#9-97

On the Evolution of Overconfidence and Entrepreneurs

May 1997
Revised August 1998

Antonio Bernardo
Anderson Graduate School of Management
University of California, Los Angeles

and

Ivo Welch
Anderson Graduate School of Management
University of California, Los Angeles
On the evolution of overconfidence and entrepreneurs*

Antonio Bernardo†

Ivo Welch‡

Abstract

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 evolutionary stable proportion trades off the survival of the group against the survival of overconfident individuals.

*This paper (UCLA WP-9/97, August 19, 1998) is available from http://linux.anderson.ucla.edu/research.html. Common US mail address: Anderson Graduate School of Management, 110 Westwood Plaza, Box 951481, Los Angeles, CA 90095, Tel: (310) 825-2508. We are grateful for comments from the editor (Lars Stole) and two anonymous referees, Jennifer Aaker, Robert Boyd, Kent Daniel, David Hirshleifer, Jack Hirshleifer, Josh Lerner, Steve Lippman, John Mamer, Terry Odean, Richard Roll, Larry Samuelson, Avanidhar Subrahmanyan, Karen Stephenson, Karen Van Nuys, and workshop participants at the University of Minnesota and Harvard Business School. Bill McKelvey especially had to suffer from giving us detailed comments on multiple drafts.

†Contact: UCLA Anderson Graduate School of Management, mailto:antonio.bernardo@anderson.ucla.edu.

‡Contact: UCLA Anderson Graduate School of Management, mailto:ivo.welch@anderson.ucla.edu.
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 that more rational individuals might not undertake. For example, overconfidence among economic entrepreneurs has been documented by Cooper, Woo, and Dunkelberg (1988). In their sample of 2,994 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 overconfident behavior, which has clearly and reproducibly been documented in laboratory settings to be irrational.\footnote{In our model, overconfidence causes entrepreneurship, not vice-versa! Thus, our paper does not require each or even all successful entrepreneurs to be irrational. Successful entrepreneurs could be better informed and thus rational; and they would still provide the information externality discussed in our paper. But the point of our paper is to explain the presence of documented irrational overconfidence in the unconditional distribution of (entrepreneurially oriented) individuals.} 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 relies on overconfidence as an underlying primitive assumption, often with little theoretical justification as to why such irrational behavior can persist.\footnote{For example, in Delong, Shleifer, Summers, and Waldmann (1991) and Daniel, Hirshleifer, and Subrahmanyam (1998), overconfidence is invoked to explain various stock return anomalies. Odean (1998) uses overconfidence to explain trading volume and reviews both the relevant psychology and financial literature. Unfortunately, the relevant literature is too large to provide a comprehensive list of references.}

Our paper offers a simple explanation for the presence of overconfidence and entrepreneurs. Economists are familiar with a variety of models in which individuals can observe one another and typically end up following the same action, yet information aggregation is poor because rational, non-entrepreneurial individuals who follow "the herd" reveal nothing about their private information. From a social perspective, this leads 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 to their own detriment but which help their groups. Our paper therefore argues that groups with some overconfident individuals have an evolutionary advantage over groups without such individuals. We identify conditions under which the benefits of entrepreneurship to the group are high and the costs to individual entrepreneurs are low, and we derive an evolutionary stable equilibrium in which overconfidence persists across generations.

To make our main point—that overconfident entrepreneurs provide socially valuable information and exploration services—more concrete, our modeling framework is based on the concept of informational cascades, introduced in Welch (1992), Banerjee (1992), and Bikhchandani, Hirshleifer, and Welch (1992). Informational cascade models are relatively simple and intuitive, and require only modest assumptions. It is easy to show that in an informational cascade, individuals are quickly locked into following previous individuals' choices, even though the actions of the crowd are based on very little information. To illustrate, consider a situation in which individuals have to make a discrete decision whether to adopt an action, do so sequentially, and have one private signal which is informative about whether the action is worthwhile or not. For simplicity, assume that an individual adopts if he can infer the presence of more favorable information than unfavorable information. Now consider the situation of the third individual in line if he has seen the previous two individuals adopt. He can (typically) infer that they held favorable information. Even if he holds one piece of unfavorable information, two pieces of favorable information outweigh it. He therefore adopts, subsequent individuals can infer nothing from his action, and—finding themselves in the same situation—will also adopt. Based on the actions of only the first two individuals, everyone thereafter will adopt, even if each and every subsequent individuals were to hold unfavorable information. The crucial conditions necessary for the formation of an informational cascade are that individuals must select one from a small number of discrete choices and that "actions speak louder than words." If these conditions are met, and if by chance a few individuals early on have incorrect information, all

---

3It is important to note that our main argument—that overconfidence and entrepreneurship actions can convey valuable information to their social group—holds also in non-informational cascade settings. The important aspect is that the benefits to the additional information are large, while the costs to entrepreneurs are small. While our argument works well in a cascade setting, it would also hold, e.g., in a two-armed bandit or experimentation setting. For example, irrational overconfidence can induce entrepreneurs to explore new options, and the outcome could be publicly observable. Thus, the presence of overconfident entrepreneurs would aid the within-group information collection and aggregation.
individuals can end up making the wrong choice. A mechanism for overthrowing informational cascades when they are incorrect would thus be socially desirable.

In our framework, overconfident entrepreneurs are one such mechanism. In our paper, an entrepreneur is an innovator, someone who puts less emphasis on what the crowd is doing and instead tends to follow his own information. In a chapter on entrepreneurship and organization, Peterson (1981) elaborates on the view of entrepreneurs in sociology, anthropology, and organizational strategy: "French economists used the term [entrepreneurs] to designate persons in the economic sphere who made innovations..." Other terms outside the commercial sphere that have been used to denote the kind of entrepreneur that we model are independent spirits, innovators, leaders, change agents, and dissidents. The presence of overconfident individuals who act on their own information and ignore the actions of other individuals in the group has already been demonstrated in laboratory settings in Anderson and Holt (1996). They also find that the effect of just one such individual acting in a different way can have profound consequences, overthrowing a (frequently incorrect) cascade. Because everyone realizes that previous opinion had been based on very few independent observations, such deviant behavior can potentially benefit a large number of subsequent individuals. In our earlier example, if the hundredth individual were an entrepreneur who holds unfavorable information and visibly rejects the action, the hundred-and-first individual would have a meaningful decision to make. Having just one entrepreneur means that subsequent crowd behavior would be based on at least four pieces of information (the first two individuals, plus the entrepreneur, plus the now non-trivial decision of the hundred-and-first individual), and not just the two pieces of information (the first two individuals). Unlike the informational cascade situation without entrepreneurs—in which the actions of just a few early individuals could determine the behavior of all subsequent individuals—the presence of some entrepreneurs permits the crowd to act on more than just these few pieces of information. Acting on more information tends to overthrow incorrect informational cascades and tends to deepen correct informational cascades. We show that when the proportion of entrepreneurs in the group increases, the probability of falling into an incorrect cascade declines.

In our context, the overconfident decision not to follow the informational cascade is suboptimal: it is more often wrong than it would be if the entrepreneur had followed the crowd. Because not following the crowd is against the entrepreneur's own self-interest, we would not expect entrepreneurs to reflect too deeply on their actions; instead, we would expect en-
trepreneurs to be "programmed" (either socially or biologically) to overestimate the quality of their own information. Fortunately, the disadvantage of being an entrepreneur is limited, because of the poor information aggregation: the cascade aggregates the information of only a small number of other individuals, which makes following one's own information instead of the cascade only slightly worse. The aforementioned entrepreneur, the hundredth individual in line, would base his decision only on his one piece of information, rather than on the two pieces of information that initiated the cascade. Our model can easily be calibrated to generate benefits to the group that are larger by a factor of 100 than the cost to the individual.

Although the social cost of having entrepreneurs is that their decisions are more often wrong than those of herd followers, a group which never produces entrepreneurs is likely to suffer from relatively frequent aggregate mistakes and is likely to be out-competed by other groups that do produce entrepreneurs. Similarly, a group that produces too many entrepreneurs is likely to lose too many individuals (their entrepreneurs) to attrition due to their poorer judgment (relative to the judgment of individuals who use all available information), and thus may be out-competed by other groups that produce fewer entrepreneurs. This allows us to derive both the socially optimal proportion of entrepreneurs ("experimentation"), which trades off the information externality against entrepreneurial attrition; and the "ESS" (evolutionary stable strategy) distribution of entrepreneurs, which trades off the relative survival of the group against the relative survival of the (altruistic) individuals that are overconfident.

The pervasiveness of overconfidence is interesting in itself, but at least of equal interest is Schumpeter's hypothesis that capitalism is driven by entrepreneurs. In fact, the role of entrepreneurs remains one of the central questions in economics today. In one sense, our paper is a rigorous exploration of one aspect of Schumpeter's hypothesis—the driving forces behind and the externalities produced by entrepreneurs. 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 modelled by our paper. But even if other forces are at work, the information externality considered in our paper is likely to contribute to the ultimate causes behind the development and persistence of entrepreneurship.
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 tradeoff between the informational externality and entrepreneurial attrition. Section 3 derives the "evolutionary stable strategy" (ESS) distribution, i.e., the tradeoff between inter-group and intra-group 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 terms of revealing their private information to their groups. Although our specific model is based on the cascades framework, it could have equally well been based on a different framework (e.g., a two-armed bandit search model, in which information is revealed only if a certain action is taken). The basic intuition and comparative statics would be similar.

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, $\theta$. The action is costless and the true value of $\theta$ is either $-1$ or $+1$, each with prior probability $1/2$. No individual can observe the true value of $\theta$, but each individual can privately observe an i.i.d. signal which is correlated with $\theta$. 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 $H$. Conversely, 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.\(^4\) Table 1 summarizes the information structure.

<table>
<thead>
<tr>
<th>Signal Value</th>
<th>Value State</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H )</td>
<td>( \theta = 1 )  ( p )  ( q )</td>
</tr>
<tr>
<td>( L )</td>
<td>( \theta = -1 ) ( q )  ( p )</td>
</tr>
</tbody>
</table>

Table 1: 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 \( \theta = 1 \) (\( \theta = -1 \)). We assume that the individual abstains \textit{if and only if} she is indifferent between adopting and rejecting. Allowing abstentions avoids tie-breaking rules or randomization when the individual is indifferent between adopting and not adopting. This assumption is made for simplicity and the results are similar when there is no abstention option. 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, Hirshleifer, and Subrahmanyam (1998)) and to the findings in Anderson and Holt (1996): Individuals are either \textit{relatively} more "skeptical" about external information or \textit{relatively} more "overconfident" about internalized information.\(^5\)

\(^4\)If \( p = 1/2 \) the information signals are completely uninformative about the true \( \theta \) while if \( p = 1 \) the information signals are perfectly informative.

\(^5\)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 relative precision of own vs. other's information. In updating a mean estimate, the relatively lower subjective variance about the internalized estimate compared to 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.
For computational simplicity, we assume that the type of each individual is public knowledge.\(^6\) We define \(\lambda\) to be the proportion of entrepreneurs in the group. Note that \(\lambda\) 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 still follows the public information.

**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 \(n^{th}\) individual's signal was \(H\), \(S_n = S_{n-1} - 1\) if all individuals can infer that the \(n^{th}\) individual's signal was \(L\), and \(S_n = S_{n-1}\) if an individual cannot infer anything about the \(n^{th}\) 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 \(n^{th}\) individual adopts if (i) \(S_{n-1} \geq 0\) and she observes \(H\); or (ii) \(S_{n-1} \geq 2\) and she observes \(L\). Stated differently, the \(n^{th}\) individual adopts if \(S_n \geq 1\).

To understand the information content of past decisions, consider the cascade scenario in Bikhchandani, Hirshleifer, and Welch (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

---

\(^6\)When the type of individual is 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 overconfidence /entrepreneurs can provide useful information to their group.
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.

**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 $p'$.\(^7\) Once the public information becomes sufficiently overwhelming, even if every individual were 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

\(^7\)The condition that entrepreneurs always follow their own information if and only if $|S_{n-1}| < k$ is equivalent to $p^{k-1} / (p^{k-1} + q^{k-1}) < p' < p^k / (p^k + q^k)$.
because, once reached, no one can infer the information signals of subsequent individuals. In effect, having entrepreneurs expands the “non-cascade 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.

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 $\tilde{V}_{R,n}(\lambda)$ denote the random payoff to the $n^{th}$ arrival if she is a normal type and let $E[\tilde{V}_{R,n}(\lambda)]$ 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,

$$
E[\tilde{V}_{R,n}(\lambda)] = Pr[\text{Adopt} | \theta = 1] \cdot Pr[\theta = 1] - Pr[\text{Adopt} | \theta = -1] \cdot Pr[\theta = -1] \\
= 1/2 \cdot [Pr(S_n \geq 1 | \theta = 1) - Pr(S_n \geq 1 | \theta = -1)] \\
= 1/2 \cdot [Pr(S_n \geq 1 | \theta = 1) - Pr(S_n \leq -1 | \theta = 1)],
$$

(1)

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

Let $\tilde{V}_{OC,n}(\lambda)$ denote the random payoff to the $n^{th}$ 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 $-k$, and if the critical state $+k$ that produces an adopt-cascade has already been reached:

$$
Pr(\text{Adopt} | \theta = 1) = p \cdot Pr(|S_{n-1}| < k | \theta = 1) + Pr(S_{n-1} = k | \theta = 1),
$$

(2)

and

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

(3)

Thus, an entrepreneur expects to receive

$$
E[\tilde{V}_{OC,n}(\lambda)] = 1/2 \cdot \left[ p \cdot Pr(|S_{n-1}| < k | \theta = 1) - q \cdot Pr(|S_{n-1}| < k | \theta = -1) \\
+ Pr(S_{n-1} = k | \theta = 1) - Pr(S_{n-1} = k | \theta = -1) \right].
$$

(4)
No closed form solution exists for these probabilities for $p \in (1/2, 1)$, but we can derive a recursion formula to compute these probabilities numerically (in the Appendix).

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

$$E[\tilde{V}(\lambda)] = \lambda \cdot \frac{1}{N} \sum_{n=1}^{N} E[\tilde{V}_{OC,n}(\lambda)] + (1 - \lambda) \cdot \frac{1}{N} \sum_{n=1}^{N} E[\tilde{V}_{R,n}(\lambda)]$$

$$= \lambda \cdot E[\tilde{V}_{OC}(\lambda)] + (1 - \lambda) \cdot E[\tilde{V}_{R}(\lambda)]. \quad (5)$$

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; and hurts in that entrepreneurs make more frequent mistakes, which hurts themselves and thus lowers the average group payoff.

**The solution: The socially optimal proportion of Entrepreneurs ($\lambda^*$)**

We begin by determining the optimal proportion of entrepreneurs from a social welfare perspective, $\lambda^*$. The social welfare function is defined as $E[\tilde{V}(\lambda)]$ in (5), which is maximized by some $\lambda^*$. The following proposition describes the impact 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 the proportion of entrepreneurs in the group ($\lambda$), with the degree of overconfidence among entrepreneurs ($k(p'; p)$), and with the size of the group $N$.

2. The limiting probability ($n \to \infty$) of being in an incorrect cascade equals $q^2/(p^2 + q^2)$ if $\lambda = 0$ and $q^k/(p^k + q^k)$ for all $\lambda > 0$.

**Proof:** See Appendix A.1.

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

There are three parameters in our model that influence the optimal proportion of entrepreneurs ($\lambda^*$): the group size ($N$), the information precision ($p$), and the degree of overconfidence ($k(p'; p)$).

Unfortunately, the effect of the three parameters on $\lambda^*$ cannot generally be solved analytically. However, numerical simulations show that certain directional influences of the parameters $p$, $p'$, and $N$ on the solution $\lambda^*$ are present—and to the extent that we cover the relevant parameter space, we can conjecture that they are pervasive.

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.

Signal Precision ($p$). Figure 3 shows that the optimal proportion of entrepreneurs decreases in the private signal precision ($p$)—except in rare border cases (not plotted). Typically, when

---

8When overconfidence is moderate, the overall group payoff can increase, then decrease, then increase again in lambda. In rare cases, the first local maximum exceeds the second local maximum for small $p$, then for larger $p$ the second local maximum dominates. Because the latter decreases in $p$, the optimal $\lambda^*$ may not monotonically decrease in $p$. Further, the switching across maxima can lead to discontinuities, e.g., as in figure 4.
$p$ is low, there are more bad cascades in the absence of information aggregation and the cost to 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 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.

Insert Figure 4 about here.

**Group Size ($N$).** Figure 4 shows that the effect of the group size ($N$) on the optimal $\lambda^*$ is ambiguous and depends on the level of overconfidence. 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 high 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, $\lambda^* = 1$ for sufficiently large $N$.

It is noteworthy that the positive information externality works well only when groups are sufficiently large. There is a threshold 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 im-
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. In such cases, we would expect to see an explosion of non-conforming, irrational behavior.\(^9\)

**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 had 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. Conversely, information aggregation is better when there is much trust and coordination.

**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)).

\(^9\)Extreme economies of scale are common to models in which more information can freely benefit everyone, and both average social welfare and the optimal number (not proportion!) of entrepreneurs always increases with $N$. However, in real life, the coefficient of relatedness (the extent to which individuals share genes with other members in their group; discussed in more detail in footnote 13) could decrease with the group size $N$. Therefore, an empiricist would have to carefully control for the mechanism by which overconfidence can survive in an evolutionary setting: On the one hand, models of altruism built on relatedness could predict lower levels of entrepreneurship in larger groups. On the other hand, as $N$ becomes larger, there could be larger potential social benefits and presumably larger social structures. Thus, a direct transfer of resources to entrepreneurs would be easier to accomplish, and models based on direct transfer of resources could predict higher levels of entrepreneurship in larger groups.
**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. 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.

**Other Externalities** Information aggregation is more valuable if payoff/network externalities are weak. (Positive payoff externalities make any choice taken by the majority the correct choice. Negative payoff externalities would induce individuals not to conform, but to act on their own information anyway.)

### 3 The public goods problem and intra-group survival

**An evolutionary stable strategy (\(A^{ESS}\))**

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. Despite some modeling complications in computing the equilibrium proportion of entrepreneurs, the ultimate outcome—that entrepreneurs can survive—is robust. This is because the benefits to the group can be several orders of magnitude greater than the costs to the individual. 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:

**Indirect genetic benefits (i.e., kinship/relatedness).** *Hamilton's rule* (e.g., Smith (1989, 169f); Boyd and Silk (1997, 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.\(^{10}\)

Consider a scenario in which 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 which contains a (family with a) mutation that gives rise to 20 moderately overconfident individuals with \(k(p'; p) = 4\).\(^{11}\) 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.\(^{12}\)

In this scenario, the expected marginal cost to 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 ~ 3.8% of the group are other entrepreneurs (the

\(^{10}\)We abstract away from diploid reproduction and consider that the direct benefits of entrepreneurship are higher for normal types than entrepreneurs, because entrepreneurs do not use public information as efficiently as normal types.

\(^{11}\)In standard biology models, simultaneous within-group appearance is assumed in that members of subgroups interact only with other members of their subgroup (although these members need not be of the same genotype). Similarly, in Eshel, Samuelson, and Shaked (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 more likely benefit other altruists.

\(^{12}\)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 is 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 which employ *Hamilton's Rule* to show that biological altruism can increase and is usually accomplished by computing payoffs over paired subgroups.
Furthermore, each of the 19 entrepreneurs garner 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 twentieth entrepreneur. As a 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 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.)

**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 $\lambda^*$. 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.\(^{14}\)

\(^{13}\)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 non-existent 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%.

\(^{14}\)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 only 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
**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.

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

**A displacement model**

*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 $\lambda$-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 $E[V_{OC}(\lambda_{g,t})]$; normal individuals receive $E[V_R(\lambda_{g,t})]$; and the average group payoff is $\lambda_{g,t}E[V_{OC}(\lambda_{g,t})] + (1 - \lambda_{g,t})E[V_R(\lambda_{g,t})]$.

*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 $\lambda_{t+1}$ is

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

also exert pressure towards other, non-overconfidence based mechanisms that help to resolve the information aggregation problem differently, e.g., with culture and conversation.
where $g^*$ denotes the winning group, $t$ a subscript for 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 $N = 100$, 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, normal individuals receive 0.2417, for a group average of 0.2378. Entrepreneurs in group B expect to receive 0.2126, normal types expect 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. With a $w$ coefficient of 1, entrepreneurs constitute $0.1 \cdot 0.2126/0.2603 \sim 8.2\%$ of the population in the next generation, which is larger than the 7.5% it was in the previous generation.\footnote{With a $w$ coefficient of 2, normal individuals would constitute only 0.008% of the next generation. Note that, in equilibrium, a higher $w$ coefficient does not necessarily induce lower survival of entrepreneurs. By rapidly displacing entrepreneurs within their own groups, groups with a large number of normal individuals may themselves be so rapidly outcompeted by groups consisting exclusively of entrepreneurs that they disappear entirely. (This is especially the case if the group optimal solution $\lambda^*$ is close to 1.) The same effect has long been recognized in biology: for example, diseases that are too efficient and kill their hosts too rapidly have difficulty spreading in the population.}

One can compute a matrix of the resulting proportion of entrepreneurs, $\lambda_{t+1}$, 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.

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.

**Equilibrium Distributions.** We can now specify our equilibrium definition. A probability distribution over lambda groups, \( f_t(\lambda) \), is an evolutionary stable strategy (ESS) if it is stable from generation \( t \) to generation \( t + 1 \):

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

The degenerate distribution \( \lambda = 0 \) with probability 1, which always yields groups with only rational types, is an ESS, because entrepreneurs cannot appear. Similarly, the degenerate distribution \( \lambda = 1 \) with probability 1 (only entrepreneurs) is an ESS. However, as explained above, no other degenerate distribution can be 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.

**Computing an ESS.** 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 \( i \in \{0, 1, 2, ..., N\} \). To compute, e.g., the number of groups with \( \lambda_{t+1} = 8/100 \) entrepreneurs in the next generation, we need to consider all possible competitive scenarios in this generation that can produce \( 8/100 \) 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 proportion \( \lambda_{t+1} = 0.082 \) entrepreneurs. To maintain discreteness, we assume that the next generation contains proportion \( 8/100 \) entrepreneurs with probability 80\% and \( 9/100 \) entrepreneurs with probability 20\%. The probability that \( \lambda_{A,t} = 0.05 \) and \( \lambda_{B,t} = 0.1 \) groups meet is \( \pi_{5,t} \cdot \pi_{10,t} \); consequently, this scenario contributes probability mass \( \pi_{5,t} \cdot \pi_{10,t} \cdot 0.8 \) to the probability of having \( \pi_{8,t+1} \) in the next generation and \( \pi_{5,t} \cdot \pi_{10,t} \cdot 0.2 \) to the probability of having \( \pi_{9,t+1} \) in the next generation. In general, to find \( \pi_{X,t+1} \), one must integrate 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 $h(\cdot)$ apports probability mass to adjacent $\lambda$ fractions to maintain the discreteness of the distribution. More specifically, define $\gamma_{i,j} = \text{int}[\lambda_{t+1}(\lambda_{i,t}, \lambda_{j,t}) \cdot N]$, the integer portion of the number of entrepreneurs and $m_{i,j} = \lambda_{t+1}(\lambda_{i,t}, \lambda_{j,t}) - \gamma_{i,j}$ the remainder. Then

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

The ESS equilibrium 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. Figure 5 graphs a set of viable ESS distributions in the example case. In this case, the group-optimal $\lambda^*$ is 1 (all entrepreneurs)—and the distribution $\lambda = 1$ with probability 1 is also the Pareto-optimal ESS.\(^\text{16}\)

This is a general result: If the group optimal proportion $\lambda^*$ is 1, a (Pareto-dominating) equilibrium exists in which no normal types can survive. Groups with only entrepreneurs wipe out all other groups, and thus normal individuals have no chance to replicate.\(^\text{17}\)

However, when $\lambda^* \neq 1$, 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 figure shows that entrepreneurs survive in the Pareto-preferred equilibrium, but the average frequency of entrepreneurs in this ESS distribution is "only" about 1-2%—an order of magnitude lower than the group-optimum of $\lambda^* = 0.425$.

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 in which only a certain proportion within the population can end up

\(^{16}\text{One could make an argument that nature has evolved mechanisms to converge on the Pareto-dominant ESS distribution, but this is beyond the scope of our paper.}\)

\(^{17}\text{This is also not artificial to our severe penalty for the losing group: even if the losing group shrank only slowly, the } \lambda = 1 \text{ group which does not experience internal selection pressure against entrepreneurs would still end up eventually displacing all other groups.}\)
being entrepreneurs. Unfortunately, a thorough comparative statics analysis is not possible because a unique ESS distribution does not exist.

**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.\(^{18,19}\) 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."

Although our model has been based on evolutionary principles, we did not base its algebra on diploid selection for two reasons: First, although this could be done in principal, it would require substantial more overhead and an algebra mostly unfamiliar to economists. Second, our basic argument (overconfidence signals useful information) applies to cultural and social evolution. For example, firms can consist of groups of individuals and resources who succeed or fail together. To the degree that firms can maintain internal entrepreneurship, they can benefit from the evolutionary informational advantage identified in our paper.\(^{20}\)

\(^{18}\)The classical examples are kin selection (e.g., parenting sacrifices and the individual who jumps into a dangerous lake to rescue a sibling), and, within a larger group, the bird that cries out a warning when a predator approaches. villages, or even higher social structures.

\(^{19}\)There is one notable difference: social groups can dissolve into other social groups. The degree to which dissolution can occur depends on the willingness of a group to accept outsiders and to share benefits with them. The rate of acceptance and the type of accepted individuals may also be subject to evolutionary pressure: groups that consist of altruists may survive better if they refuse to accept outsiders, especially non-altruistic outsiders! It is not clear whether permitting between-group transfers speeds or slows the convergence to the group-optimal \(\lambda\).

\(^{20}\)Firms can both assure mutual presence of entrepreneurs (so that entrepreneurs benefit one another) and appropriate internal transfer payments.
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 problem with this argument is that the benefit to overconfidence rests on the willingness of other individuals to let themselves be deceived—and it is not clear why they would allow themselves to be deceived. Those non-entrepreneurial individuals who see through the deception would be more likely to survive, relative to individuals who buy into “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 non-conforming. 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. 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 above-mentioned risky benefits, there are certainly other real-world facets of entrepreneurship that we have omitted. Still, the infor-

\[21\text{Models in which individuals can either follow or learn (but not both) and in which the marginal cost/benefits to both activities are equal in equilibrium can be found in Boyd and Richerson (1988) and Rogers (1988). This typically results in learning that is suboptimal from a group-perspective. Biological models in which some, but not all individuals pursue an activity in their own self-interest, and in which an optimum interior proportion develops, can be found in Smith (1974), Cornell and Roll (1981), and Weibull (1995).}\]
mational benefits of entrepreneurship discussed in our paper are likely to persist in a much richer model than was 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 which 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 then 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

Recursion formula for state-time probabilities

Let $\pi^n_s$ 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 $E[\tilde{V}_{R,n}(\lambda)]$ and $E[\tilde{V}_{OC,n}(\lambda)]$ in equations 1 and 4, respectively. The recursion formulae for deriving these probabilities are as follows:

$$
\begin{align*}
\pi^n_0 &= p\pi^{n-1}_1 + q\pi^{n-1}_0 \\
\pi^n_1 &= \lambda p\pi^{n-1}_2 + q\pi^{n-1}_0 \\
\pi^n_2 &= \lambda q\pi^{n-1}_2 + p\pi^{n-1}_0 \\
\pi^n_{-2} &= \lambda p\pi^{n-1}_{-3} + q\pi^{n-1}_{-1} + (1 - \lambda)\pi^{n-1}_{-2} \\
\pi^n_{-1} &= \lambda q\pi^{n-1}_{-3} + p\pi^{n-1}_{-1} + (1 - \lambda)\pi^{n-1}_{-2} \\
\pi^n_s &= \lambda q\pi^{n-1}_{s+1} + \lambda p\pi^{n-1}_{s-1} + (1 - \lambda)\pi^{n-1}_s \quad \text{for} \ 2 < |s| < k - 1 \\
\pi^n_{-k+1} &= \lambda q\pi^{n-1}_{-k+2} + (1 - \lambda)\pi^{n-1}_{-k+1} \\
\pi^n_{k-1} &= \lambda p\pi^{n-1}_{k-2} + (1 - \lambda)\pi^{n-1}_{k-1} \\
\pi^n_{-k} &= \lambda q\pi^{n-1}_{-k+1} + \pi^{n-1}_{-k} \\
\pi^n_k &= \lambda p\pi^{n-1}_{k-1} + \pi^{n-1}_k
\end{align*}
$$

(A1)

with the starting value given by $\pi^0_0 = 1$ if $k \geq 4$.

Proof of proposition 1

Proof of Part 1: Fix $n$ and let $X_i = 1 \ (X_i = -1)$ if the $i^{th}$ 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 $n^{th}$ arrival if $\lambda = \lambda_1$ and let $Y_2(n)$ denote the random number of publicly inferred signals by the $n^{th}$ arrival if $\lambda = \lambda_2$. Without loss of generality, let $\lambda_1 > \lambda_2$.

The state of publicly observed information at the $n^{th}$ 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 $n^{th}$ 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 (θ = 1). Thus, the probability of making an incorrect decision when \( \lambda = \lambda_j \) is:

\[
\begin{align*}
\Pr[S_n^j \geq 1|\theta = -1] \cdot \Pr[\theta = -1] &+ \Pr[S_n^j \leq 0|\theta = 1] \cdot \Pr[\theta = 1] \\
= \frac{1}{2} \cdot \Pr[S_n^j \geq 1|\theta = -1] &+ \frac{1}{2} \cdot \Pr[S_n^j \leq 0|\theta = 1] \\
= \frac{1}{2} \cdot \Pr[S_n^j \leq -1|\theta = 1] &+ \frac{1}{2} \cdot \Pr[S_n^j \leq 0|\theta = 1].
\end{align*}
\]

(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,

\[
\begin{align*}
\frac{1}{2} \cdot \Pr[S_n^1 \leq -1|\theta = 1] &+ \frac{1}{2} \cdot \Pr[S_n^1 \leq 0|\theta = 1] \\
= \frac{1}{2} \cdot E[1_{S_n^1 \leq -1|\theta = 1} + 1_{S_n^1 \leq 0|\theta = 1}] \\
= \frac{1}{2} \cdot E[E[1_{S_n^1 \leq -1|\theta = 1} + 1_{S_n^1 \leq 0|\theta = 1}|Y_1(n)]] \\
= \frac{1}{2} \cdot E[f(Y_1(n))] \\
\leq \frac{1}{2} \cdot E[f(Y_2(n))] \\
= \frac{1}{2} \cdot \Pr[S_n^2 \leq -1|\theta = 1] &+ \frac{1}{2} \cdot \Pr[S_n^2 \leq 0|\theta = 1].
\end{align*}
\]

(A3)

where the inequality follows from the fact that \( f(\cdot) \) is a decreasing function because the \( X_i \) are iid 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 \( \lambda \), simply let \( Y_1(n) \) denote the random number of publicly inferred signals by the \( n^{th} \) arrival if \( k = k_1 \) and let \( Y_2(n) \) denote the random number of publicly inferred signals by the \( n^{th} \) 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 iid with \( E[X_i] > 0 \) the probability of making an incorrect decision decreases with each successive arrival when \( \lambda \) and \( k \) are fixed. Thus, the greater is \( N \) the smaller is the \textit{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)).

Q.E.D.
References


Figure 1 shows 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. 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 inner-most 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 plots the optimal proportion of entrepreneurs ($\lambda^*$) as a function of the degree of overconfidence ($k$) for two values of individuals' signal precision $p$ and group size $N$. The figure shows that when the degree of overconfidence becomes more extreme, it is better for groups to contain fewer entrepreneurs. When overconfidence is modest, however, it can be optimal for groups to contain only entrepreneurs.
Figure 3 plots the optimal proportion of entrepreneurs ($\lambda^*$) as a function of each individual's private signal precision ($p$). Holding group size constant (see legend), the optimal proportion of entrepreneurs decreases with private signal precision; over a certain range, the optimal proportion of entrepreneurs can be either one or zero.
Figure 4 plots the optimal proportion of entrepreneurs ($\lambda^*$) as a function of group size ($N$) for various degrees of overconfidence ($k$) and holding private signal precision constant at $p = 0.6$. (A higher $p$ would shift the respective functions towards the south-east.) When overconfidence is infinite, the optimal proportion of entrepreneurs is zero for small $N$, then reaches a maximum, and finally asymptotes towards zero. When 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 this critical population size.
Figure 5 plots some evolutionary stable strategy distributions ($\pi_i, \forall i$) when population size $N$ is 100, signal precision $p$ is 0.51, and overconfidence $k(p'p)$ is a modest 4. In this example, the socially optimal proportion of entrepreneurs is 100%. Although the (unplotted) Pareto-optimal ESS distribution has mass 100% at $\lambda = 1$, the plotted distributions are also stable from an evolutionary perspective. For example, the rightmost plotted distribution shows that if 1% of groups in the population have 90 entrepreneurs, 21.2% have 91 entrepreneurs, and 77.8% have 92 entrepreneurs, then so will the next generation.
**Figure 6** plots some evolutionary stable strategy distributions when population size $N$ is 100, signal precision $p$ is 0.51, and overconfidence ($k$) is a large 12. 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.