes less than the benefits to the group. In many situations, it is difficult to imagine that alternative mechanisms (e.g., cultural mechanisms, such as large-scale coordination by *credible* communication) have lower social or individual costs—aside from the fact that they were not feasible when evolution shaped our psyche.
4. Because economics has been faced with such puzzling psychological biases, it has developed a chasm between a "behavioral literature," which takes documented psychological biases as primitives but rarely offers an explanation for why these biases are so pervasive, and a "rational literature," which de facto argues that only individually rational behavior can survive in an evolutionary or market setting and which consequently often tends to discount even near-rational behavior. The use of optimal group-selection

mechanisms offends both camps equally, but (or perhaps because) it holds the promise of reconciliation between them.

In the end, one goal of our paper is to develop a disciplined approach to the investigation of seemingly irrational inferences and behavior, based on group selection principles. In particular, the point of our paper is not to show that informational cascades can (and have been documented to) be broken by overconfident behavior, but that overconfident behavior creates a large positive externality on public information aggregation and small costs on its perpetrators. In a group selection framework, this allows overconfidence to survive. As such, our theory has the potential both to explain why we are overconfident and to offer new insights into the question under what circumstances overconfidence is most likely to be useful and thus appear. Our paper does *not* propose to deemphasize self-interested behavior and individual selection—indeed, that is the stronger force when payoffs are equal. But group selection, in which the cost to the irrational individual is very low and the benefit to the group is very high, can help us understand *documented individually irrational biases* that are otherwise difficult to explain.

Our paper now proceeds as follows: The formal model in Section 2 derives the socially optimal proportion of entrepreneurs. It is purposely kept as simple and focused as possible. It ends with a brief summary of factors influencing the trade-off between the informational externality and entrepreneurial attrition. Section 3 derives the stationary distribution, i.e., the trade-off between intergroup and intragroup selection. It also discusses arguments pro and con group selection—familiar to readers of the biology literature—as they pertain to our model. Section 4 discusses alternative explanations for the presence of overconfidence and entrepreneurs. Briefly, overconfidence could also be explained as a signal that helps individuals convince others of high ability; entrepreneurship could also be explained as a high-risk but value-maximizing alternative. These explanations are not only different from those advanced in our model, but also (and more importantly) are complementary to our own explanation. Section 5 concludes.

2. The Model

We now develop a simple model to illustrate that overconfidence can impose only small costs on entrepreneurs (individuals that put too much weight on their own information) but provide large benefits in revealing their private information to their groups. Although our specific model is based on the *cascades* framework, it could have equally

TABLE I.
INFORMATION STRUCTURE

	Value State	
Signal Value	$\theta = 1$	$\theta = -1$
H	p	q
L	q	p

well been based on a different framework (e.g., a two-armed-bandit search model). The basic intuition and comparative statics would be similar. Our goal is to show that overconfident behavior can create an externality that improves public information aggregation.

2.1 Available Information

Assume that members of a group of N risk-neutral individuals choose in sequence whether or not to take an action with uncertain value θ. The action is costless, and the true value of θ is either -1 or $+1$, each with prior probability 1/2. No individual can observe the true value of θ, but each individual can privately observe an i.i.d. signal that is correlated with θ. For simplicity, we assume that if $\theta = 1$, then each individual observes a private signal H (high) with probability $1 > p > 1/2$ and a signal L (low) with probability $q = 1 - p$. Thus, if $\theta = +1$, individuals are more likely to observe H. Likewise, if $\theta = -1$, then individuals observe the signal L with probability p and the signal H with probability q. In this setting, a higher value of p implies that the signal is more informative. Table I summarizes the information structure.

The group of individuals is sequenced randomly and exogenously. Each individual chooses a publicly visible action, either to adopt (A), to reject (R), or to abstain from decision. Adopting (rejecting) has higher expected payoffs if it is more likely that the state is $\theta = 1$ ($\theta = -1$). Without loss of generality, we assume that the individual abstains if and only if she is indifferent between adopting and rejecting. There are two types of individuals in this model, who differ only in one respect:

- *Normal individuals* are fully rational in that they base their decisions optimally on both publicly available information and their individual private information.
- *Entrepreneurs* base their decisions on both publicly available and their own information, but they do not put enough weight (in Bayesian terms) on the public information. This definition of overconfidence

conforms to the definition employed in recent finance models (Daniel et al., 1998) and to the findings in Anderson and Holt (1996): Individuals are either *relatively* more skeptical about external information or *relatively* more overconfident about internalized information.[1]

For computational simplicity, we assume that the type of each individual is public knowledge.[2] We define λ to be the proportion of entrepreneurs in the group. Note that λ is not necessarily the incidence of overconfident actions—even an overconfident entrepreneur can find himself in a situation in which the public information is so overwhelming that even he follows it.

2.2 Normal Individuals' Decision Rule

The decision rule for normal individuals optimally uses their private information signal and information contained in the decisions of individuals that arrived earlier. The payoff and information structure is such that normal individuals adopt if $\theta = +1$ is more likely than $\theta = -1$. In our setup, this occurs when individuals can infer that more H-signals have been observed than L-signals.

Let S_n be the number of H-signals less the number of L-signals that can be inferred by all individuals from the actions of the first n arrivals. Thus, $S_n = S_{n-1} + 1$ if all individuals can infer that the nth individual's signal was H, $S_n = S_{n-1} - 1$ if all individuals can infer that the nth individual's signal was L, and $S_n = S_{n-1}$ if an individual cannot infer anything about the nth individual's signal. It is straightforward to show that, within our information structure, the state of information at any stage n is completely summarized by S_n.

The normal individuals' optimal decision rule is as follows: The nth individual adopts if (i) $S_{n-1} \geq 0$ and she observes H, or (ii) $S_{n-1} \geq 2$

1. Ross and Sicoly (1979) outline why information availability and attribution can have an egocentric bias, and much of their argument would naturally apply to an egocentric bias on judgment about the relative precision of own vs. others' information. In updating a mean estimate, the lower relative subjective variance about the internalized estimate than that of the external estimate would then lead subjects to put too much weight on their internal information and too little weight on external information. While such behavior (by security analysts) has been documented in Abarbanell and Bernard (1992) and Batchelor and Dua (1992), experimental studies of overconfidence that we are aware of show only that subjects' internal information generates assessments with too narrow a range. We are unaware of experimental studies that test the *relative* overconfidence about internalized vs. new external information.

2. When the types of individuals are unknown, each conforming individual could be an entrepreneur or a normal individual. This causes the public state S_n (defined below) to drift (slowly) as more individuals conform. The algebra gets more complex (see Anderson and Holt, 1996), but the intuition of our paper remains: information aggregation is poor, and overconfident entrepreneurs can provide useful information to their group.

and she observes L. Stated differently, the nth individual adopts if $S_n \geq 1$.

To understand the information content of past decisions, consider the cascade scenario in Bikhchandani et al. (1992), in which there are no entrepreneurs ($\lambda = 0$). If the first arrival observes the signal H, then the conditional expected value of adopting is $p - q > 0$ and therefore the first individual adopts. Even though the first individual's information is private, all individuals can infer from the first individual's action that she observed the signal H and thus $S_1 = 1$. Now suppose the second individual observes the signal L. Conditional on the sequence of signals HL, the expected value of adopting is 0 and the individual is indifferent between adopting and rejecting, and, by assumption, the individual abstains. Consequently, if the first individual adopts and the second abstains, then all individuals know that HL has occurred; thus $S_2 = 0$. However, if the second individual also adopts, all individuals know that HH has occurred; thus $S_2 = 2$, and the third individual adopts regardless of her private information. Because this action is uninformative, the fourth individual also adopts regardless of her private information, and this continues for all future individuals. Similarly, when $S_{n-1} = -2$ all subsequent arrivals reject. $S_n = +2$ and $S_n = -2$ are absorbing states, and all future arrivals will conform, adopting or rejecting, respectively. The group gets entrenched in *good* or *bad* cascades. In a good cascade, everyone gets locked into adopting (rejecting) if $\theta = 1$ ($\theta = -1$). In a bad cascade, everyone gets locked into rejecting (adopting) if $\theta = 1$ ($\theta = -1$). The probability of such a bad cascade can be quite high; for example, if $p = 0.51$, it approaches 48% even in large groups.

2.3 Entrepreneurs' Decision Rule

Entrepreneurs also use both publicly available information and their own private information, but place too much weight on the latter. We assume that entrepreneurs believe their signal has precision $p' > p$. For a given p, this leads them to follow their own signal if $|S_{n-1}| < k$ and to behave like normal individuals and follow the crowd if $|S_{n-1}| \geq k$, where the critical state k increases monotonically with p'. [The condition that entrepreneurs always follow their own information if and only if $|S_{n-1}| < k$ is equivalent to $p^{k-1}/(p^{k-1}+q^{k-1}) < p' < p^k/(p^k+q^k)$]. Once the public information becomes sufficiently overwhelming, even if every individual is overconfident, irrationally overconfident actions cease in our model and entrepreneurs suppress their eagerness to ignore the public information in favor of their more limited private information. If $p' = 1$, the critical state k is infinity, and public information never tempers overconfident actions. If $p' = p$, then $k = 2$ and entrepreneurs are like normal individuals, regardless of the current state.

The decision rule for entrepreneurs is similar to those of normal individuals. The entrepreneur ignores public information and follows the private information (adopt if H, reject if L) if the state of information prior to their arrival satisfies $|S_{n-1}| < k$. However, if $|S_{n-1}| \geq k$, entrepreneurs follow their predecessor. The information states k and $-k$ are absorbing states because, once they are reached, no one can infer the information signals of subsequent individuals. In effect, having entrepreneurs expands the "noncascade action interval" between the absorbing cascades states from -2 and $+2$ to $-k$ and $+k$. Once $+k$ or $-k$ is reached, the informational cascade becomes unbreakable even in the presence of entrepreneurs.

2.4 Payoffs

We now define the payoffs and *ex ante* welfare for both types of individual and for the group overall. We begin with the normal types. Let $\widetilde{V}_{\mathrm{R},n}(\lambda)$ denote the random payoff to the nth arrival if she is a normal type, and let $E[\widetilde{V}_{\mathrm{R},n}(\lambda)]$ denote its unconditional expectation. Because the model is symmetric, we only have to consider the cases when she adopts. First, suppose the true value state is $\theta = 1$, in which case adopters get a payoff of 1. Rational types adopt if $S_n \geq 1$, which occurs with probability $\Pr(S_n \geq 1|\theta = 1)$. Yet, if the true value state is $\theta = -1$, adopters receive a payoff of -1, which occurs with probability $\Pr(S_n \geq 1|\theta = -1)$. Thus,

$$
\begin{aligned}
E[\widetilde{V}_{\mathrm{R},n}(\lambda)] &= \Pr[\text{Adopt } |\theta = 1]\Pr[\theta = 1] \\
&\qquad - \Pr[\text{Adopt } |\theta = -1]\Pr[\theta = -1] \\
&= \frac{1}{2}[\Pr(S_n \geq 1|\theta = 1) - \Pr(S_n \geq 1|\theta = -1)] \\
&= \frac{1}{2}[\Pr(S_n \geq 1|\theta = 1) - \Pr(S_n \leq -1|\theta = 1)], \qquad (1)
\end{aligned}
$$

because $\Pr(S_n \geq 1|\theta = -1) = \Pr(S_n \leq -1|\theta = 1)$.

Let $\widetilde{V}_{\mathrm{OC},n}(\lambda)$ denote the random payoff to the nth arrival if she is an entrepreneur. An entrepreneur adopts either if she receives a private high signal and the state is not above k or below $-k$, or if the critical state $+k$ that produces an adopt cascade has already been reached:

$$
\begin{aligned}
\Pr(\text{Adopt } \mid \theta = 1) &= p\Pr(|S_{n-1}| < k \mid \theta = 1) \\
&\qquad + \Pr(S_{n-1} = k \mid \theta = 1), \qquad (2)
\end{aligned}
$$

and

$$\Pr(\text{Adopt } \mid \theta = -1) = q \Pr(|S_{n-1}| < k \mid \theta = -1) + \Pr(S_{n-1} = k \mid \theta = -1). \tag{3}$$

Thus, an entrepreneur expects to receive

$$\begin{aligned} E[\widetilde{V}_{\mathrm{OC},\,n}(\lambda)] = \frac{1}{2}[&p\Pr(|S_{n-1}| < k \mid \theta = 1) \\ &- q\Pr(|S_{n-1}| < k \mid \theta = -1) \\ &+ \Pr(S_{n-1} = k \mid \theta = 1) \\ &- \Pr(S_{n-1} = k \mid \theta = -1)]. \end{aligned} \tag{4}$$

No closed-form solution exists for these probabilities for $p \in (\frac{1}{2}, 1)$, but we can derive a recursion formula to compute them numerically (in the Appendix).

The overall group payoff is the expected payoff to individuals in the group,

$$\begin{aligned} E[\widetilde{V}(\lambda)] &= \lambda \cdot \frac{1}{N}\sum_{n=1}^{N} E[\widetilde{V}_{\mathrm{OC},\,n}(\lambda)] + (1-\lambda) \cdot \frac{1}{N}\sum_{n=1}^{N} E[\widetilde{V}_{\mathrm{R},\,n}(\lambda)] \\ &= \lambda E[\widetilde{V}_{\mathrm{OC}}(\lambda)] + (1-\lambda)E[\widetilde{V}_{\mathrm{R}}(\lambda)]. \end{aligned} \tag{5}$$

From a group perspective, the presence of entrepreneurs in the group helps in that it releases more information, which makes it more likely that most individuals choose correctly; it hurts in that entrepreneurs make more frequent mistakes, which hurts themselves and thus lowers the average group payoff.

2.5 The Solution: The Socially Optimal Proportion of Entrepreneurs (λ^*)

We begin by determining the optimal proportion of entrepreneurs from a social welfare perspective, $\lambda^\star$. The social welfare function is defined as $E[\widetilde{V}(\lambda)]$ in (5), which is maximized by some $\lambda^\star$. The following proposition describes the effect of entrepreneurs—a positive externality—on normal individuals:

Proposition 1:

1. *The* (ex ante) *probability that a normal individual makes an incorrect decision decreases with increasing proportion of entrepreneurs in the group* (λ)*, with increasing degree of overconfidence among entrepreneurs* ($k(p'; p)$)*, and with increasing size of the group* (N)*.*
2. *The limiting probability* ($n \to \infty$) *of being in an incorrect cascade equals* $q^2/(p^2+q^2)$ *for* $\lambda = 0$ *and* $q^k/(p^k+q^k)$ *for all* $\lambda > 0$*.*

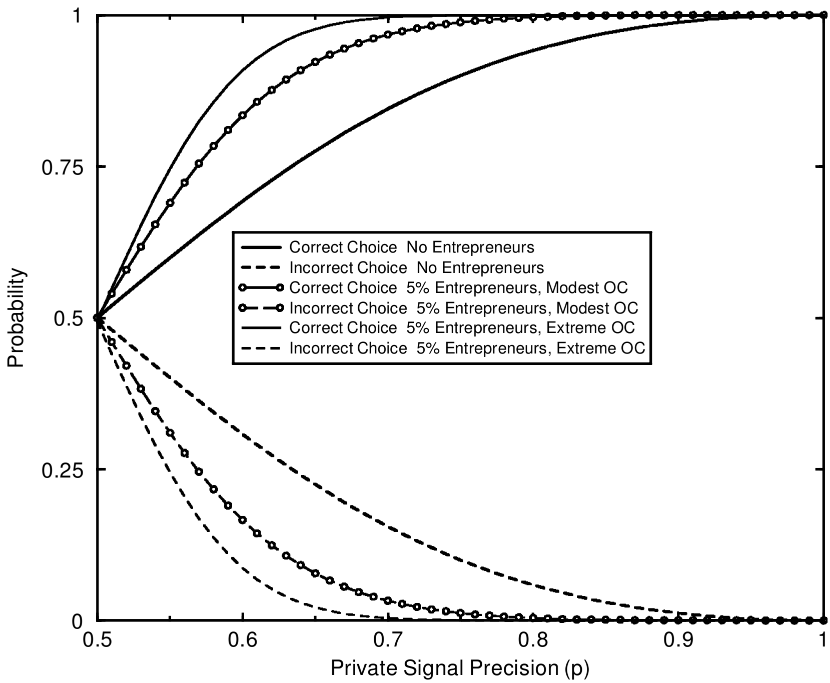

FIGURE 1. THE PROBABILITIES THAT NORMAL INDIVIDUALS LATE IN THE QUEUE END UP IN A CORRECT AND IN AN INCORRECT INFORMATIONAL CASCADE, AS A FUNCTION OF THE PRIVATE SIGNAL PRECISION (p) IN A GROUP OF N = 250 *INDIVIDUALS*

Proof. See Appendix A.1. □

Figure 1 shows the effect of the proportion of entrepreneurs (λ) on the decisions of the normal individuals. Having entrepreneurs shifts probability mass from being in an incorrect cascade to being in a correct cascade.

The probability of not being in a cascade approaches 0 very rapidly. For example, the 10th individual is in a cascade with probability 97.5% if $p = 0.6$, 99% if $p = 0.75$, and 99.98% if $p = 0.9$. The 50th individual is in a cascade with probability in excess of 99.99999% for these three p's. The innermost two lines are the probabilities of ending up in a right or a wrong cascade in the straight informational-cascade-without-entrepreneurs scenario. The outer lines are the same probabilities in the presence of 5% entrepreneurs with modest overconfidence [$k(p'; p) = 12$] and extreme overconfidence [$k(p'; p) = \infty$].

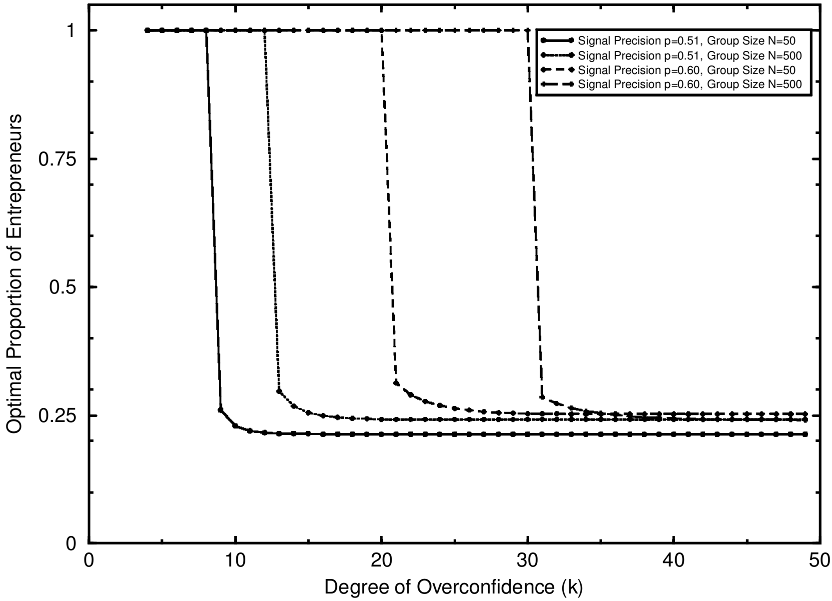

FIGURE 2. THE OPTIMAL PROPORTION OF ENTREPRENEURS (λ^*) AS A FUNCTION OF THE DEGREE OF OVERCONFIDENCE (k) FOR TWO VALUES OF INDIVIDUALS' SIGNAL PRECISION p AND GROUP SIZE N

2.6 Comparative Statics

There are three parameters in our model that influence the optimal proportion of entrepreneurs ($\lambda^\star$): the group size N, the information precision p, and the degree of overconfidence $[k(p'; p)]$. Unfortunately, the effects of the three parameters on $\lambda^\star$ cannot generally be found analytically. However, numerical simulations show that certain directional influences of the parameters p, p', and N on the solution $\lambda^\star$ are present—and to the extent that we cover the relevant parameter space, we can conjecture that they are pervasive.

2.6.1 Degree of Overconfidence (p'). Figure 2 shows that the optimal proportion of entrepreneurs in the group decreases as the degree of overconfidence increases. When overconfidence is extreme, entrepreneurs make more mistakes but provide more information than moderately overconfident entrepreneurs. However, the marginal value of the extra information is small, because it arrives when the public state of information is already very informative. Thus, it is beneficial to the group to have fewer entrepreneurs as overconfidence becomes more extreme. The figure also shows that if overconfidence is modest, it can be socially optimal for the group to consist entirely of entrepreneurs. When overconfidence is modest, entrepreneurs provide extra information when it is most valuable and make few extra mistakes relative to normal types.

We also compared social welfare for all degrees of overconfidence (p'). The social welfare function is typically highest for intermediate levels of overconfidence. For fixed p, greater overconfidence p' is better for larger N. For fixed N, less overconfidence p' is better for larger p. Interestingly, Hirshleifer and Luo (2001) and Wang (2001) demonstrate a similar result in different settings. (Obviously, holding everything else constant, increasing N and p improves the social welfare function.)

2.6.2 Signal Precision. Figure 3 shows that the optimal proportion of entrepreneurs decreases in the private signal precision (p)—except in rare border cases (not plotted). Holding group size constant,

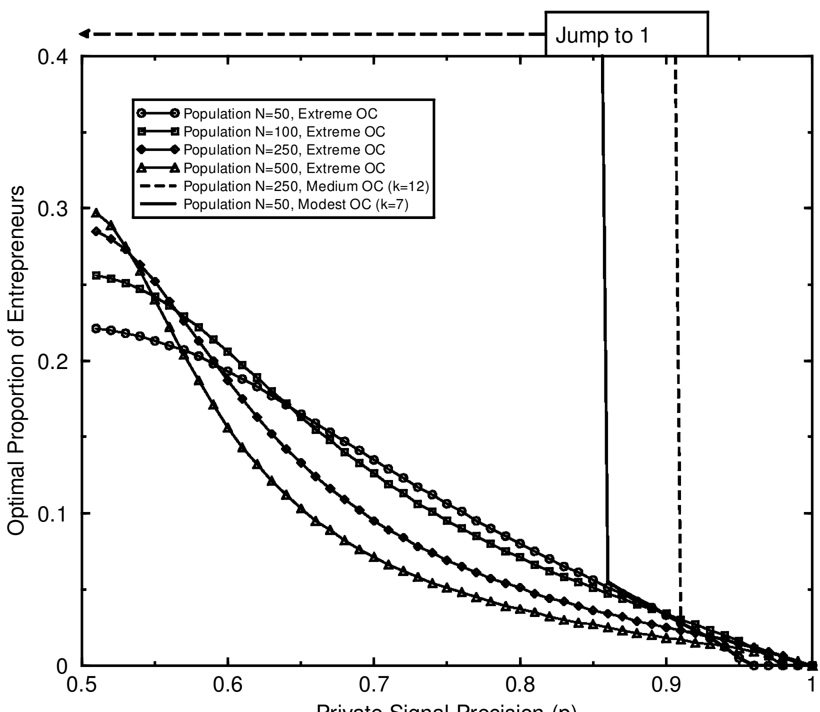

FIGURE 3. THE OPTIMAL PROPORTION OF ENTREPRENEURS (λ^*) AS A FUNCTION OF EACH INDIVIDUAL'S PRIVATE SIGNAL PRECISION (p)

the optimal proportion of entrepreneurs decreases with increasing private signal precision; over a certain range, the optimal proportion of entrepreneurs can be either one or zero. Typically, when p is low, there are more bad cascades in the absence of information aggregation and the cost of being an entrepreneur is lower. This favors the presence of entrepreneurs. The figure also shows that large groups require more entrepreneurs than small groups when the private signal precision is either very low or very high. Thus, the effect of group growth on entrepreneurship depends on signal precision: for very low and very high private information precision, group growth translates into an increased optimal proportion of entrepreneurs. For intermediate private information precision, group growth translates into a smaller, optimal proportion of entrepreneurs.

2.6.3 Group Size (N). Figure 4 shows that the effect of the group size (N) on the optimal $\lambda^\star$ is ambiguous and depends on the level of overconfidence. A higher p would shift all the functions towards

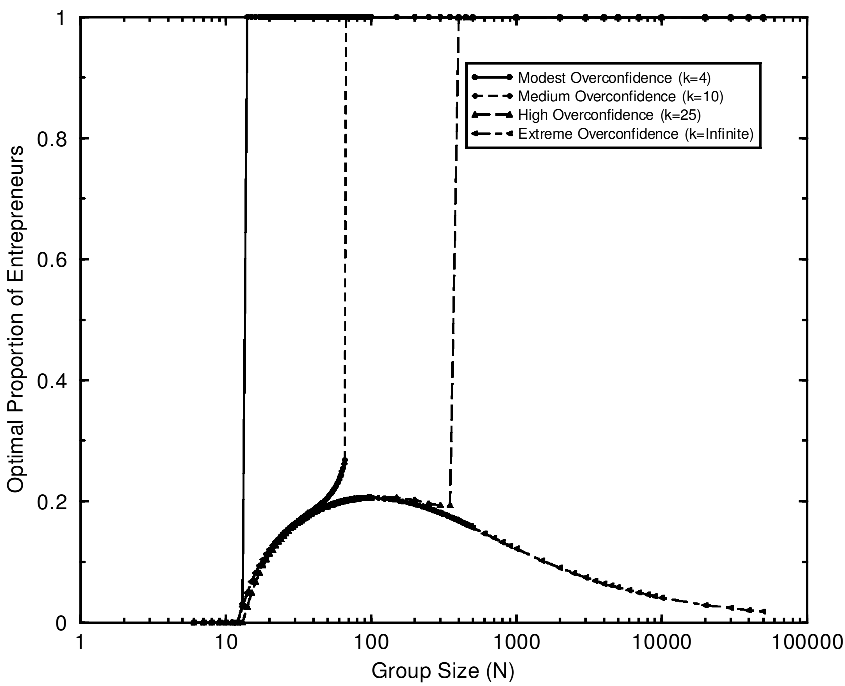

FIGURE 4. OPTIMAL PROPORTION OF ENTREPRENEURS (λ) AS A FUNCTION OF GROUP SIZE N FOR VARIOUS DEGREES OF OVERCONFIDENCE (k) AND HOLDING PRIVATE SIGNAL PRECISION CONSTANT AT $p = 0.6$*

the southeast. When overconfidence is infinite, the optimal proportion of entrepreneurs is zero for small N, then reaches a maximum, and finally asymptotes towards zero. When the overconfidence k is finite, the optimal proportion closely tracks the infinite-overconfidence proportion below a critical population size, and quickly converges to 100% just above that size.

First, consider the case where entrepreneurs exhibit extreme overconfidence. Such entrepreneurs provide a positive externality only to the normal types in the group, because all other entrepreneurs completely ignore the public information. When the group is small, entrepreneurs provide few benefits because there are few individuals following them who can take advantage of the positive information externality. Consequently, a small proportion of entrepreneurs is optimal. Similarly, if the group is very large, even a small proportion of entrepreneurs can represent a large absolute number of entrepreneurs (additional pieces of information), almost assuring a correct cascade. Again, a small proportion of entrepreneurs is optimal. The optimal proportion of entrepreneurs is highest for intermediate group sizes. Now consider the case where entrepreneurs are not perfectly but only moderately overconfident ($k \neq \infty$). The above arguments continue to be true—except that entrepreneurs act like normal individuals once a sufficient number of other entrepreneurs have appeared. At that point, entrepreneurs make no additional mistakes (relative to normal individuals) and thus impose no extra costs on the population. Because the payoffs to both types are increasing in the probability of being in the correct cascade, which increases with the proportion of entrepreneurs, $\lambda^\star = 1$ for sufficiently large N.

It is noteworthy that the positive information externality works well only when groups are sufficiently large. There is a minimum group size necessary for groups to benefit from overconfident behavior. This suggests economies of scale: small groups are intrinsically poorly suited towards taking advantage of the information externality, and may be poorly equipped, e.g., to deal with new situations in which information aggregation is especially important. They cannot afford to risk the loss of entrepreneurs in ordinary situations. The steep slope at the minimum N also suggests that tests of our theory would do well to focus on situations in which group size increased from a very small to a slightly larger number of entrepreneurs (holding relatedness constant). In such cases, we would expect to see an explosion of nonconforming, irrational behavior.

2.7 Omitted Influences

Although the formal model was based on the specific concept of informational cascades, the point of our paper is to argue that overconfidence could have evolved as a device that helps groups to overcome poor information aggregation and to explore their environment better. Offering only a model, this paper has to ignore a number of other influences that can be important. To the extent that other factors can reduce information aggregation or increase the usefulness of the information, the marginal value of having more overconfidence would increase, and we would expect to see more entrepreneurs:

- *Experimentation.* Information aggregation is particularly poor when a situation is unique and choices are discrete, so that individuals cannot repeat and experiment with different choices. As reasoned above, because information aggregation is poorer when experimentation is not feasible, we would expect to see more entrepreneurs and overconfidence.
- *Communication.* Information aggregation is particularly poor when direct talk (conversation) fails, when it is too cumbersome and costly, or when it is not credible. It is better when there is much trust and coordination. Finally, there should be more overconfidence in social species/societies then in solitary species/societies.
- *Ordering.* Information aggregation is particularly poor if the most informed (possibly, most prestigious) individual acts first and thereby induces all subsequent individuals to conform (see Zhang, 1997).
- *Information Costs.* Information aggregation is particularly poor if individuals have to purchase information instead of being freely endowed with it. (To remedy this, we would have to define an entrepreneur as someone who altruistically chooses to purchase information and act on it.)
- *Memory.* Information aggregation is particularly poor if individuals can only observe the most recent individuals, rather than everyone. Then the information in the dissent of a single entrepreneur may be quickly forgotten; it would take a set of consecutive entrepreneurs to signal to the group that a different action would be better.
- *Changing environment.* Information aggregation is more valuable if the environment is stable enough for the actions of previous individuals to be informative.

In our informational cascade model, agents observe the actions of the other agents but not their information. But noncascade settings can offer similar findings, as long as the informational group benefits are large relative to the costs to the entrepreneur. For example, Bolton and Harris (1999) examine a two-armed bandit problem with N agents in which each agent observes the actions *and* information generated by the experimentation of others. They demonstrate that free-rider problems yield suboptimal experimentation (relative to the social optimum). The presence of overconfident individuals would help to mitigate these free-rider problems. Cao and Hirshleifer (2000) extend the simple cascades model to allow early adopters to communicate the payoffs (but not the signals) they received. They show that informational cascades will still occur with positive probability. Moreover, they show that it is possible that observing payoffs and actions of predecessors can reduce average welfare compared to the simple case where only actions can be observed. Again, the presence of overconfident individuals would help to mitigate these free-rider problems.

3. The Public-Goods Problem and Intragroup Survival

3.1 A Stationary Distribution

In the previous model, selection occurred only at the group level, in effect assuming that all individuals within a group were clones. In this section we sketch an environment in which the *public-goods* problem (suboptimal information acquisition from a social perspective because it is in everyone's interest to have someone else acquire the information) exposes individuals to both group and individual selection. This, combined with the fact that the benefits to the group can be several orders of magnitude greater than the costs to the individual, allows entrepreneurs to survive robustly in equilibrium. For example, if the signal precision is $p = 0.51$ and the group contains 500 individuals, the expected group benefit to having a first entrepreneur with $k(p'; p) = 4$ is approximately 114 times larger than the expected cost to this individual.

An entrepreneur can benefit from such large incremental group payoffs/survival to the extent that her genes are "in the same boat" (likely to cosurvive) with those of her other group members in at least three ways:

3.1.1 Indirect Genetic Benefits

(i.e., Kinship/Relatedness). *Hamilton's rule* (e.g., Smith, 1989, pp. 169f; Boyd and Silk, 1997, pp. 260ff) is commonly used in evolutionary genetics to show that a small set of related altruists can initially increase in numbers when they appear within a large population of normal individuals. We now show that a variant of this rule can apply to our entrepreneurs.

Consider a scenario in which the signal precision is $p = 0.51$, and the population consists of a large number of groups, each with $N = 500$ normal individuals and no entrepreneurs. Now suppose that one group appears that contains a (family with a) mutation that gives rise to 20 moderately overconfident individuals with $k(p'; p) = 4$.[3] In a population of many groups composed of normal individuals, the average fitness of normal individuals *in the population* is not much influenced by the presence of 20 entrepreneurs; but the average fitness of the entrepreneurs in the population—100% of whom enjoy the benefits of the presence of the other entrepreneurs—does increase dramatically. It is this disproportionate benefit that allows entrepreneurs to increase in frequency in the overall population of *both* groups, even though their frequency in their own group may decrease.[4]

In this scenario, the expected marginal cost of being the twentieth entrepreneur is 0.0035. The expected marginal benefit to the other 499 individuals from having this twentieth entrepreneur totals 0.11. Yet, only $19/499 \sim 3.8\%$ of the group are other entrepreneurs (the *coefficient of relatedness*).[5] Furthermore, each of the 19 entrepreneurs garners slightly less benefit from the presence of this entrepreneur than a normal individual, because they tend to act more based on their own information than on public information. Adding up the expected benefit to entrepreneurs gives a total gain to the other 19 entrepreneurs of 0.00755 from the presence of the marginal 20th entrepreneur. For the group of twenty individuals, the total cost of overconfidence to all entrepreneurs is 0.129; the total benefit is 0.156. Consequently, in the next generation, entrepreneurial types displace some normal types

3. In standard biology models, simultaneous within-group appearance is assumed in that members of sibgroups interact only with other members of their sibgroup (although these members need not be of the same genotype). Similarly, in Eshel et al. (1998), altruists survive only if they are clustered together with other altruists, but the method to produce spacially correlated distributions of types is different: Altruists can appear by imitating other altruists around them, which allows one altruist to be more likely to benefit other altruists.

4. Both the consideration of small changes in characteristics (such as our introduction of just modest overconfidence) and the consideration of a simultaneous appearance of just a few altruistic individuals within the same group are standard practice in evolutionary biology. If group benefits were not captured disproportionately by other entrepreneurs, altruism (overconfidence) would quickly disappear. Put differently, if the main beneficiaries were egoists, the altruistic type would likely disappear before it could help enough another altruistic types to garner a survival benefit. Such a proximity-of-types condition is necessary in calculations that employ Hamilton's rule to show that biological altruism can increase and is usually accomplished by computing payoffs over paired sibgroups.

5. Relatedness on the order of 5–10% is not implausible. Because marriage occurs primarily in proximity (geographical, cultural), there is more genetic similarity among groups than suggested by gene dispersion by random mating. This is readily visible in some persistent local, regional, national, ethnic, and racial physical traits. For example, the gene for dark skin is ubiquitous in sub-saharan Africa and nonexistent in Europe. For a more random set of polymorphic genes/traits subject more to genetic drift than selection, Lewontin (1974) estimates that 85.4% of the genetic variance among humans is between individuals, 8.3% between populations, and 6.3% is between races. Cavalli-Sforza (1969) finds that individuals in alpine villages in the Parma valley of Italy display a genetic similarity of about 3%.

within the overall population. (It is not important to the argument that normal types from the group containing the entrepreneurs will also displace normal types in other groups.)

3.1.2 Direct Transfers. As in all public-goods problems, it is in the group's interest to find a mechanism to enhance the *internalization* of the positive spillover provided by entrepreneurs. In the national economic sphere, internalizing mechanisms could be patent and copyright protection, or even public recognition and social standing. In a firm or institution, a governing body might be able to directly subsidize or discourage entrepreneurial activity. In small social groups, individuals displaying no intention to explore the environment could be ostracized (Hirshleifer and Rasmusen 1989).

It is reasonable to presume that the within-group sharing arrangement is itself subject to evolutionary pressures, and therefore likely to evolve towards solutions favoring the outcome that enhances group survival. Consequently, we would expect to see group institutions evolve that facilitate long-run solutions closer to the group-optimal $\lambda^{\star}$. Yet, if optimal institutions can evolve, they could potentially reduce the need for biological, irrational overconfidence, and augment entrepreneurship with incentives (transfer subsidies) that are optimal from a group perspective.[6]

3.1.3 Direct Payoff Participation. Groups can share in their success through economies of scale and equitable distribution, e.g., in their joint hunting of large prey or conquest of new territory. While this cannot in itself stop the shrinking *proportion* of entrepreneurs, it can increase the absolute *number* of entrepreneurs. To the extent that groups with entrepreneurs can maintain relative faster growth than groups without them, an equilibrium can emerge. This is explored below.

3.1.4 Summary. In sum, the large discrepancy between group benefits and individual costs makes it easy to construct models in which overconfident entrepreneurs can survive. The next subsection sketches one such possible model.

6. The issue of social/cultural mechanisms as replacements for biological mechanisms is itself rather interesting, and the subject of much debate between sociologists and biologists (e.g., Rogers, 1988). For our paper, we note that the evolution of social institutions is fairly recent in the history of *Homo sapiens*, while imitation has been documented in many vertebrate species (see, e.g., Gibson and Hoglund, 1992). Finally, social mechanisms could also exert pressure towards other, non-overconfidence-based mechanisms that help to resolve the information aggregation problem differently, e.g., with culture and conversation.

3.2 A Displacement Model

3.2.1 Distribution of Types within Groups. We now sketch a model to compute an equilibrium distribution of entrepreneurs. In each generation t, we pit two groups ($g \in [A, B]$) against one another, and we pit individuals within groups against one another. We assume that each group consists of N individuals drawn from an underlying population of groups. Let $f_t(\lambda)$ denote the probability density of λ-groups in the population at generation t, and define $\lambda_{A,t}$ and $\lambda_{B,t}$ to be the realized proportion of entrepreneurs in groups A and B. Groups and individuals compete a large number of times in each generation, so that their payoffs are the expected payoffs computed in Section 2: overconfident individuals in group g receive payoffs $E[V_{\mathrm{OC}}(\lambda_{g,t})]$; normal individuals receive $E[V_{\mathrm{R}}(\lambda_{g,t})]$; and the average group payoff is $\lambda_{g,t}E[V_{\mathrm{OC}}(\lambda_{g,t})] + (1 - \lambda_{g,t})E[V_{\mathrm{R}}(\lambda_{g,t})]$.

3.2.2 The Contest. The winning group displaces the losing group in the next generation. Within the winning group, individuals survive in proportion to their relative payoffs. Consequently, the proportion of entrepreneurs in the next generation, λ_{t+1}, is

$$\lambda_{t+1}(\lambda_{A,t}, \lambda_{B,t}) = \frac{\{\lambda_{g^*,t}E[V_{\mathrm{OC}}(\lambda_{g^*,t})]\}^{w}}{\{\lambda_{g^*,t}E[V_{\mathrm{OC}}(\lambda_{g^*,t})]\}^{w} + \{(1-\lambda_{g^*,t})E[V_{\mathrm{R}}(\lambda_{g^*,t})]\}^{w}}, \tag{6}$$

where g^* denotes the winning group, t is a subscript for the generation, and w modulates the relative efficiency by which overconfident individuals are replaced by normal individuals from one generation to the next. For example, assume groups A and B are of size $N = 100$, the signal precision is $p = 0.6$, and overconfidence is a modest $k(p'; p) = 4$. Also, suppose groups A and B with $\lambda_A = 0.05$ and $\lambda_B = 0.1$, respectively, compete. Entrepreneurs in group A expect to receive a payoff of 0.1644, and normal individuals to receive 0.2417, for a group average of 0.2378. Entrepreneurs in group B expect to receive 0.2126, and normal types to receive 0.2656, for a group average of 0.2603. Consequently, group B survives, but entrepreneurs within this group are in a less favorable position. This latter effect is a standard within-group *replicator dynamic.* The group competition, however, counterbalances the individual selection pressure. With a coefficient w of 1, entrepreneurs constitute $0.1 \times 0.2126/0.2603 \sim 8.2\%$ of the population in the next generation, which is larger than the 7.5% it was in the previous generation.

One can compute a matrix of the resulting proportion of entrepreneurs, λ_{t+1}, as a function of the proportion of entrepreneurs in groups A and B. This matrix has certain general properties. The proportion of entrepreneurs in the next generation declines in cells close to the diagonal of the matrix, because when the two groups have a similar proportion of overconfident types, only individual selection pressure remains. Group selection plays no role when both groups are of equal quality. However, the generational decline in entrepreneurs is zero if there are either no or only entrepreneurs (and it is small nearby). In sum, if the frequency of entrepreneurs across groups has no variance, then group effects cannot persist and only pure groups have a chance of survival.

On moving away from the diagonal cells—i.e., increasing the difference in the proportion of entrepreneurs in the two groups—as one group gets closer to the optimal proportion of entrepreneurs than its competitor, it tends to win. When the group with more entrepreneurs is fitter than its competition, group selection favors more entrepreneurs to counterbalance individual selection in favor of more normal types. Thus, in such matrix cells, the next generation may or may not contain a greater proportion of entrepreneurs than the average proportion from both groups in the previous generation. Finally, in off-diagonal matrix cells in which the group with fewer entrepreneurs is fitter, both individual and group selection effects reduce the proportion of entrepreneurs in the next generation.

3.2.3 Equilibrium Distributions. A probability distribution over λ-groups, $f_t(\lambda)$, is a *stationary distribution* if[7]

$$f_{t+1}(\lambda) = f_t(\lambda). \tag{7}$$

The degenerate distribution $\lambda = 1$ with probability 1 (only entrepreneurs) is a stationary distribution. However, as explained above, no other degenerate distribution is necessarily an equilibrium, because two equal competing groups face no group pressure, and the proportion of types would change in favor of normal individuals in the next generation.

3.2.4 Computing a Stationary Distribution. To illustrate an equilibrium, consider the previously used example with groups of size $N = 100$, entrepreneurs of type $k(p'; p) = 4$, and information precision $p = 0.6$. There are $N + 1 = 101$ possible group arrangements. Define $\pi_{i,t}$ to be the probability that $\lambda = i/N$ for $i \in \{0, 1, 2, \ldots, N\}$. To compute, e.g., the number of groups with a proportion $\lambda_{t+1} = 8/100$ of entrepreneurs in the next generation, we need to consider all possible competitive scenarios in this generation that can produce 8/100

7. This is a different definition from the standard evolutionarily stable strategy (ESS) definition, because we are also concerned with across-group dynamics and not only with within-group dynamics.

entrepreneurs. For example, we showed above that if $\lambda_{A,t} = 0.05$ and $\lambda_{B,t} = 0.1$, then the next generation will have a proportion $\lambda_{t+1} = 0.082$ of entrepreneurs. To maintain discreteness, we assume that the next generation contains a proportion 8/100 of entrepreneurs with probability 80% and 9/100 with probability 20%. The probability that $\lambda_{A,t} = 0.05$ and $\lambda_{B,t} = 0.1$ groups meet is $\pi_{5,t}\pi_{10,t}$; consequently, this scenario contributes probability mass $\pi_{5,t} \times \pi_{10,t} \times 0.8$ to the probability of having $\pi_{8,t+1}$ in the next generation, and $\pi_{5,t} \times \pi_{10,t} \times 0.2$ to the probability of having $\pi_{9,t+1}$ in the next generation. In general, to find $\pi_{x,t+1}$, one must sum probabilities over all possible competitive scenarios. Thus,

$$\pi_{x,t+1} = \sum_{i=0}^{100}\sum_{j=0}^{100} \pi_{i,t}\pi_{j,t}h(x;i,j), \tag{8}$$

where $h(\cdot)$ apportions probability mass to adjacent λ-fractions to maintain the discreteness of the distribution.

More specifically, define $y_{i,j} = \text{int}[\lambda_{t+1}(\lambda_{i,t}, \lambda_{j,t}) \cdot N]$, the integer portion of the number of entrepreneurs, and $m_{i,j} = \lambda_{t+1}(\lambda_{i,t}, \lambda_{j,t}) - y_{i,j}$, the remainder. Then

$$h(x;i,j) = \begin{cases} 1 - m_{i,j} & \text{if } x = y_{i,j}, \\ m_{i,j} & \text{if } x = y_{i,j} + 1, \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

The stationary distribution requires that $\pi_{x,t+1} = \pi_{x,t}$ for all x. The solution to this system of $N + 1$ nonlinear equations and unknowns is *not* necessarily unique. *Evolution need not necessarily lead to a unique outcome.* Figure 5 graphs a set of viable stationary distributions in the example case. In this case, the group-optimal $\lambda^\star$ is 1 (all entrepreneurs)—and the distribution $\lambda = 1$ with probability 1 is also the Pareto-optimal stationary distribution.

This is a general result:

Proposition 2: *If the group-optimal proportion of entrepeneurs, $\lambda^\star$, is 1, a (Pareto-dominating) equilibrium exists in which no groups with any normal types can survive.*

Groups with only entrepreneurs face no individual selection pressure and eventually wipe out all other groups; consequently, normal individuals have no chance to replicate. (This is not an artifact of our severe penalty for the losing group: even if the losing group shrank only slowly, the $\lambda = 1$ group, which does not experience internal

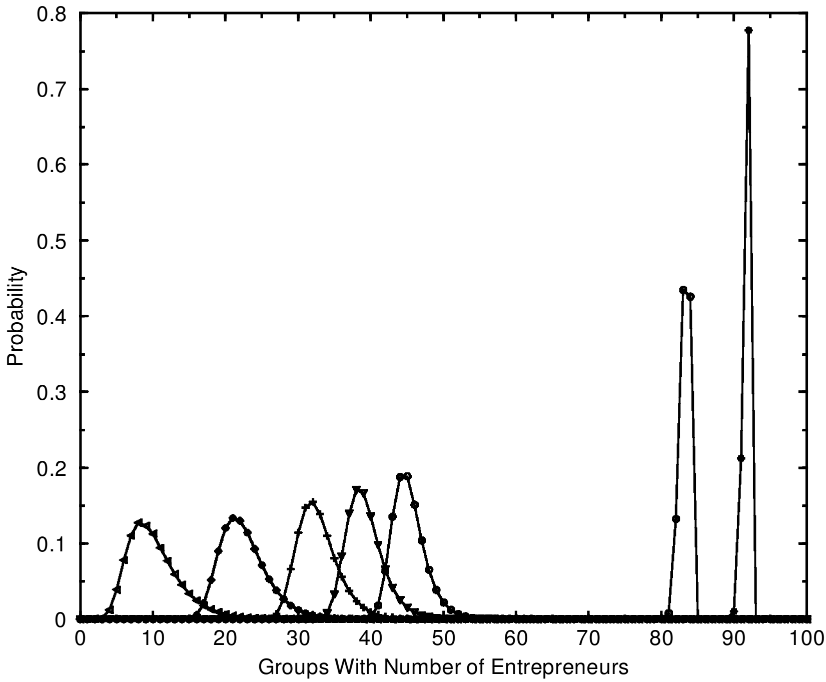

FIGURE 5. SOME STATIONARY STRATEGY DISTRIBUTIONS ($\pi\forall$) WHEN THE POPULATION SIZE N IS 100, SIGNAL PRECISION p IS 0.51, AND OVERCONFIDENCE $k(p';p)$ IS A MODEST 4

selection pressure against entrepreneurs, would still end up eventually displacing all other groups.)

However, when $\lambda^\star \neq 1$, selection pressure against entrepreneurs is strong. For example, Figure 6 graphs a set of equilibria when $p = 0.51$, $k = 12$, and $N = 100$. (These parameters imply that entrepreneurs tend to act with close to extreme overconfidence.) The socially optimal proportion of entrepreneurs is 42.5%. Yet, the Pareto-best distribution of groups contains on average only about 1 entrepreneur per group. The figure shows that entrepreneurs survive in the Pareto-preferred equilibrium, but the average frequency of entrepreneurs in this stationary distribution is "only" about 1–2%—an order of magnitude lower than the group optimum of $\lambda^\star = 0.425$.

These examples show that overconfidence and entrepreneurship can survive in an evolutionary setting. There is a large parameter space in which either all individuals can end up being entrepreneurs or only a certain proportion within the population can end up being entrepreneurs. Unfortunately, a thorough comparative statics analysis is not possible because a unique stationary distribution does not exist.

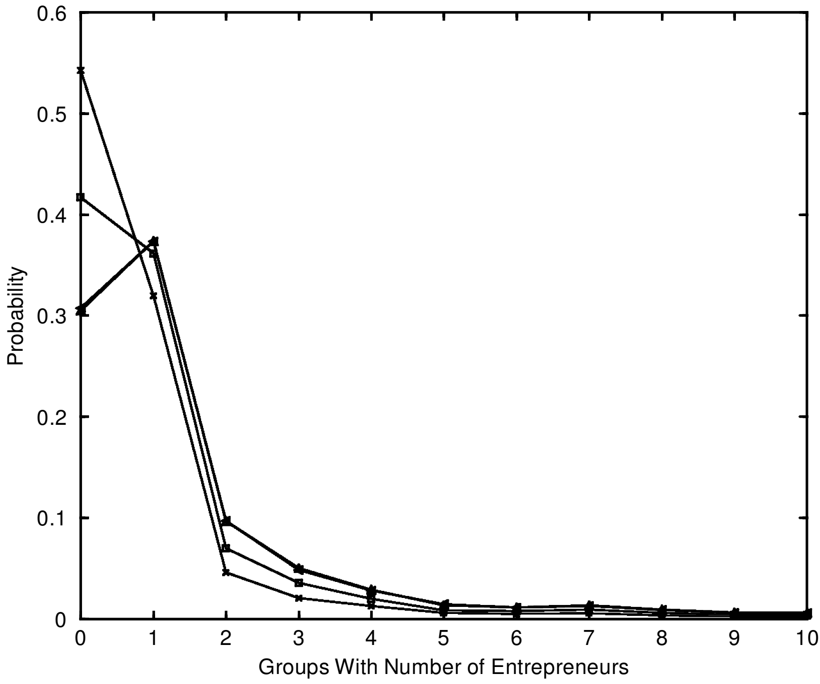

FIGURE 6. SOME STATIONARY STRATEGY DISTRIBUTIONS WHEN THE POPULATION SIZE N IS 100, THE SIGNAL PRECISION p IS 0.51, AND THE OVERCONFIDENCE k IS A LARGE 12

3.3 Group Selection in the Social Sciences

It is generally recognized that although genes are the unit of biological selection, they require *vehicles of selection*. These vehicles are the degree to which genes find themselves "in the same boat" with other genes as far as survival is concerned. The vehicles can be cells (e.g. cancer cells), individuals, kin, villages, tribes, nationalities, ethnicities, races, or even species. The appropriate question is not whether group vehicles are logically possible, but whether selection at a higher organizational level can be sufficiently important to overwhelm selection at a lower organizational level. The empirical evidence indicates that group selection can be an important force, especially in human social structures. Yet, Wilson and Sober (1994) lament that "the most recent developments in biology have not yet reached the human behavioral sciences, which still know group selection primarily as the bogey man of the 60's and 70's." They also argue that "social structures . . . have the effect of reducing fitness differences within groups, concentrating natural selection (and functional organization) at the group level."

4. Alternative Theories Explaining Overconfidence and Entrepreneurs

Our theory has argued that overconfident entrepreneurs are useful, because they broadcast information and thus break the poor information aggregation intrinsic to conformity.

We are aware of only one alternative explanation for seemingly irrational overconfidence, proposed in Trivers (1985) and Hirshleifer (1997): when trying to deceive others that they are of higher ability, individuals' credibility is enhanced if they are themselves convinced of this higher ability. One concern with this argument is that the benefit to overconfidence rests on one's own inability or on other individuals' willingness to let themselves be deceived. Yet, if discovery costs are not too high, those nonentrepreneurial individuals who see through the deception will be more likely to survive than individuals who buy into the "overconfidence equals ability" argument. Fortunately, our own argument for the presence of overconfidence is synergistic: groups that allow themselves to be "deceived"—permitting entrepreneurs to procreate as frequently—receive extra information, which in turn enhances the group's chances for survival. The willingness to be deceived may in effect be a transfer of resources from normal types to entrepreneurs.

The alternative prevailing view of entrepreneurship is that entrepreneurs are tempted by high payoffs associated with nonconforming. This view considers the innovation to be an activity in the entrepreneur's self-interest. In such a setting, if there are diminishing returns to entrepreneurial activity, an optimal, interior proportion of entrepreneurs may arise.[8] This hypothesis is testably different from our hypothesis, in which entrepreneurs are overconfident, make mistakes on average, and suffer in terms of expected payoffs. More realistically, entrepreneurship is likely to be the outcome of both factors: (1) lower risk aversion and the lure of payoffs substantially higher than those available to the majority, and a higher tendency for some such risk-loving individuals to survive within their group, and (2) a genetic overconfidence of entrepreneurs, and a higher tendency for groups with some such individuals to survive.

In addition to the informational externality and the abovementioned risky benefits, there are certainly other real-world facets of

8. Models in which individuals can either follow or learn (but not both) and in which the marginal costs/benefits to the two activities are equal in equilibrium can be found in Boyd and Richerson (1988) and Rogers (1988). This typically results in learning that is suboptimal from a group perspective. Biological models in which some, but not all, individuals pursue an activity in their own self-interest, and in which an optimum interior proportion develops, can be found in Smith (1974), Cornell and Roll (1981), and Weibull (1995).

entrepreneurship that we have omitted. Still, the informational benefits of entrepreneurship discussed in our paper are likely to persist in a much richer model than is considered in our paper.

5. Conclusion

Our paper argues that overconfidence (and with it certain forms of entrepreneurship) can persist because overconfident behavior broadcasts valuable private information to the group—information that would be lost if rational individuals instead just "followed the herd." We explored the costs and benefits to individuals and groups in a simple setting. A group with too few entrepreneurs falls too easily into an incorrect choice (in which the entire crowd follows the wrong path), because of poor aggregation of information across individuals. A group with too many entrepreneurs has too many individuals relying only on their own information and making mistakes too often, and thus suffers from high attrition. The social optimum trades off the information externality against this attrition. We also identified a set of influences (e.g., group size, information precision, degree of overconfidence, type of decision, etc.) that influences the optimal proportion of entrepreneurs (degree of overconfidence) in groups.

Unfortunately, individual selection tends to discriminate against entrepreneurs. From an economic perspective, this is a particularly severe form of a public-goods problem. We showed that the benefits to the group can easily be two orders of magnitude larger than the costs to the entrepreneur. We briefly discussed multiple mechanisms by which entrepreneurs may participate in such large payoffs, and offered one such model that trades off the tendency of overconfident individuals to disappear against the tendency of groups without overconfident individuals to disappear. We showed that there are situations in which groups exclusively consisting of entrepreneurs can drive out mixed groups, and other situations in which entrepreneurs can survive in mixed groups.

We have mixed feelings about group selection, which is rarely found in economic theory today. Clearly, the first best set of explanations for behavior patterns should be based on individual rationality. But when there is no rational explanation for a widely observed deviation from rationality, the next best set of explanations should be based on arguments in which this behavior is helpful to the individual's group (and preferably of limited harm to the individual altruist himself). Having invoked such a group selection argument allowed us to explain and justify the presence of an otherwise seemingly irrational, unjustifiable and ubiquitous behavior pattern—overconfidence—while still imposing a discipline on the types of "reasonable" behavioral anomalies we would permit in a model.

In conclusion, there are many facets to the presence and benefits of entrepreneurs, the entrepreneurial spirit, and overconfidence, not all of which are captured by our model. But we believe that our theory has captured one important aspect of overconfidence and entrepreneurial culture—a possibly rare but persistent presence of individuals who provide information to their group—in a simple, reasonable, and intuitive model.

Appendix. Recursion Formula and Proofs

A.1 Recursion Formula for State-Time Probabilities

Let π_s^n denote the probability of being in state s after n signals that can be inferred if the true state is $\theta = 1$. These probabilities are required to compute $E[\widetilde{V}_{\mathrm{R},\,n}(\lambda)]$ and $E[\widetilde{V}_{\mathrm{OC},\,n}(\lambda)]$ in equations (1) and (4), respectively. The recursion formulae for deriving these probabilities are as follows:

$$
\begin{aligned}
\pi_0^n &= p\pi_{-1}^{n-1} + q\pi_1^{n-1},\\
\pi_{-1}^n &= \lambda p\pi_{-2}^{n-1} + q\pi_0^{n-1},\\
\pi_1^n &= \lambda q\pi_2^{n-1} + p\pi_0^{n-1},\\
\pi_{-2}^n &= \lambda p\pi_{-3}^{n-1} + q\pi_{-1}^{n-1} + (1-\lambda)\pi_{-2}^{n-1},\\
\pi_2^n &= \lambda q\pi_3^{n-1} + p\pi_1^{n-1} + (1-\lambda)\pi_2^{n-1},\\
\pi_s^n &= \lambda q\pi_{s+1}^{n-1} + \lambda p\pi_{s-1}^{n-1} + (1-\lambda)\pi_s^{n-1} \qquad \text{for} \quad 2 < |s| < k-1, \qquad \text{(A1)}\\
\pi_{-k+1}^n &= \lambda q\pi_{-k+2}^{n-1} + (1-\lambda)\pi_{-k+1}^{n-1},\\
\pi_{k-1}^n &= \lambda p\pi_{k-2}^{n-1} + (1-\lambda)\pi_{k-1}^{n-1},\\
\pi_{-k}^n &= \lambda q\pi_{-k+1}^{n-1} + \pi_{-k}^{n-1},\\
\pi_k^n &= \lambda p\pi_{k-1}^{n-1} + \pi_k^{n-1}
\end{aligned}
$$

with the starting value given by $\pi_0^0 = 1$ if $k \geq 4$.

A.2 Proof of Proposition 1

Proof of Part 1. Fix n, and let $X_i = 1$ ($X_i = -1$) if the ith signal is high (low) and it can be publicly inferred. Fix $k(p'; p) = k$, let $Y_1(n)$ denote the random number of publicly inferred signals by the nth arrival if $\lambda = \lambda_1$, and let $Y_2(n)$ denote the random number of publicly inferred signals by the nth arrival if $\lambda = \lambda_2$. Without loss of generality, let $\lambda_1 > \lambda_2$.

The state of publicly observed information at the nth arrival is $S_n^1 = \sum_{i=1}^{Y_1(n)} X_i$ if $\lambda = \lambda_1$ and $S_n^2 = \sum_{i=1}^{Y_2(n)} X_i$ if $\lambda = \lambda_2$.

The *ex ante* probability that a rational individual makes an incorrect decision if she is the nth arrival is equal to the probability that she accepts the project when the project is an incorrect one ($\theta = -1$) plus the probability that she does not accept the project when the project is a good one ($\theta = 1$). Thus, the probability of making an incorrect decision when $\lambda = \lambda_j$ is

$$\begin{aligned}
&\Pr[S_n^j \geq 1|\theta = -1]\Pr[\theta = -1] + \Pr[S_n^j \leq 0|\theta = 1] \cdot \Pr[\theta = 1] \\
&\quad = \tfrac{1}{2}\Pr[S_n^j \geq 1|\theta = -1] + \tfrac{1}{2}\Pr[S_n^j \leq 0|\theta = 1] \\
&\quad = \tfrac{1}{2}\Pr[S_n^j \leq -1|\theta = 1] + \tfrac{1}{2}\Pr[S_n^j \leq 0|\theta = 1].
\end{aligned} \tag{A2}$$

If $\lambda_1 > \lambda_2$ then $Y_1(n)$ is stochastically larger than $Y_2(n)$; i.e., $\Pr[Y_1(n) > a] \geq \Pr[Y_2(n) > a]\,\forall a$ and $\forall n$. Consequently,

$$\begin{aligned}
&\tfrac{1}{2}\Pr[S_n^1 \leq -1|\theta = 1] + \tfrac{1}{2}\Pr[S_n^1 \leq 0|\theta = 1] \\
&\quad = \tfrac{1}{2}E[1_{(S_n^1 \leq -1|\theta=1)} + 1_{(S_n^1 \leq 0|\theta=1)}] \\
&\quad = \tfrac{1}{2}E[E[1_{(S_n^1 \leq -1|\theta=1)} + 1_{(S_n^1 \leq 0|\theta=1)}|Y_1(n)]] \\
&\quad = \tfrac{1}{2}E[f(Y_1(n))] \\
&\quad \leq \tfrac{1}{2}E[f(Y_2(n))] \\
&\quad = \tfrac{1}{2}\Pr[S_n^2 \leq -1|\theta = 1] + \tfrac{1}{2}\Pr[S_n^2 \leq 0|\theta = 1],
\end{aligned} \tag{A3}$$

where the inequality follows from the fact that $f(\cdot)$ is a decreasing function because the X_i are i.i.d. with $E[X_i] > 0$ and $Y_1(n)$ is stochastically larger than $Y_2(n)$. See Ross (1983, Proposition 8.1.2).

For fixed λ, simply let $Y_1(n)$ denote the random number of publicly inferred signals by the nth arrival if $k = k_1$, let $Y_2(n)$ denote the random number of publicly inferred signals by the nth arrival if $k = k_2$, and w.l.o.g. let $k_1 > k_2$. Now the remainder of the proof is identical to the proof above.

Finally, since X_i are i.i.d. with $E[X_i] > 0$, the probability of making an incorrect decision decreases with each successive arrival when λ and k are fixed. Thus, the greater is N, the smaller is the *ex ante* probability that a rational type will make an incorrect decision. □

Proof of Part 2. Follows directly from the solution to the gambler's ruin problem (see Ross, 1983, p. 115f). □

References

Abarbanell, J.S. and V.L. Bernard, 1992, "Tests of Analysts' Overreaction/Underreaction to Earnings Information as an Explanation for Anomalous Stock Price Behavior," *The Journal of Finance*, 47(3), 1181–1208.

Ainslie, G., 1975, "Specious Reward: A Behavioral Theory of Impulsiveness," *Psychological Bulletin*, 82(4), 463–496.

Anderson, L.R. and C.A. Holt, 1996, "Information Cascades in the Laboratory," *American Economic Review*, 87(5), 847–862.

Banerjee, A.V., 1992, "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, 107(3), 797–818.

Batchelor, R. and P. Dua, 1992, "Conservatism and Consensus-Seeking among Economic Forecasters," *Journal of Forecasting*, 11(2), 169–181.

Becker, G.S., 1976, "Altruism, Egoism, and Genetic Fitness: Economics and Sociobiology," *Journal of Economic Literature*, 14(3), 817–826.

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

Bolton, P. and C. Harris, 1999, "Strategic Experimentation," *Econometrica*, 67(2), 349–374.

Boyd, R. and P.J. Richerson, 1988, "An Evolutionary Model of Social Learnings: The Effects of Spatial and Temporal Variation," in T.R. Zentall, J. Bennett, and G. Galef, eds., *Social Learning*, Hillsdale, NJ: Lawrence Erlbaum Associates.

—— and J.B. Silk, 1997, *How Humans Evolved*, New York: W.W. Norton.

Busenitz, L.W. and J.B. Barney, 1997, "Differences between Entrepreneurs and Managers in Large Organizations: Biases and Heuristics in Strategic Decision-Making," *Journal of Business Venturing*, 12(1), 9–30.

Cao, H. and D. Hirshleifer, 2000, "Conversation, Observational Learning, and Informational Cascades," Working Paper, Ohio State University.

Cavalli-Sforza, L.L., 1969, "Genetic Drift in an Italian Population," *Scientific American*, 221, 30.

Cooper, A.C., C.A. Woo, and W. Dunkelberg, 1988, "Entrepreneurs Perceived Chances for Success," *Journal of Business Venturing*, 3, 97–108.

Cornell, B. and R. Roll, 1981, "Strategies for Pairwise Competitions in Markets and Organizations," *Bell Journal of Economics*, 12(1), 201–216.

Daniel, K., D. Hirshleifer, and A. Subrahmanyam, 1998, "A Theory of Overconfidence, Self-Attribution, and Security Market Under- and Over-reaction," *The Journal of Finance*, 53(5), 1839–1886.

DeBondt, W.F.M. and R.H. Thaler, 1995, "Financial Decision-Making in Markets and Firms: A Behavioral Perspective," in R. Jarrow, V. Maksimovic, and W. Ziemba, eds., *Finance*, Elsevier North Holland, Amsterdam, Handbook in Operations Research and Management Science, Vol. 9, Chap. 13, 385–410.

Delong, B.J., A. Shleifer, L. Summers, and R.J. Waldmann, 1991, "The Survival of Noise Traders in Financial Markets," *Journal of Business*, 64(1), 1–20.

Eshel, I., L. Samuelson, and A. Shaked, 1998, "Altruists, Egoists, and Hooligans in a Local Interaction Model," *American Economic Review*, 88(1), 157–179.

Gibson, R.M. and J. Hoglund, 1992, "Copying and Sexual Selection," *TREE*, 7(7), 229–232.

Hirshleifer, D., 1997, "An Adaptive Theory of Self Deception," Working Paper, University of Michigan, Graduate School of Business.

—— and G. Luo, 2001, "On the Survival of Overconfident Traders in a Competitive Securities Market," *Journal of Financial Markets*, 4, 73–84.

—— and E. Rasmusen, 1989, "Cooperation in a Repeated Prisoners' Dilemma with Ostracism," *Journal of Economic Behavior and Organization*, 12(1), 87–106.

Hirshleifer, J., 1977, "Economics from a Biological Viewpoint," *Journal of Law and Economics*, 20(1), 1–52.

—— 1987, "On the Emotions as Guarantors of Threats and Promises," in J. Dupre, ed., *The Latest on the Best: Essays on Evolution and Optimality*, Cambridge, MA: The MIT Press 307–326.

—— and J.C. Martinez-Coll, 1988, "What Strategies Can Support the Evolutionary Emergence of Cooperation?" *Journal of Conflict Resolution*, 32(2), 367–398.

Lewontin, R.C., 1974, *The Genetic Basis of Evolutionary Change*. New York: Columbia University Press.

Manove, M., 1998, "Entrepeneurs, Optimism, and the Competitive Edge," Working Paper, Boston University.

Odean, T., 1998, "Volume, Volatility, Price and Profit When All Traders Are Above Average," *The Journal of Finance*, 53(6), 1887–1934.

Rogers, A.R., 1988, "Does Biology Constrain Culture?" *American Anthropologist*, 819–831.

——, 1994, "Evolution of Time Preference by Natural Selection," *American Economic Review*, 84(3), 461–481.

Ross, M. and F. Sicoly, 1979, "Egocentric Biases in Availability and Attribution," *The Journal of Personality and Social Psychology*, 37, 322–336.

Ross, S.M., 1983, *Stochastic Processes*, New York: Wiley.

Smith, J.M., 1974, "The Theory of Games and the Evolution of Animal Conflicts," *Journal of Theoretical Biology*, 209–219.

——, 1989, *Evolutionary Genetics*, Oxford: Oxford University Press.

Trivers, R., 1985, *Social Evolution*, Menlo Park, CA: Benjamin/Cummings.

Waldman, M., 1994, "Systematic Errors and the Theory of Natural Selection," *American Economic Review*, 84(3), 482–497.

Wang, A., 2001, "Overconfidence, Investor Sentiment, and Evolution," Working Paper, Rice University; *Journal of Financial Intermediation*, forthcoming.

Weibull, J.W., 1995, *Evolutionary Game Theory*, Cambridge, MA: The MIT Press.

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

Wilson, D.S. and E. Sober, 1994, "Re-introducing Group Selection to the Human Behavioral Sciences," *Behavioral and Brain Sciences*, http://www.cogsci.soton.ac.uk/bbs/; Scientific American, July 1996.

Zhang, J., 1997, "Strategic Delay and the Onset of Investment Cascades," *Rand Journal of Economics*, 28(1), 188–205.