Overconfidence and Excess Entry: An Experimental Approach

By COLIN CAMERER AND DAN LOVALLO*

Psychological studies show that most people are overconfident about their own relative abilities, and unreasonably optimistic about their futures (e.g., Neil D. Weinstein, 1980; Shelly E. Taylor and J. D. Brown, 1988). When assessing their position in a distribution of peers on almost any positive trait—like driving ability (Ola Svenson, 1981), income prospects, or longevity—a vast majority of people say they are above the average, although of course, only half can be (if the trait is symmetrically distributed).[1]

This paper explores whether optimistic biases could plausibly and predictably influence economic behavior in one particular setting—entry into competitive games or markets. Many empirical studies show that most new businesses fail within a few years. For example, using plant-level data from the U.S. Census of Manufacturers spanning 1963–1982, Timothy Dunne et al. (1988) estimated that 61.5 percent of all entrants exited within five years and 79.6 percent exited within ten years. Most of these exits are failures (see also Daniel Shapiro and R. S. Khemani, 1987; Dunne et al., 1989a, b; Paul A. Geroski, 1991; John R. Baldwin, 1995).

Some possible explanations for the high rate of business failure are reviewed below. In this paper we consider the hypothesis that business failure is a result of managers acting on the optimism about relative skill they exhibit in surveys (e.g., James March and Zur Shapira, 1987). This hypothesis is worth exploring because it is consistent with so much psychological evidence, and because optimistic overentry will persist if the performance feedback necessary to correct it is relatively noisy, infrequent, or slow.

The idea that overconfidence causes business entry mistakes has, of course, been suggested before (e.g., Richard Roll, 1986) but has not been directly tested by measuring economic decisions and personal overconfidence simultaneously. To link the two we created an experimental setting with basic features of business entry situations. In the experiments, the success of entering subjects depends on their relative skill (compared to other entrants). Most subjects who enter think the total profit earned by all entrants will be negative, but their own profit will be positive. The findings are consistent with the prediction that overconfidence leads to excessive business entry.

The experiments also develop a paradigm in which business entry and other skill-based competitions (e.g., labor-market tournaments) could be studied further. The paradigm extends typical economics experiments by including a potentially potent psychological variable—relative skill perceptions—and also extends typical psychology experiments on overconfidence by adding financial incentives for judging one's

* Camerer: Division of the Humanities and Social Sciences 228-77, California Institute of Technology, Pasadena, CA 91125 (e-mail: camerer@hss.caltech.edu); Lovallo: Wharton School, University of Pennsylvania, Philadelphia, PA 19104. Help and comments were received from Daniel Kahneman, Marc Knez, Matthew Rabin, David Teece, Dick Thaler, participants at the MacArthur Foundation Preferences Group, the 1995 J/DM Society, workshops at the Universities of Chicago and Colorado, UCLA, Harvard Business School, and Wharton, and several anonymous referees. Gail Nash provided superb secretarial help and Roberto Weber provided research assistance. The research was funded by National Science Foundation Grant No. SBR 95-11001.

[1] There are interesting exceptions—most people demurely say they are not in the very top decile or quintile, but merely above average; for many traits, women are less optimistic than men (and even overly pessimistic; e.g., Eleanor E. Maccoby and Carol N. Jacklin, 1974); and clinically depressed patients are not optimistic (e.g., Lauren B. Alloy and Anthony H. Ahrens, 1987). The latter finding calls into question the common psychiatric presumption that "realistic" people are well adjusted and happy, and also raises the question of whether unrealistic optimism might be evolutionarily adaptive (e.g., Lionel Tiger, 1979) or socially beneficial (Giovanni Dosi and Lovallo, 1997). Michael Waldman (1994) shows how such optimism could be evolutionarily stable, and mentions conditions under which gender differences like those observed empirically could arise.

skill accurately and a clear definition of the skill one is judging.[2]

Of course, experimental data are hardly conclusive evidence that overconfidence plays a role in actual entry decisions by firms. A bigger scientific payoff comes when experimental observations suggest a new phenomenon that might be studied in the field. Our data suggest a new phenomenon we call ''reference group neglect.'' Excess entry is much larger when subjects volunteered to participate knowing that payoffs would depend on skill. These self-selected subjects seem to neglect the fact that they are competing with a reference group of subjects who all think they are skilled too. (Neglecting the increased level of competition is like the neglect of adverse selection which leads to the ''winner's curse'' in bidding.)

I. Possible Explanations for Entrant Failure

There are three primary explanations for the frequency of entrant failure. The first explanation is that failures are frequent because entrants have only brief opportunities to make money. In this view, failures are actually hit-and-run entries that are profitable but brief.

A second explanation is that business entries are expensive lottery tickets with positively skewed returns. In this view, although most firms expect to lose money and fail, entry still maximizes expected profits because the payoffs to success are very large. There are two variants of this argument: First, if small-business owners are risk preferring or get psychic income from running businesses, then the expected utility from entering might be high even if expected profit is low. Second, it is well known from multiarmed bandit problems that when sampling from unknown distributions of possible payoffs (such as career paths or profitable industries), it may pay to sample from ''arms'' with negative expected payoffs if the possible payoffs from those arms is large (because sampling provides information about which arms to choose in the future). Models of occupational choice provide a clear life-cycle prediction based on this sampling motive for entry, since people should bear the risk of failure early in their careers, but not later (e.g., Robert A. Miller, 1984).

The third explanation is that many entry decisions are mistakes, made by boundedly rational decision makers. Firms could mistakenly enter too often for two different reasons—they know their own skills but fail to appreciate how many competitors there will be (they have ''competitive blind spots''), or they forecast competition accurately but overconfidently think their firm will succeed while most others will fail.

In a natural setting it is difficult to distinguish between these three explanations for high failure rates. The overconfidence explanation is particularly hard to establish because it predicts that firms will enter even if they expect negative industry profits. But even if cumulative industry profits are actually negative at some point in time, it is possible positive returns will roll in later (or the industry simply made a large unpredictable forecasting mistake). So it is hard to imagine how to establish conclusively that *expected* industry returns were negative.

While more field research is surely worthwhile, some progress might be made in the laboratory. In an experiment, everything needed to distinguish the three theories—entry decisions, forecasts of industry profits, and forecasts of the number of total entrants—can be measured. If subjects forecast positive industry profits and enter, the rational-entry theories appear correct. If subjects forecast positive industry profits, but they underestimate the amount of entry and industry profit turns out to be negative, then the blind spots story appears correct. If subjects accurately forecast negative industry profits, and enter anyway, then the overconfidence explanation appears correct.

II. Experimental Design

Our experiments extend a paradigm first used by Daniel Kahneman (1988), Jim Brander, and Richard Thaler, then explored more throughly by Amnon Rapoport and colleagues.

[2] Earlier studies showed that overconfidence is larger when traits are defined ambiguously—''driving ability'' is more ambiguous than ''ability to brake quickly to avoid an accident'' (David Dunning et al., 1989; Lovallo, 1996).

In their game, N players choose simultaneously, and without communicating, whether to enter a market or not. The market "capacity" is a preannounced number, c. If players stay out they earn a payment K. If the total number of entrants is E, the entrants each earn $K + rK(c - E)$ (with $rK > 0$). The optimal behavior is simple: Players want to enter only if the number of expected entrants (including themselves) is less than the capacity c. If they do enter, players prefer the number of entrants to be as small as possible. The interesting questions are whether the right number of players enter (is E around c?), whether E changes with c, and how players figure out whether to enter or not.

Kahneman (1988) was surprised to see that the number of entrants, E, was typically in the range $[c - 2, c + 2]$ even though subjects could not communicate or coordinate their decisions in any explicit way. "To a psychologist," he wrote, "it looks like magic." Rapoport (1995) replicated the results using Ph.D. students playing for much larger stakes. He also found that subjects entered a bit too frequently at first, but gradually E converged very close to c. E and c were highly correlated across trials. Extensions by James Sundali et al. (1995) and Rapoport et al. (1998a) replicated the earlier findings. Rapoport et al. (1998b) introduced probabilistic payoffs and showed that deviations from equilibrium entry could be parsimoniously explained by nonlinear transformations of entry probabilities.

Our experiments extend this paradigm in four ways: Payoffs depend on a subject's rank (relative to other entrants); ranks depend on either a chance device, or on a subject's skill; subjects in some experiments are told in advance that the experiment depends on skill (and hence, more skilled subjects presumably self-select into the experiment); and subjects forecast the number of entrants in each period.

Skill-dependent payoffs are the crucial new design feature. The early experiments capture an important aspect of entry—tacit coordination among potential entrants to avoid excess entry—but all entrants earned the same amount. In naturally occurring settings, some entrants win and others lose, due at least partly to differences in managerial skill (see Kenneth R. MacCrimmon and Donald A. Wehrung, 1986). Besides being more realistic, differences in payoffs based on skill allow the possibility that overconfidence will lead to excess entry.

TABLE 1—RANK-BASED PAYOFFS

	Payoff for successful entrants as a function of "c"			
Rank	2	4	6	8
1	33	20	14	11
2	17	15	12	10
3		10	10	8
4		5	7	7
5			5	6
6			2	4
7				3
8				2

Table 1 shows how payoffs depend on a subject's rank and on the market capacity c. The top c entrants share \$50 proportionally, with higher-ranking entrants earning more. All entrants ranking below the top c lose \$10. For example, if the market capacity $c = 2$, then the highest-ranked entrant receives \$33, the second highest-ranked entrant receives \$17, and any lower-ranked entrant loses \$10. (Subjects are staked \$10 initially.) Notice that if the number of entrants is exactly $c + 5$, then the total payoff to all entering subjects ("industry profit") is zero; if there are more than $c + 5$ entrants, the average entrant loses money.

Actual ranks are assigned in two different ways: Each subject is ranked by a random drawing, and also ranked according to his relative performance on a skill or trivia task. Skill ranks are determined by how many questions subjects answer correctly on a sample of 10 logic puzzles (sessions 1–2) or trivia questions about sports or current events (sessions 3–8). It is important to stress that subjects' ranks were not determined until the end of the experiment, *after* they made all their entry decisions in both the skill and random conditions.

Here are the steps in each experimental session:

1. Before the experiment, subjects were recruited using either standard recruiting instructions or "self-selection" instructions.

TABLE 2—MARKET CAPACITY "c" VALUES

Rounds	Experiment 1	Experiment 2	Experiments 3–6	Experiments 7 & 8
1	2	8	2	4
2	4	4	6	2
3	8	2	4	6
4	6	6	4	8
5	4	4	2	6
6	2	2	6	4
7	8	8	4	2
8	6	6	6	8
9	4	4	2	6
10	6	2	6	4
11	8	8	4	2
12	2	6	2	8

In the self-selection condition, subjects were asked if they would like to volunteer for an experiment in which performance on sports or current events trivia would determine their payoff, and people who were very good might earn a considerable sum of money. (They were also reminded in the experimental instructions that all subjects were recruited this way.)

2. Subjects were seated in a large classroom where they could not see each other's materials. The instructions were read aloud and a comprehension test was given to guarantee understanding of the payoff table. The two types of ranking systems were explained and subjects were shown examples of the skill questions, along with sample answers. Subjects were informed that there would be two sequences of 12 rounds for each condition—one for the random rank and another for the skill rank. Subjects were also informed that the decisions they made for one of the rounds, chosen randomly, would determine their payoff.

Individual rounds proceed as follows:

3. Subjects were told whether skill or random ranks are being used in that round, and the capacity c. Table 2 shows the capacities used in each round. The same sequence of capacities was used in the two consecutive conditions within a session.
4. Subjects privately forecasted how many entrants they expected would enter (including themselves) in the round. They earned \$0.25 for each forecast that was correct. These forecasts distinguish the hypothesis that too many subjects enter because they underestimate the number of competitors ("blind spots") from the hypothesis that subjects forecast entry accurately, but entrants all think they are above average.
5. Subjects made their entry decisions privately and simultaneously.[3]
6. Entry decisions were recorded and subjects were told how many total entrants there were in the round. Thus, the only feedback that subjects received after each round is the total number of entrants for each period.
7. At the end of the experimental session, after all of the rounds in both conditions were played, subjects either solved puzzles or took the trivia quiz, and their skill rank was determined and announced. Then one of the subjects randomly chose one of the 24 rounds and subjects' earnings from that round were computed and paid to them.

It is important to reiterate that the only feedback subjects got throughout the session of 24 rounds was the total number of entrants per

[3] In one session, not reported here, we allowed decisions to be made sequentially. This means that a subject moving after $c + 5$ entrants have already entered knows for sure that the total payment to subjects will be negative; entering in that condition is the strongest possible evidence that subjects are relatively overconfident. Roughly the same number of subjects entered in that session, but too few data are available from the single session to draw firmer conclusions.

TABLE 3—DESCRIPTION OF EXPERIMENTS

Experiment #	Sample	n	Selection procedure	Rank order
1	Chicago, undergraduates	12	random	R/S
2	Chicago, undergraduates	14	random	S/R
3	Wharton, undergraduates	16	random	R/S
4	Wharton, undergraduates	16	random	S/R
5	Wharton, undergraduates	16	self-selection	R/S
6	Wharton, undergraduates	16	self-selection	S/R
7	Chicago, M.B.A.'s	14	self-selection	R/S
8	Wharton, M.B.A.'s	14	self-selection	S/R

round. This design was chosen to model initial entry behavior by firms that do not learn much about their competitive advantage until after they incur substantial nonsalvageable fixed costs. The question of how post-entry feedback about performance impacts subsequent behavior is interesting, of course—it is certainly likely that overconfidence would be diminished if subjects were given a separate skill test and told their ranks after each round. But it is natural to begin by establishing whether overconfidence is present in the first place, before turning to the question of what forces make it go away.

The procedures described above were used in eight sessions. Table 3 summarizes differences in treatment variables across sessions.[4] In half of the experimental sessions the random-rank condition rounds were conducted first; in the other half the skill-rank rounds were first. Four sessions involved self-selected subjects (who knew trivia skill would help) and four sessions did not.

A. *Equilibrium Predictions*

Assuming risk neutrality, there are many pure-strategy Nash equilibria in which $c + 4$ or $c + 5$ subjects enter (the fifth subject is indifferent since he or she expects to earn zero from entering). Since the pure-strategy equilibria are necessarily asymmetric, it is hard to see how they might arise without communication or some coordinating device, like history, sequential moves, or public labels distinguishing subjects. There is also a unique symmetric mixed-strategy equilibrium in which (risk-neutral) players enter with a probability close to $(c + 5)/N$ (see Lovallo and Camerer, 1996).

Relaxing the assumption of risk neutrality, there is no way to determine the equilibrium number of entrants without measuring or making specific assumptions about subjects' risk preferences.[5] The random-rank condition gives an empirical estimate of observed equilibrium without having to impose any a priori assumption about risk preferences. Since subjects participate in both random- and skill-rank conditions, their decisions in the random-rank condition act as a within-subject control for risk preferences. The *difference* in the number of entrants in the random and skill conditions is the primary measure of interest.

III. Results

A. *Does Overconfidence About Skill Increase Entry?*

Table 4 lists the total amount of money earned by subjects (''industry profit'') per

[4] Business students, especially M.B.A.'s, are an appropriate sample because many go on to start businesses or participate in corporate entry decisions (e.g., entrepreneurship is the fifth most popular major among Wharton M.B.A.'s).

[5] An alternative is to try to induce risk neutrality (or some other specific degree of risk aversion) by paying subjects in units of probability (see Joyce E. Berg et al., 1986). We chose to use the random-rank condition because the probability procedure does not induce risk neutrality reliably (see Reinhard Selten et al., 1995; cf., Vesna Prasnikar, 1996), and the random-rank condition is equally theoretically valid, and simpler.

TABLE 4—INDUSTRY PROFIT BY ROUND

Experiment #	n	Rounds 1	2	3	4	5	6	7	8	9	10	11	12	Total
Profit for random-rank condition														
1	12	50	50	20	30	40	30	20	50	30	40	20	40	420
2	14	0	−10	10	20	−10	10	20	10	0	0	30	20	100
3	16	10	50	20	40	10	20	30	40	20	40	30	20	330
4	16	0	10	10	20	10	−10	0	10	20	10	0	20	100
5	16	20	10	10	10	0	0	30	20	−10	0	0	0	90
6	16	30	20	10	0	−10	30	20	10	10	30	10	20	180
7	14	10	20	40	20	30	40	−30	40	10	0	0	20	200
8	14	20	10	0	30	30	0	10	10	20	10	20	40	200
Profit for skill-rank condition														
1	12	50	0	20	10	30	10	20	10	40	10	10	30	240
2	14	0	−10	10	20	−10	10	20	10	0	0	30	20	100
3	16	10	20	10	20	0	10	20	10	10	30	20	10	180
4	16	0	0	20	20	10	−30	10	−10	−10	10	−20	0	0
5	16	−30	−20	−20	−10	−40	−10	−30	0	−30	−10	−20	0	−220
6	16	10	−40	−20	−30	−10	−30	−10	−20	−20	−10	0	0	−180
7	14	−40	−10	−10	0	−20	−10	−40	0	0	0	−10	0	−140
8	14	10	−10	−10	−10	−20	−20	−20	0	−20	10	−20	−20	−130

round in each experimental session, by rank condition. Recall that if c subjects enter, total industry profit is \$50. If $c + 5$ enter, total profits are 0.

The main question is whether there is more entry (and lower industry profit) when people are betting on their own relative skill rather than on a random device. The answer is "Yes": In the majority of the random-rank rounds (74/96 or 77 percent) industry profit is strictly positive[6] and total profit is negative only six times (6 percent). Average industry profit across rounds is \$16.87. In contrast, in the skill-rank rounds industry profit is strictly positive in only 37 rounds (40 percent) and negative in 41 (42 percent). Average profit across the skill-rank rounds is −\$1.56. The difference in average profits between the conditions is \$18.43, which is about two extra entrants per round in the skill conditions (about a third of the number expected to *not* enter).

A powerful statistical test of significance exploits the yoked design by comparing industry profit in each pair of skill-rank and random-rank periods in exactly the same periods of experimental sessions t and $t + 1$ (for $t = 1, 3, 5, 7$). In this comparison, each pair of periods has exactly the same location in experimental time and the same value of c, and differ only in whether ranks were due to skill or chance. (Fixed effects of periods, self-selection, and subject pool are all controlled for by this comparison.) A matched-pair t-test using these comparisons yields $t = -7.43$ (dof = 95, $p < 0.0001$). Industry profits under skill-based entry are clearly lower.

The next question is whether reference group neglect produces a larger skill-random entry differential in the experiments with self-selected subjects. The answer appears to be "Yes." In sessions without self-selection (1–4), the average per-period industry profit is \$19.79 and \$10.83 for the random and skill

[6] This is also consistent with tacit collusion among risk-neutral players, since having exactly c entrants is the collusive solution (but is not a Nash equilibrium), or with some degree of risk aversion or (more likely) loss aversion.

TABLE 5—AVERAGE DIFFERENCE IN EXPECTED PROFITS PER ENTRANT BETWEEN RANDOM AND SKILL CONDITIONS

Measure	Experiment 1	Experiment 2	Experiment 3	Experiment 4	Experiment 5	Experiment 6	Experiment 7	Experiment 8	Total
$\Pi_r - \Pi_s$	1.635	0.477	−1.19	0.24	1.62	2.49	3.16	1.80	1.31
	(1.98)	(1.41)	(1.72)	(2.41)	(1.32)	(1.27)	(1.61)	(1.20)	(2.04)
# of S's with	10/12	10/13	3/11	7/14	12/13	12/13	13/13	11/12	78/101
$\Pi_r - \Pi_s < 0$	(83)	(77)	(27)	(50)	(92)	(92)	(100)	(92)	(77)
(percent)									
# of S's with	0/12	0/13	0/12	2/15	12/15	15/16	12/14	11/14	52/111
$\Pi_s < 0$	(0)	(0)	(0)	(13)	(80)	(94)	(86)	(79)	(47)
(percent)									

conditions, respectively—a difference of \$9.14, or about one extra entrant in the skill-based rounds. In sessions with self-selection (5–8) profit is \$13.96 in the random condition and −\$13.13 in the skill condition, which results in an entry differential of \$27.10—about three times as large as in the sessions without self-selection. Furthermore, in the experiments with self-selection, industry profits are positive in only 3 of the 48 skill-rank periods, compared with 34 of 48 in the non-self-selected sessions. A matched-pairs test comparing the skill-random profit differentials for matched periods between sessions 1–4 and 5–8 strongly rejects the hypothesis that differentials are the same in sessions with and without self-selection ($t(94) = -4.08$, $p < 0.001$). Reference group neglect clearly makes the overconfidence effect stronger.

B. *Expected Earnings Differences in Skill and Random Rounds*

The matched-pairs tests illustrate the effect of overconfidence on entry and demonstrate that self-selection makes the effect stronger. But these tests do not carefully control for all alternative explanations.[7] For example, the blind spots hypothesis suggests that excessive entry in the skill conditions may be due to players underforecasting how many others will enter.

To test this hypothesis, we use subject j's forecast F_{ijt} to compute the profit that subject j expects the average entrant to earn in round t of experiment i. If the capacity is c_{it} in that particular period, then the "expected average profit"—the amount of profit subject j thinks the average entrant will earn—is $(50-10*(F_{ijt} - c_{it}))/F_{ijt}$, which we denote by $E_j(\Pi_{ijt})$. This method effectively separates the blind spots hypothesis from the overconfidence hypothesis. Suppose, for example, that in skill conditions subjects are more apt to enter because they think fewer people will enter, not because they feel they are more skilled. Then their $E_j(\Pi_{ijt})$ values will be larger in the skill condition. Including $E_j(\Pi_{ijt})$ in an entry regression will then wipe out the effect spuriously attributed to skill.

If entering subjects are more overconfident in the skill rounds, then their expected average profits $E(\Pi_{ijt})$ will be smaller than in random rounds because the skilled subjects expect to earn more than the average entrant and, hence, are willing to enter even when the expected average profit is low. To test this prediction, Table 5 reports the difference between expected average profits in random rounds (denoted Π_r) and the same statistic in skill rounds (Π_s), using only the rounds in which a subject entered. The table shows three different measures for each session: The mean difference $\Pi_r - \Pi_s$ averaged across entering subjects, the number and percentage of subjects who have a negative mean (i.e., who expect less average profit in skill periods), and the number and percentage of subjects whose expected aver-

[7] Gender could be confounded with self-selection, too, since women may be less likely to volunteer for tasks which reward expertise in sports trivia (and are usually found to be less overconfident than men, in general). We controlled for this by only recruiting male subjects in sessions 3–8. Thus, the logit analysis of sessions 3–8 effectively controls for gender.

age profit is negative, on average, across skill periods.

In sessions without self-selection (1–4) the mean difference $\Pi_r - \Pi_s$ is generally positive and modestly significant—60 percent of the subjects expect to earn less in skill periods, but only a few subjects (4 percent) actually expect losses in skill periods. In the sessions with self-selection (4–8) the statistics are more striking: There are large, modestly significant average differences $\Pi_r - \Pi_s$ in all four sessions (almost all subjects expect to earn less in skill periods than in random periods), and 85 percent of the subjects have negative expected average profits in skill periods. The large majority of subjects in the self-selection sessions seem to be saying, "I expect the average entrant to lose money, but not me!"

C. *Regression Estimates of the Overconfidence Effect*

Another way to see the size and significance of all the variables' effects at once is a logit regression in which the dependent variable is subject j's 0-1 entry decision (enter = 1) in round t of experiment i, D_{ijt}. The logit includes controls for period-specific intercepts (to capture any period-by-period influences on entry), a subject pool dummy (MBA = 1), a self-selection condition dummy (RNG = 1), capacity c, and a skill-rank dummy (Skill = 1).

Table 6 shows the results of the logit regression of entry decisions.[8] Period-specific dummy variables were never significant and are not reported. Curiously, $E(\Pi_{ijt})$ enters with a negative sign, implying that when subjects expect high average profit they enter less often. This odd result is robust to several specifications but it does not disrupt inferences about skill.[9] The dummy variable MBA enters positively but the MBA*Skill interaction does not.

Most importantly, the effect of the Skill condition variable is significantly positive (t = 2.48) in the full model but the interaction of self-selection and skill, RNG*Skill, is insignificant. The middle column drops the uninteresting variables MBA, MBA*Skill, and the insignificant main effect of RNG. Then RNG*Skill becomes significant (t = 1.90), confirming that self-selection significantly increases the tendency to enter more frequently when payoffs depend on skill.[10] The right column excludes the RNG*Skill interaction, which increases the estimated coefficient and significance of the pure skill effect (t = 4.83). Comparing the log-likelihoods with and without RNG*Skill also shows that including it improves fit significantly (χ^2 = 3.6, p = 0.05), corroborating the results of the t-test.

D. *Additional Analyses: Forecasts and Equilibrium Behavior*

Since subjects forecasted the number of entrants in each period, we can test whether their forecasts reflect rational use of available information (see Lovallo and Camerer [1996] for more details). Forecasts are slightly biased: In random conditions subjects forecast about 0.30 entrants too high, and in skill conditions they forecast 0.50 entrants too low (the latter bias is significantly negative at $p <$ 0.05). We have no explanation for these small

[8] The regression uses only data from sessions 3–8 because sessions 1–2 used a different task (logic puzzles rather than trivia) and did not include self-selection as a treatment, so including it does not give much extra power for estimating the effect of RNG.

[9] We also included capacity c as a series of dummy variables (to capture nonlinearity in the effect of c on entry), and included interactions between $E(\Pi_{ijt})$ and Skill, and between $E(\Pi_{ijt})$ and c. None of these specifications improved the fit substantially or eliminated the significant negative coefficient on $E(\Pi_{ijt})$. We suspect the result occurs because when subjects plan to enter, they also forecast a lot of entry, so the expected average profit $E(\Pi_{ijt})$ is lower when they enter. This could be due to a "false consensus" in which subjects use their own decision as a clue about what others will do (and, because of optimism, they do not let their forecast inhibit their own entry).

[10] Excluding the RNG*Skill interaction from the second model raises the estimated coefficient and significance of the pure skill effect (0.450, t = 4.83), which suggests that the precision of the estimates of skill and RNG*Skill in the middle-column specification are a lot lower because of the colinearity between skill and RNG*Skill. Comparing the log-likelihoods with and without RNG*Skill also shows that including it improves fit significantly (χ^2 = 3.6, p = 0.05).

TABLE 6—LOGIT ESTIMATION OF ENTRY EQUATION (SESSIONS 3–8, n = 2,204)

	Dependent variable: Entry (=1)					
Variable	Estimate	(t-statistic)	Estimate	(t-statistic)	Estimate	(t-statistic)
Intercept	−0.887	(−3.28)	−0.855	(−3.40)	−0.865	(−3.21)
c	0.233	(6.93)	0.257	(7.99)	0.258	(8.06)
$E(\pi_{ijt})$	−0.129	(−6.93)	−0.126	(−6.65)	−0.144	(−8.72)
Skill	0.375	(2.48)	0.286	(2.26)	0.450	(4.83)
MBA	0.283	(1.76)				
RNG	0.011	(0.01)				
MBA*Skill	0.196	(0.83)				
RNG*Skill	0.078	(0.35)	0.299	(1.90)		
Log-likelihood	−1366.8		−1372.5		−1374.3	
Percent correct	64.84		64.02		64.34	

biases and do not attach much economic significance to them. For most subjects, forecasts pass the standard rationality tests because forecast errors are not predicted by observable information (i.e., by previous errors, or by the current forecast level). When errors *are* predictable, they tend to flip in the opposite direction of previous errors, and errors are positively correlated with levels (i.e., when forecasts are high, they are too high so the forecast error is positive).

Compared to other economics experiments in which paid forecasts have been gathered (cf., Camerer, 1995 pp. 609–12), the informational rationality of these forecasts is quite good. This fact is important because it means subjects are not *generally* irrational in processing information and they do not overenter because they underforecast the amount of competition. They are just overconfident about their relative skill.

The time series of matched-pair skill-random differentials in entry has a slight downward trend across periods. This raises the important question of whether the effect of overconfidence on entry would disappear if the experiment were run longer. A helpful way to forecast the answer is to fit a time-series model which estimates the long-run differential by extrapolating from 12 periods of data to what would happen if the experiment were run forever (see Camerer, 1987). Our working paper reports estimates from three different models. One model assumes partial adaptation of deviations from long-run equilibrium. Two other models assume those deviations drop with the reciprocal or reciprocal square root of the period number. The three techniques yield estimated differentials of 1.96, 1.79, and 1.34 (all of which are highly significant).

These numbers suggest that even if the experiment was repeated for a much longer time, one or two more subjects would enter when their payoffs depend on skill, relative to the number who enter than when payoffs are random. Keep in mind that an average of five or six subjects are predicted to stay out in each period (depending on the design). Two extra entrants means that more than a third of the number who are predicted to stay out actually enter.

IV. Discussion

Empirical studies show a high rate of business failure. We explored whether overconfidence about relative ability is part of the explanation for excessive failure by creating experimental entry games in which entrants' payoffs depend on their skill.

When subjects' post-entry payoffs are based on their own abilities, individuals tend to overestimate their chances of relative success and enter more frequently (compared to a condition in which payoffs do not depend on skill). The more surprising finding is that overconfidence is even stronger when subjects self-select into the experimental sessions, knowing their success will depend partly on their skill (and that others have self-selected too). In these sessions, there is so much entry that the average subject loses money in 34 out of 48 periods, and earns money in only four periods. This result suggests a new phenomenon specific to competition, "reference group neglect"—the tendency to underadjust to changes in the reference group one competes with.

Reference group neglect is one byproduct of a psychological phenomenon called the "inside view" (Kahneman and Lovallo, 1993). An inside view forecast is generated by focusing on the abilities and resources of a particular group, constructing scenarios of future progress, and extrapolating current trends. In contrast, an "outside view" ignores special details of the case at hand, constructs a class of cases similar to the current one, and guesses where the current case lies in that class (cf., Kahneman and Tversky, 1979). The inside view tells a colorful story; the outside view recites statistics. In the inside view, there is no special role for anticipation of the number of competitors or their abilities. In the outside view, the fact that most entries fail cannot be ignored.

Reference group neglect was nicely expressed by Joe Roth, chairman of Walt Disney Studios, when he was asked why so many expensive big-budget movies are released on the same weekends (such as Memorial Day and Independence Day). Roth replied:

> Hubris. Hubris. If you only think about your own business, you think, "I've got a good story department, I've got a good marketing department, we're going to go out and do this." *And you don't think that everybody else is thinking the same way.* In a given weekend in a year you'll have five movies open, and there's certainly not enough people to go around. (Emphasis ours; *Los Angeles Times,* 1996 p. F8.)

A. *Some Testable Economic Implications*

Experimental results are especially useful when they suggest implications which are testable with naturally occurring data. If people are generally overconfident about their relative abilities, then in industries or professions where overconfidence is likely to be largest, industry profits or total wages (including costs of training) may be negative. This brings us full circle to the empirical facts about business failure, and the difficulty of clearly establishing negative industry profit. The key to empirical tests which distinguish overconfidence from other explanations is to find variables which predict levels of overconfidence and see if they correlate with the tendency for overall profit to be negative. For example, when the criterion for success is more vague, people or firms should be more likely to overcompete, since ambiguity permits excess optimism. This implies that in professions where success can be achieved by different types of people, or industries with highly differentiated products, excess entry is more likely. For example, the skills required to be a successful model seem to be narrower than the skills required to be a successful actor. If so, your waiter at a Los Angeles restaurant is more likely to be an aspiring actor than an aspiring model.

The overconfidence hypothesis also predicts that people will prefer performance-based incentives schemes more often than standard theory predicts. Standard theory predicts that as output variance rises, principals who can bear risk should offer less output-sensitive contracts to agents (who presumably dislike risk). Overconfidence predicts that agents will be relatively insensitive to risk; indeed, when risk is high their overconfidence might lead them to *prefer* riskier contracts because they think they can beat the odds. There is some evidence from sharecropping that the standard prediction is wrong, and the overconfidence prediction may be right. Crops with larger yield variation are more likely to be farmed with cash leases, where farmers pay a fixed fee to lease the land and bear all the crop risk themselves (e.g., Douglas W. Allen and Dean Lueck, 1995). Other evidence from franchising and mining show that risk variables play a small role in contract determination; the existence of overconfidence may explain why.

Reference group neglect predicts that when agents compete based on skill, they will be insufficiently sensitive to the quality of competition. This has at least three testable implications.

First, people will gather too little data about the nature of their competitors when deciding whether to enter.[11]

Second, reference group neglect predicts that people will be insensitive to whether their competitors are *forced* to compete or *choose* to compete. Empirical tests could compare a situation in which entry is dictated by regulation or law, with a similar situation in which people can opt in or out. Reference group neglect predicts a higher failure rate in the latter cases.[12]

Third, in hierarchical tournaments where "winners" at one level advance to the next level, reference group neglect predicts that overconfidence will get stronger and stronger as people advance.[13] As workers win each level of the tournament, their success is certainly a positive signal of ability relative to the tournament losers left behind, but every other winner has received the same positive signal too. If winners neglect the fact that competition increases at each level, they will become more overconfident at each new level. Educational attainment might be an example. Freshman students at a highly selective college will be overconfident (neglecting the large increase in competition from their high school to college) and the effect will get stronger if they go to graduate school. Promotion tracks in businesses could produce the same pattern of snowballing overconfidence: Perhaps as cream rises to the top, hubris does too.

Reference group neglect predicts an opposite bias when workers lose tournaments and consider whether to enter "consolation" tournaments with other losers. Losing workers will be *under*confident if they neglect how easy the new competition is. For example, recently fired workers will have unusually long spells of unemployment (compared to the spell length predicted by optimal search theory), if they lick their wounds rather than compete with other recently fired workers by searching.

In addition to reference group neglect and its testable implications, an important implication of our study is methodological. In some settings with uncertainty, it is sensible to characterize economic agents as making decisions about random events, and use chance devices in the lab to mimic such events. However, when agents are betting on their own abilities, assuming that random luck and skill are the same is a mistake (cf., Linda Babcock and George Loewenstein, 1997). Indeed, we reach different conclusions about equilibrium predictions when we use skill-based payoffs instead of random payoffs—they enter more when betting on their skill. This is not to say that the subjects behave irrationally—indeed, they forecast the number of competitors quite well, and most pass tests of expectational rationality. They are simply overconfident; and the inside view which creates that confidence leads them to neglect the quality of their competition.

[11] For example, prospective doctoral students often do not ask about exam failure rates or what jobs all graduates get, and academics are surprisingly unfamiliar with acceptance rates of different journals they submit articles to. These "outside view" statistics only make cameo appearances in the success stories people project for themselves.

[12] An example is the difference between an army that drafts citizens to serve and a regime in which soldiers volunteer. Reference group neglect predicts that soldiers who aspire to become high-ranking officers will be too optimistic about their chances when they volunteer (because many volunteers will share the same ambition). A similar test could turn on differences between compulsory and voluntary education, or differences between required core courses and electives (students will be more overconfident about their success in the latter courses).

[13] Possible examples include political competition, "rat races" in professional firms, athletic competition, educational competition for entrance into elite colleges and graduate schools, and entertainment industries where entry is easy and a few "superstar" workers earn large rewards (such as acting and screenwriting).

REFERENCES

Allen, Douglas W. and Lueck, Dean. "Risk Preferences and the Economics of Contracts." *American Economic Review*, May 1995 (*Papers and Proceedings*), *85*(2), pp. 447–51.

Alloy, Lauren B. and Aherns, Anthony H. "Depression and Pessimism for the Future: Biased Use of Statistically Relevant Information in Predictions for Self and Others." *Journal of Personality and Social Psychology*, February 1987, *52*(2), pp. 366–78.

Babcock, Linda and Loewenstein, George. "Explaining Bargaining Impasse: The Role of Self-Serving Biases." *Journal of Economic Perspectives*, Winter 1997, *11*(1), pp. 109–26.

Baldwin, John R. *The dynamics of industrial competition*. Cambridge: Cambridge University Press, 1995.

Berg, Joyce E.; Daley, Lane A.; Dickhaut, John W. and O'Brien, John R. "Controlling Preferences for Lotteries on Units of Experimental Exchange." *Quarterly Journal of Economics*, May 1986, *101*(2), pp. 281–306.

Camerer, Colin F. "Do Biases in Probability Judgment Exist in Markets? Experimental Evidence." *American Economic Review*, December 1987, *77*(5), pp. 981–97.

______. "Individual Decision Making," in John H. Kagel and Alvin E. Roth, eds., *Handbook of experimental economics*. Princeton, NJ: Princeton University Press, 1995, pp. 587–703.

Dosi, Giovanni and Lovallo, Dan. "Rational Entrepreneurs or Optimistic Martyrs? Some Considerations on Technological Regimes, Corporate Entries, and the Evolutionary Role of Decision Biases," in R. Garud, P. Nayyar, and Zur Shapira, eds., *Technological innovation: Oversights and foresights.* Cambridge: Cambridge University Press, 1997, pp. 41–70.

Dunne, Timothy; Roberts, Mark J. and Samuelson, Larry. "Patterns of Firm Entry and Exit U.S. Manufacturing Industries." *Rand Journal of Economics*, Winter 1988, *19*(4), pp. 495–515.

______. "Firm Entry and Postentry Performance in the U.S. Chemical Industries." *Journal of Law and Economics*, October 1989a, *32*(2), pp. 233–71.

______. "The Growth and Failure of U.S. Manufacturing Plants." *Quarterly Journal of Economics*, November 1989b, pp. 671–98.

Dunning, David; Meyerowitz, Judith A. and Holzberg, Amy D. "Ambiguity and Self-Evaluation: The Role of Idiosyncratic Trait Definitions in Self-Serving Appraisals of Ability." *Journal of Personality and Social Psychology*, December 1989, *57*(6), pp. 1082–90.

Geroski, Paul A. "Some Data-Driven Reflections on the Entry Process," in Paul A. Geroski and J. Schwalbach, eds., *Entry and market contestability: An international comparison.* Oxford: Basil Blackwell, 1991, pp. 282–96.

Kahneman, Daniel. "Experimental Economics: A Psychological Perspective," in R. Tietz, Wulf Albers, and Reinhard Selten, eds., *Bounded rational behavior in experimental games and markets*. New York: Springer-Verlag, 1988, pp. 11–18.

Kahneman, Daniel and Lovallo, Dan. "Timid Choices and Bold Forecasts: A Cognitive Perspective on Risk Taking." *Management Science*, January 1993, *39*(1), pp. 17–31.

Kahneman, Daniel and Tversky, Amos. "Intuitive Prediction: Biases and Corrective Procedures," in S. Makridakis and S. C. Wheelwright, eds., *Forecasting*. TIMS Studies in Management Science, 1979, pp. 313–27.

***Los Angeles Times*.** "Going After the Big One." December 31, 1996, pp. F1, F8.

Lovallo, Dan. "Essays on the Psychology of Competition." Ph.D. dissertation, University of California, Berkeley, 1996.

Lovallo, Dan and Camerer, Colin F. "Overconfidence and Excess Entry: An Experimental Approach." Working Paper No. 975, California Institute of Technology, 1996.

Maccoby, Eleanor E. and Jacklin, Carol N. *The psychology of sex differences*. Stanford, CA: Stanford University Press, 1974.

MacCrimmon, Kenneth R. and Wehrung, Donald A. *Taking risks*. New York: Free Press, 1986.

March, James and Shapira, Zur. "Managerial Perspectives on Risk and Risk Taking." *Management Science*, November 1987, *33*(11), pp. 1404–18.

Miller, Robert A. "Job Matching and Occupational Choice." *Journal of Political Economy*, December 1984, *92*(6), pp. 1086–120.

Prasnikar, Vesna. "Binary Lottery Payoffs: The Degree of Controlling Risk Aversion." Working paper, University of Pittsburgh, 1996.

Rapoport, Amnon. "Individual Strategies in a Market-Entry Game." *Group Decision and*

Negotiation, March 1995, *4*(2), pp. 117–33.

Rapoport, Amnon; Seale, Darryl. A.; Erev, Ido and Sundali, James A. "Equilibrium Play in Large Group Market Entry Games." *Management Science*, January 1998a, *44*(1), pp. 129–41.

Rapoport, Amnon; Seale, Darryl A. and Ordóñez, Lisa. "Weighted Probabilities in Large Group Coordination: Experimental Evidence from Market Entry Games." Discussion paper, University of Arizona, Department of Management and Policy, 1998b.

Roll, Richard. "The Hubris Hypothesis of Corporate Takeovers." *Journal of Business*, April 1986, *59*(2), pp. 197–216.

Selten, Reinhard; Sadreih, Karim and Abbink, Klaus. "Money Does Not Induce Risk Neutral Behavior, but Binary Lotteries Do Even Worse." Discussion Paper No. B-343, University of Bonn, 1995.

Shapiro, Daniel and Khemani, R. S. "The Determinants of Entry and Exit Reconsidered." *International Journal of Industrial Organization*, March 1987, *5*(1), pp. 15–26.

Sundali, James; Rapoport, Amnon and Seale, Darryl A. "Coordination in Market Entry Games with Symmetric Players." *Organizational Behavior and Human Decision Processes*, November 1995, *64*(2), pp. 203–18.

Svenson, Ola. "Are We All Less Risky and More Skillful Than Our Fellow Drivers?" *Acta Psychologica*, 1981, *47*(2), pp. 143–48.

Taylor, Shelly E. and Brown, J. D. "Illusion and Well-Being: A Social Psychological Perspective on Mental Health." *Psychological Bulletin*, March 1988, *103*(2), pp. 193–210.

Tiger, Lionel. *Optimism: The biology of hope*. New York: Simon and Schuster, 1979.

Waldman, Michael. "Systematic Errors and the Theory of Natural Selection." *American Economic Review*, June 1994, *84*(3), pp. 482–97.

Weinstein, Neil D. "Unrealistic Optimism About Future Life Events." *Journal of Personality and Social Psychology*, November 1980, *39*(5), pp. 806–20.